DATA REDUCTION, RADIAL VELOCITIES AND STELLAR PARAMETERS FROM SPECTRA IN THE VERY LOW SIGNAL-TO-NOISE DOMAIN

LUCA MALAVOLTA

PhD Thesis

Scuola di Dottorato in Astronomia
Dipartimento di Fisica e Astronomia
Università degli Studi di Padova

Prof. Giampaolo Piotto

January 2013

Luca Malavolta: *Data Reduction, Radial Velocities and Stellar Parameters from Spectra in the Very Low Signal-to-Noise Domain,* PhD Thesis, © January 2013

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Fisica e Astronomia

SCUOLA DI DOTTORATO DI RICERCA IN ASTRONOMIA
INDIRIZZO UNICO
CICLO XXV

DATA REDUCTION, RADIAL VELOCITIES AND STELLAR PARAMETERS FROM SPECTRA IN THE VERY LOW SIGNAL-TO-NOISE DOMAIN

Direttore della Scuola: Ch.mo Prof. Giampaolo Piotto
Supervisore: Ch.mo Prof. Giampaolo Piotto

Dottorando: Luca Malavolta

ABSTRACT

Telescopi di grandi dimensioni usualmente rendono disponibili dei programmi per la riduzione dati che restituiscono all'astronomo dati già pronti per l'analisi scientifica, e sempre più spesso gli astronomi si appoggiano a questi programmi per evitare un lavoro lungo e difficile. I programmi di riduzione dati standard sono però progettati per restituire buoni risultati su dati con buon Rapporto Segnale Rumore (RSR), e spesso i problemi legati alla riduzione di dati a basso RSR non sono presi in considerazione, con il risultato che le informazioni che contengono non sono adeguatamente utilizzate.

Negli ultimi anni il nostro gruppo di ricerca ha collezionato migliaia di spettri osservati con lo strumento GIRAFFE collegato al *Very Large Telescope* dell'Osservatorio Europeo del Sud in Cile, con lo scopo di determinare la distanza geometrica e lo stato dinamico di diversi Ammassi Globulari Galattici, ma in definitiva l'analisi è stata ostacolata da errori sistematici nella riduzione e calibrazione dei dati e nella misura delle velocità radiali. Inoltre questi dati non sono mai stati utilizzati per determinare altre informazioni come temperatura e metallicità delle stesse, poiché considerati troppo rumorosi per questo tipo di analisi.

In questa tesi concentriamo la nostra attenzione sulla riduzione dati ed analisi di spettri con bassissimo RSR. Il set di dati che analizziamo in questa tesi è composto da 7250 spettri per 2771 stelle dell'ammasso globulare M 4 (NGC 6121) ottenute con GIRAFFE nell'intervallo spettrale $5145 - 5360$Å. Stelle della parte superiore del Ramo delle Giganti Rossi fino alla Sequenza Principale sono state osservate in condizioni molto differenti, comprese notti con luna piena, fino ad raggiungere un valore limite di RSR $\simeq 10$ per molti spettri.

La nostra analisi incomincia con i passi di base della riduzione dati ed estrazione degli spettri, adattando tecniche ben testate in altri campi (come la fotometria) ma ancora non ben sviluppate in spettroscopia. Continuiamo con il migliorare la soluzione della dispersione in lunghezza d'onda la correzione per piccoli spostamenti nelle velocità radiali di riferimento tra i dati di calibrazione presi durante il giorno e le osservazioni scientifiche seguendo un approccio completamente differente rispetto a quello ESO. Analizziamo poi la miglior maniera per effettuare la sottrazione del cielo e la normalizzazione del continuo, le due più importanti fonti rispettivamente di rumore ed errori sistematici nella misura delle velocità radiali nell'analisi chimica degli spettri. L'alto numero di spettri del nostro dataset richiede un approccio automatico ma robusto, che non manchiamo di fornire. Determiniamo infine per il nostro campione

di stelle velocità radial con una precisione mai vista in precedenza per dati di questo tipo e ritroviamo gli stessi parametri atmosferici di altri lavori svolti su stelle più brillanti, con dati a risoluzione spettrale maggiore e su intervalli di lunghezza d'onda dieci volte superiori a quello dei nostri dati.

Nell'ultimo capitolo della tesi affrontiamo una problematica simile ma da una prospettiva completamente differente. Spettri ad alta risoluzione e ad alto RSR ottenuti con lo spettrografo HARPS sono stati usati per calibrare i parametri atmosferici stellari in funzione delle caratteristiche di funzioni di cross-correlazione specificatamente costruite includendo linee spettrali con diversa sensibilità ai parametri atmosferici stellari. Questi strumenti sono stati progettati per essere facilmente implementati un programma di riduzione dati, pur tuttavia senza sacrificare l'accuratezza dei parametri determinati anche per spettri a basso Rapporto Segnale Rumore.

ABSTRACT

Large astronomical facilities usually provide data reduction pipeline designed to deliver ready-to-use scientific data, and too often astronomers are relying on this to avoid the most difficult part of an astronomer job Standard data reduction pipelines however are usually designed and tested to have good performance on data with average Signal to Noise Ratio (SNR) data, and the issues that are related with the reduction of data in the very low SNR domain are not taken int account properly. As a result, informations in data with low SNR are not optimally exploited.

During the last decade our group has collected thousands of spectra using the GIRAFFE spectrograph at Very Large Telescope (Chile) of the European Southern Observatory (ESO) to determine the geometrical distance and dynamical state of several Galactic Globular Clusters but ultimately the analysis has been hampered by systematics in data reduction, calibration and radial velocity measurements. Moreover these data has never been exploited to get other informations like temperature and metallicity of stars, because considered too noisy for these kind of analyses.

In this thesis we focus our attention on data reduction and analysis of spectra with very low SNR.

The dataset we analyze in this thesis comprises 7250 spectra for 2771 stars of the Globular Cluster M 4 (NGC 6121) in the wavelength region $5145 - 5360$Å obtained with GIRAFFE. Stars from the upper Red Giant Branch down to the Main Sequence have been observed in very different conditions, including nights close to full moon, and reaching SNR $\simeq 10$ for many spectra in the dataset.

We will first review the basic steps of data reduction and spectral extraction, adapting techniques well tested in other field (like photometry) but still under-developed in spectroscopy. We improve the wavelength dispersion solution and the correction of radial velocity shift between day-time calibrations and science observations by following a completely different approach with respect to the ESO pipeline. We then analyze deeply the best way to perform sky subtraction and continuum normalization, the most important sources respectively of noise and systematics in radial velocity determination and chemical analysis of spectra. The huge number of spectra of our dataset requires an automatic but robust approach, which we do not fail to provide. We finally determine radial velocities for the stars in the sample with unprecedented precision with respect to previous works with similar data and we recover the same stellar atmosphere parameters of other studies performed on the same cluster but on

brighter stars, with higher spectral resolution and wavelength range ten times larger than our data.

In the final chapter of the thesis we face a similar problem but from a completely different perspective. High resolution, high SNR data from the High Accuracy Radial Velocity Planet Searcher spectrograph (HARPS) in La Silla (Chile) have been used to calibrate the atmospheric stellar parameters as functions of the main characteristics of Cross-Correlation Functions, specifically built by including spectral lines with different sensitivity to stellar atmosphere parameters. These tools has been designed to be quick and to be easy to implement in a instrument pipeline for a real-time determination, nevertheless they provide accurate parameters even for lower SNR spectra.

CONTENTS

LIST OF TABLES

1 INTRODUCTION

1.1 HISTORICAL INTRODUCTION

Isaac Newton (1643-1727) was the first to realize that the light of the Sun could be dispersed into its components with the help of the prism, in the effort to understand how the rain drops split the white sunlight into its colors.

Atomic absorption lines in the solar spectra were classified and characterized for the first time by Joseph von Fraunhofer (1787-1826) in 1817. The terrestrial nature of these lines was excluded a few years later when he found that some lines are absent in the Sun and present in Sirio's and other very bright spectra, by the use of a diffraction grating.

The real nature of spectral lines was finally uncovered by Gustav Kirchhoff (1824-1887) and Robert Bunsen, who realized that a cool gas in front of hot source is absorbing light at specific wavelengths, dependent on the chemical composition of the gas.

In 1886 the observation of lines in the solar spectrum that had never been observed in laboratory led Sir Norman Lockyer (1836-1920) to the introduction of an unknown element, Helium, twenty-five years before the first detection on Earth.

Since then spectroscopy has influenced almost every field of Astronomy. When an object is moving in the radial direction with respect to the observer, the spectra is subjected to a shift in wavelength that is easily determinable with the Restricted Theory of Relativity. This behavior, independent of the physical properties of the observed object, has been used to derive the kinematical properties of neighbor stars as well as to derive the Hubble's Law (further galaxies are receding more rapidly) and to understand the real nature of Quasars, i. e. galaxies which spectrum has been highly shifted because of their high distance (hence velocity). In the last decade, an increasing precision of the radial velocity measurements thanks to innovations in design and construction of astronomical spectrographs has led to the discovery of planets around other stars. Spectral lines change their depth and shape according to the physical properties of the atmospheres of the stars, or *photospheres*, and the amount of the chemical elements that are forming the lines. The study of properties derived from spectral lines has shed light on the chemical evolution and ultimately the formation history of our Galaxy and the Local Group. Spectra of distant galaxies, whose single stars are too faint to be observed individually, can be investigated by comparison with the weighted sum of spectra

from stars with known characteristics. Spectroscopy plays a fundamental role in the study of the chemical composition of asteroids, comets and planets in our Solar System and around other stars.

This thesis focuses on one of the most recurrent use of spectroscopy in Astronomy: the determination of radial velocities, physical properties and chemical composition of stars in our Galaxy.

1.2 PUSHING TOWARDS THE LOW SIGNAL TO NOISE RATIO LIMIT

Spectroscopy has always suffered an intrinsic disadvantage with respect to photometry. A photometric band collects light from a wavelength range that is far more extended than a single resolution element of a spectrum: the V band in the photometric system of Johnson and Morgan [1953] has a Full Width at Half Maximum (FWHM) of 880Å, while a spectrograph with resolving power of $R \equiv \Delta\lambda/\lambda = 1000$ can resolve features that are $\simeq$ 1Å apart at the same wavelength (5400Å). Low resolution ($R < 1000$) is preferred when studying the overall flux distribution or very broad features like molecular bands; accurate determination of chemical abundances requires a higher resolving power to separate a sufficient number of lines and perform Equivalent Width analysis, while high precision spectrographs for exoplanets detection usually work at $R \simeq 100000$. Ultimately spectra at $R > 200000$ have been taken for the brightest stars in the sky (including the Sun) to study the shape of spectral lines and to build atlas of spectral features. The loss of light between the entrance of the spectrograph in the focal plane of the telescope and the detector contributes to lower significantly the limiting magnitude of a spectrograph as R increases.

It appears clear that many objects that are easily observable by imaging are close to the limit of observability in spectroscopy. For example, obtaining the Color Magnitude Diagram (CMD) down to the Main Sequence (MS) of a Galactic Globular Cluster (GCC) is easily feasible with a mid-size telescope like the 2.2-m MPG/ESO at La Silla, while a single spectra at medium resolution of a MS Turn-Off (TO) star of the same cluster requires a 8-meter telescope and several hours of exposure to reach the minimum Signal to Noise Ratio (SNR) required for chemical analysis.

With larger facilities like the European Extremely Large Telescope or the James Webb Space Teelscope still under development, the Astronomy community has to rely on current facilites to observe faint objects that show anomalous or unexpected behavior and need spectral analysis to be understood. In many cases pipelines and analysis technique have been developed regardless of the spectral properties of the observed objects, and tested only on bright objects, so that the direct application to faint targets does not fully exploit the available informations in the data. Indeed, in many instances astronomers will

needlessly discard spectra whose low SNR hides a lot of very useful data on stellar properties that standard tools are not able to read correctly.

In this thesis we focus our attention on spectra low SNR gathered with the GIRAFFE spectrograph, part of the Fibre Large Array Multi Element Spectrograph (FLAMES), mounted at the Very Large Telescope. During the last decade several thousands of spectra of stars of Galactic Globular Clusters (GGC) have been collected with this instrument, mostly by our group, to understand the internal cinematic of these objects and to determine their geometrical distance by comparing their internal velocity dispersion with the internal proper motions of the stars. Previous studies from [Milone et al., 2006] and [Sommariva et al., 2009] have shown that the radial velocities obtained using the ESO Giraffe standard reduction [Blecha et al., 2000] and the Ancillary Data Analysis Software [Royer et al., 2002] are affected by several systematics. Such systematics affects also the comparison of noisy spectra with stellar synthesis, since weak spectral lines are not found where they are expected to be.

We have at our disposal a dataset of 7250 spectra in a wavelength range of 214Å and centered at 5250Å, for 2771 photometrically-selected stars of the Galactic Globular Cluster M 4 (NGC 6121), from the Upper Red Giant Branch down to the Main Sequence of the Cluster. Despite the collecting area of a 8-meter telescope in one of the best astronomical site in the world and the use of a middle-resolution spectrograph ($\Delta\lambda/\lambda = 24000$) to carry out observations, most of the spectra have SNR $\simeq 10$.

Standard analysis techniques has not been designed, and probably never tested, to work at this level of noise. We have developed and tested a series of tools designed to deal with very low SNR data, with two main goals in mind:

- the determination of high-precision, systematic-free radial velocities;
- extraction of stellar atmospheric parameters and chemical abundance.

Data reduction and analysis play a fundamental role in reaching these objectives. We give a brief overview of the main issue we will encounter in this thesis.

Spectral extraction from very low SNR *spectra*

The classical algorithm for spectral extraction adopted by the majority of the pipelines is the Optimal Extraction by Horne [1986]. This algorithm is based on two assumptions: the line profile in cross-dispersion direction is well known, and the position of the peak flux in the spatial direction of the *science* spectra is known with a good precision.

The combination of these two pieces of information is used to give to each pixel a weight based on the fact that more information on the spectral flux is contained near the peak of the line profile.

In fiber-fed spectrographs however the shape is usually approximated with a more or less simple analytic function, which coefficients are derived by fitting this function to the line profile of a continuum lamp spectra.

Analytic models are not able to reproduce shape asymmetries, unless very complicated models are used. Model parameters are determined by minimization of the sum of the squares of model-observations residuals weighted for the SNR of each data point (χ^2 minimization), so an asymmetry in the line profile that has not been taken into account in the model will alter the model globally.

The result is that the shape is not accurately described by the model and the extracted flux will be afflicted by a higher noise, caused simply by the use of inaccurate weights in the optimal extraction.

Another issue is given by the position of the spectra in the CCD. The position of the science spectra in the cross-dispersion direction is assumed to be the same measured for the continuum lamp spectra. Continuum lamp calibration frames are usually taken several hours apart, at the beginning or the end of the night, and if the spectrograph is not thermally isolated and in vacuum, internal variation in pressure and temperature of the air results in a movement of the position of the spectra. As for the shape of the line profile, assuming an inaccurate position for the center of the spectrum results in a noisier extracted spectrum.

Similar problems have been already encountered in photometry, but successfully accounted for. The Empirical Point Spread Function approach described for the first time in Anderson and King [2000] has been adapted to the spectroscopic case to determine accurate shape and fibre position along the cross-dispersion direction in the frame. With this approach, the derived cross-dispersion profile does not rely on any analytic assumption and intra-pixel variations are automatically taken into account. Being developed for astrometry, this technique ensures a very precise determination of the fibre position.

Sky subtraction

One of the major drawbacks of echelle spectrograph is the impossibility of measuring the sky flux separately from the target spectra during the exposure.

Clear sky can reach a surface brightness of $\simeq$ 17.5mag/arcsec^2 at 30° from the Full Moon, so it usually does not represent a problem for bright and moderately faint objects if observations are properly scheduled. Unfortunately it is not always possible to choice the preferred time to carry out observations, and often the astronomer has

to deal with data where the sky flux contributes consistently to the observed flux.

In fiber-fed spectrograph the problem is partially solved by dedicating a few fibers to sky observations. When available, sky spectra are co-added together and the resulting spectrum is subtracted from the science spectra. This correction does not come for free: every fibre has a different transmission efficiency, and the error on its determination enters in both co-addition of sky spectra and more heavily when the resulting spectrum has to be subtracted to the science one. Even after co-addition, sky spectra are intrinsically very noisy and their noise will be unavoidably added to the science spectra. Finally, each fiber has a different wavelength dispersion solution (given by the geometry of the instrument), so the individual sky spectra must be rebinned a first time to be co-added, and a second time to match the wavelength scale of the object (or alternatively the object spectra and the sky one must be rebinned at the same wavelength scale). Rebinning a low SNR spectrum is a complex task since it can introduce correlated noise.

We propose two complementary approaches to remove the contribute of the sky from our spectra without sacrificing their quality. The first one is to determine the fraction of sky flux with respect to the stellar one by comparing the numerical Cross-Correlation function of the observed spectrum to the combination of the CCF of a synthetic spectrum with atmosphere parameters close to the ones of the observed object and the CCF of a solar spectrum, with the reasonable assumption that the Moon-illuminated sky spectra resembles the Solar spectra. The second one is to fit a combination of solar spectrum and synthesis directly to the observed spectrum, using the spectral lines of the sky as proxy for the sky contribution.

If the observed target has a radial velocity close to zero (before the barycentric correction), the fraction of sky flux can be determined by integrating the observed sky spectra over several Ångstroms and rescaling the solar spectra accordingly.

In both cases, the two model spectra (solar spectrum and synthesis) are rebinned into the wavelength scale of the observed spectrum and instrumental effects are applied, leaving the science spectrum untouched and thus without decreasing its SNR.

Continuum determination

The determination of the continuum flux of a spectrum plays a fundamental role in atmosphere parameters determination and chemical analysis, in either Equivalent Width measurements and synthesis matching. Assuming a lower continuum then the real value for example will lead to systematically underestimated abundances of a given element, even if the equivalent widths are precisely measured.

Continuum determination is usually performed only once before the chemical analysis. The most common method is to determine a first guess of the continuum function by fitting a few pre-selected regions free of lines, with some human interaction to exclude evident outliers yet included in the fit.

From this point of view, the data at our disposal have revealed to be very challenging. With several thousand of spectra to be analyzed, an interactive procedure had to be excluded.

If the spectra cover a very small wavelength range, as it is usual for individual orders in an echelle spectrograph, the continuum regions may be too short and not distributed across the whole range. Not only our spectra extend over 214Å only, but they are also characterized by strong features near the blue edge of the spectra. Moreover in low SNR spectra it is very difficult to discriminate faint lines from the noise, and an automatic continuum determination performed by taking all the flux values over a given threshold will result in a biased determination towards lower continuum fluxes.

Selecting the regions that exceed a given threshold in flux from a spectral synthesis and rescaling the normalization factor by the loss of flux computed using the synthetic spectrum may seem the best way to go, but in doing so we would introduce a dependence on the atmospheric stellar parameters used to create the synthesis itself, since the selected points would change with the synthesis parameters. This bias is avoided by determining the continuum each time that the observed spectrum is compared with a synthesis, using the same synthesis to select the regions to be used for continuum determination. This is feasible thanks to the automatic nature of our approach.

Radial velocities and photosphere parameters at low SNR

Spectral extraction, sky subtraction and continuum determination are finally put together to determine the radial velocities and atmosphere stellar parameters of our stars. We have decided to follow a different approach for the determination of the wavelength dispersion solution and for the simultaneous wavelength-calibration fibre (used to correct for radial velocity shifts in the data between the day-time calibration and the science observations), resulting effectively in an improvement of the precision respect to the ESO pipeline. We also tested the most common radial velocity measurements technique to determine which is the optimal one for short wavelength range spectra in the low SNR regime. Regarding to the stellar atmosphere parameters determination, the great attention in data reduction and analysis has been paired with a careful selection of atomic lines parameters and model atmospheres. In order to avoid the introduction of any false correlations into the spectrum noise, all the steps that usually lead to sky-corrected, rest-frame, continuum-normalized stellar spectra have

been performed in the opposite direction: deliberate degredation of the *perfect* synthetic spectra to best match the real extracted spectra. In this way the (already low) SNR of our spectra has not been affected by further reprocessing.

1.3 FAST ATMOSPHERE PARAMETERS DETERMINATION

When observing faint objects, it is common practice to split the total time of observation in several exposures. If the goal of such observations is to obtain a time series of a given observable, e. g. radial velocities for faint companions or exoplanets search, this splitting in deliberate. In the case of exoplanet search the situation is even more extreme: bright targets are observed in many short exposures spanned over years, individual spectra are very noisy and it is possible to determine the atmospheric parameters of the observed stars only several months after the first observations.

For such time-series spectroscopic data it may be useful to have a tool that is capable of deriving the fundamental atmosphere parameters rapidly and accurately even after a single exposure. If the derived parameters differ significantly from the ones used to select the targets of a survey, for example, the target can be dropped almost immediately and the saved time can be spent over other targets, thus optimizing the scientific outcome of the observations. This is a very likely situation when using very popular instruments, i. e. when the number of requested nights exceeds the available time.

For this purpose, a new technique based on a set of Cross-Correlation Functions built using lines more sensitive to a given parameter have been developed. Temperature, Gravity and Metallicity are calibrated using the parameters of several CCFs, so that the implementation on a pipeline only requires the inclusion of the used masks and the coefficients of the calibration.

In developing this technique we take advantage of the extensive database of spectra collected with the High Accuracy Radial Velocity Planet Searcher (HARPS) spectrograph. The database at our disposal comprises 1111 solar-type stars from several HARPS long-term programs [Mayor et al., 2003, Lo Curto et al., 2010, Santos et al., 2011] whose atmospheric parameters have been already determined using equivalent widths measurements [Adibekyan et al., 2012]. These stars are used to link the CCFs characteristics with the photosphere parameters and build a calibration to be used with spectra at lower SNR, where equivalent width measurements are not reliable anymore but the CCFs have still good signal.

1.4 STRUCTURE OF THE THESIS

To solve the problems described above we follow a pragmatic approach, where the described techniques are developed and tested on cases in which the spectroscopic data have been previously analyzed and described in the published literature. This allows us to compare our new spectroscopic manipulations with the old ones, to test whether we have really improved the information content of the extracted data.

In Chapter 2 the application of the empirical PSF is applied to spectroscopy in conjunction with the Optimal extraction method. Additionally, we reconsider from basic principles the often-neglected details of basic data reduction techniques.

Chapter 3 deals with the wavelength calibration of not-stabilized fiber-fed spectrograph. Three measurement techniques for radial velocity determination, namely *Classical CCF* in Fourier space, CORAVEL-type *Binary Mask* and *Template matching*, are contrasted to show their behavior in the low SNR regime and to derive the dynamical properties of the cluster.

In Chapter 4 our approach for sky subtraction and continuum normalization is described in detail. Stellar atmospheric parameters are determined for the stars in our sample by comparison with a grid of synthetic spectra created for this purpose. The general idea of applying every correction to synthetic spectra and leaving the observed ones untouched to preserve their SNR is strictly followed here.

In Chapter 5 we use HARPS-GTO data sample to derive and test a stellar parameter determination technique based on CCFs, designed to be fast and easy to be implemented within existing pipelines.

2

EMPIRICAL SPECTRAL EXTRACTION.

Standard data reduction pipelines are usually designed and tested to have good performance on average SNR data, neglecting the issues that are showing up when moving to lower SNR domains. Their black-box nature allows only a minimal fine-tuning, so the difficulties in the analysis of noisy spectra are worsened by a sub-optimal data reduction. In the following sections, we will take a careful look to often-neglected details of basic data reduction techniques, like dark master determination from a reduced number of frames or correction for pixel-to-pixel gain variations, and we will provide several ad-hoc solutions. We then apply to spectroscopy the well-known technique of the *empirical PSF*, which has already delivered high-precision astrometry and photometry of stars in Globular Clusters [Anderson et al., 2006, Piotto et al., 2007].

2.1 INSTRUMENT DESCRIPTION

The Fibre Large Array Multi Element Spectrograph (FLAMES) is a multi-object, intermediate and high resolution spectrograph mounted at the Nasmyth A platform of Kueyen (UT2) telescope of VLT. The interface with the telescope focus is accomplished by a Fibre Positioner where the fibres can be moved to access a corrected field of view with a diameter of 25 arcmin. The available fibres are linked to two instruments: 8 single fibres go to the UVES spectrograph (Red arm), while three different types of feeding fibres systems are connected to GIRAFFE. the latter is a medium-high spectrograph ($R \simeq 5600 - 39000$ depending on the setting) for the entire visible range, with two gratings and several filters available, but only one order or a fraction of it can be observed at once. The three fibre system feeding GIRAFFE are named IFU, ARGUS and MEDUSA. The first two are integral field units, the third one allows one to observe up to 132 objects (including fibres dedicated to the sky) simultaneously, plus five fibres for simultaneous Thorium-Argon calibration spectra.

In the next section the data pre-reductions steps and a new spectral extraction algorithm are presented. Spectral extraction has been tested on MEDUSA data, in particular spectra obtained with HR9 settings, while the pre-reduction steps are common to the three fibre-feeding system since they have the CCD in common.

2.2 BASIC STEP IN DATA REDUCTION

We briefly review the artifacts that affect CCD data acquisition and how we dealt with them in our analysis.

2.2.1 *Read-Out Noise*

During the CCD readout the amount of charge stored in each pixel is amplified and then passed to the Analog to Digital Converter to be converted to an integer number of electrons. Beside the discretization error, usually $\pm 0.5\ e^{*}-$, this process usually introduces an error called *Read-Out Noise* (RON) that is characteristic of the specific CCD, and it is usually expressed as a single mean value for the whole data frame.

The data values are expressed in Analog-to-Digital Units (ADU), also called Digital Numbers; the maximum value of the ADU is given by the electronics, e.g. a 16-bit ADC cannot measure describe more than $2^16 = 65536$ ADU, which is called the *digital saturation limit*. Since a single pixel can hold up to several hundreds of thousand of electrons, a conversion factor called *gain* is introduced in order to extend the lower and upper limits of the measurable number of electrons (i.e. to access a larger portion of the dynamic range of the CCD):

$$g = N_{e^-}/ADU$$

Increasing the gains allows one to increase the effective dynamical range of the sensor, but it will also increase the noise associated to a single measurements since error calculation must be performed using the physical values (i.e. ADU must be converted in electrons). For example the SNR in Poissonian regime is given by:

$$SNR = \frac{n_{e^-}}{\sqrt{f_{e^-}}} = \frac{n_{adu}}{\sqrt{g * f_{adu}}} \tag{1}$$

The read-noise should always be greater than the gain factor, in order to be resolved in ADU.

2.2.2 *Amplifier Bias level*

The amplifier bias is the zero level of the electronics, i.e. the value measured by the electronics when no photons are detected on the CCD. Usually, several columns of the data frame are dedicated to the measurements of this level, in a region called `pre-scan` if it is read before the transfer of the CCD counts, or `over-scan` if it is read after that the CCD data have been transferred. In some systems these regions are obscured parts of the CCD (in this case also the Dark Current is included), in other cases these regions are virtual, i.e. these zones are not constituted by physical pixels but are added by the

electronics at reading time (in this case, no information about the dark current is available).

Generally speaking the bias level is not constant with time, but it may depend mainly on the temperature of the CCD at exposure time and in second place by external factor e. g. variations of voltage in the electronics. Bias slightly varies in the opposite direction to the reading one, so a single value in general cannot be use for the whole frame but its trend can still be determined by pre/over-scan.

For each frame, we determined pre-scan and over-scan through two steps: all the values in a row are averaged with a $2.5-\sigma$ rejection algorithm (reading direction is horizontal); resulting values are interpolated with a low-order Chebyshev polynomial of the first kind. We opted for a Chebyshev polynomials instead of the classical ones since the orthogonal basis of the former provide a better fit to the irregular shape of the bias.

2.2.3 *Electronic (Read-Time) Bias*

Reading the full frame is not an instantaneous process and during this time interval voltage variations in the electronics may systematically introduce structures across the frame. These kinds of structures can be easily spotted in zero integration time exposures that are called *bias frame* `biasframe`. Since the S/N of a single bias (once the amplifier bias has been removed) is very low and fluctuation can be hidden by read-out noise, usually several bias frames are averaged together to get a *master bias*.

Differently from amplifier bias level, the electronic bias is not always present, e.g. can be absent if electronics have been designed very well. Although it is advisable to always obtain a master bias to check if such structures are present, a series of checks should be performed to decide if removing the bias frame is worthwhile:

1. the mean pixel count after over-scan correction should be zero
2. the count distribution function should be a Gaussian with width equal to the read-noise and symmetric respect to the center
3. a visual inspection of the master frame is always recommended

In our case, we have at our disposal five bias frames for each night, taken before the beginning of the observations. The master bias is obtained computing the median of the over-scan-corrected bias frames pixel-by-pixel, with the Root Mean Square (RMS) of the deviations from the median stored in the associated master error bias. An overall value $f_{\mathtt{bias}} \pm \sigma_{f_{\mathtt{bias}}}$ is obtained simply by computing the median of all the values in the master bias and the associated error frame; when a pixel value of the bias frames exceeds $3.5\sigma_{f_{\mathtt{bias}}}$ times the mean value, the pixel is flagged as hot one.

We analyzed all the master bias obtained in the period of our observations, including a few nights before and after the observations to check the stability of the ccd, and we saw the absence of any pattern or variation at any spatial frequency and confirmed the stability of the electronics with time for both the CCDs. We decided not to use the master frame in the subsequent steps of the data reduction, but we keep the information regarding the anomalous pixels and the read-out noise gathered in this step of the reduction process.

2.2.4 *Dark current*

This is an additive noise that comes from thermally generated electrons recaptured by the CDD itself, and in a modern CCD it is usually very low (few electrons per hour). Dark current is linear with time, so to correct for it the usual approach is to take several frames (for co-addition and cosmic rays removal) with the shutter closed, and using an exposure time adequate for the science program (usually the longest exposure time from the science ones). Since gathering the dark frames can be time consuming, especially for large facility where several science programs are running during the night, and the dark current is most of the time negligible, often the dark frame subtraction is skipped.

When a pixel produces a vastly higher rate of thermal electrons, it takes the name of *hot pixels*. If the production rate is excessive i. e. saturation is reached even with short exposure times,electrons can flow to the neighbor pixels in the same column . Correction for this effect is very difficult if a proper set of dark frames is not available, so we decided to simply flag these pixels as unusable.

The first CCD installed on GIRAFFE (named *Bruce* by the Eso Quality Control Group) was affected by a bright glowing area on its upper right edge. According to the GIRAFFE *CCD Detector Systems Detector Design and Performance Report*, the glowing at the CCD top section starts to be visible in exposures longer than 60s, and it is related to charge injection by the CCD guard ring. This charge injection is caused by a manufacturing defect on the CCD and cannot be removed by the CCD electronics. This effect is proportional to the exposure time and also exhibits a dependency on temperature, so a frame-by-frame correction is required. This glow is not always present (i.e., it disappeared from 072003 to 072004), so every scientific frame has to be visually inspected to determine if the dark correction is necessary or not. When necessary, the correction is performed also on the calibration frames, i.e. detector flat fields, fibre flat-fields and wavelength calibration frames.

ESO decided to gather only three dark frames roughly once in month with a fixed exposure time of 3600 seconds. A simple median like the one performed for the master bias determination is not

the best option due to the lower number of dark frames and the cosmic rays that unavoidably affect long exposures. A dark master is obtained in this way:

- Each frame is corrected for over-scan and rescaled to a reference exposure time (set as 3600 seconds since this is the most common value used for dark exposure)
- For each pixel, a median of the values from the frames and the RMS of the deviations is determined; the same values are then averaged using a 2.5σ rejection, using the median and the rms just determined, to obtain a master dark and the associated master error dark.

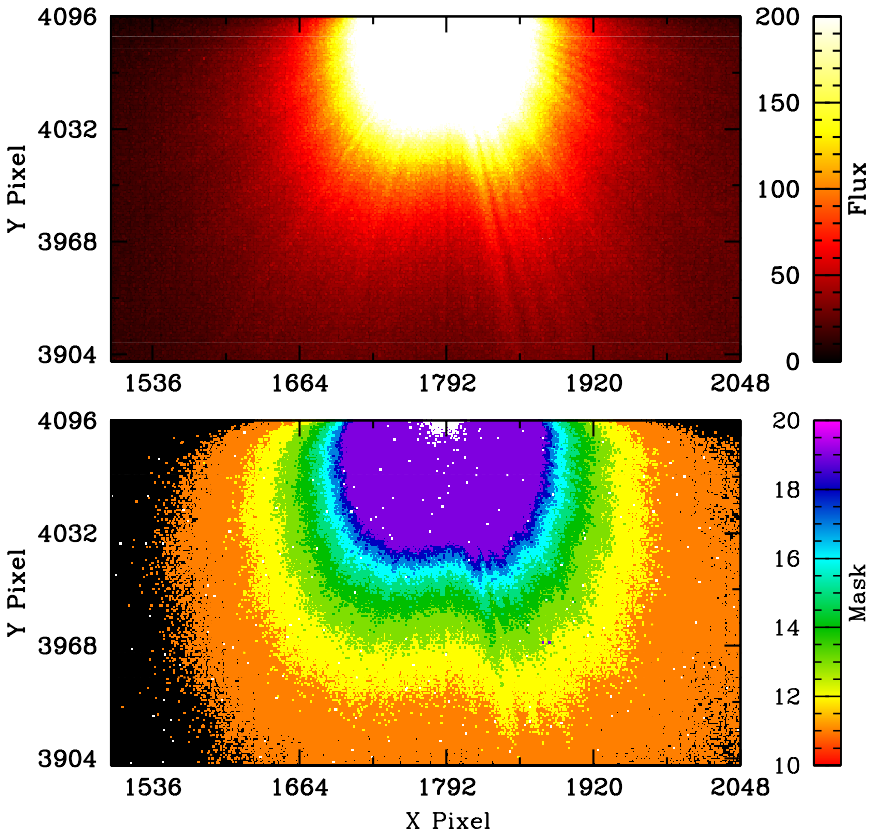

Figure 1: Upper panel: the bright glowing area on the upper right edge of the CCD in a 3600s dark frame exposure. Lower panel: each pixel is flagged with a number that tells how many times the flux in the master dark exceeds by five times σ the local mean.

Hot pixel masking is treated in a different way than done for the bias frame. Pixels that can be flagged as *hot* in 1 hour exposure can still have a dark current lower than the RON in shorter exposures, and we don't want to discard useful information. This is particularly true for our data, where the glow is affecting an important portion

of the frame (upper panel of Figure 1). We first divide the frame in several sections for local mean and σ determination, since we don't want the glow to affect the determination of these two values in *clean* zones. We then flag each pixel with a number, that tells how many times the flux in the master dark exceeds by five times σ the mean determined in the corresponding quadrant (lower panel of Figure 1) The hot pixel mask is rescaled according to the exposure time of the frame to be corrected.

Since dark frames are taken only once every month, we have decided to interpolate for intermediate period. We made a conservative choice to retain as flagged the pixel that were flagged only in one of the two epochs of interpolation. After dark correction, sometimes a residual of the glow is still visible in science frames, probably due to its dependency with time and temperature of the CCD. We will deal with this source of stray light before spectral extraction.

Last observation with Bruce was taken the 13^{th} of March 2008 before being replaced by a new CCD (called *Carreras*), in operation since the 26^{th} of May of the same year. The new CCD does not show any defect in the dark frames, and the correction for dark master is no longer required.

2.2.5 *Pixel Gain Variation*

In a perfect system, the number of measured electrons on the CCD would be directly proportional to the number of photons that are supposed to be collected by the pixel in a given position for a given source. In reality, there can be unpredicted loss of light both due to interferences in the optical path and to anomalous response of the electronics. Going from low spatial frequencies to higher ones, these are the most common situations encountered in spectroscopy:

- dust particles or defects in the optical path that cause discernible patterns on the data frame;
- fringing problems, i.e. wavelength dependent interference that creates global interference patterns, mostly visible at infrared wavelength;
- differences in the Quantum Efficiencies (QE, efficiency of conversion of incident photons into electrons) from pixel to pixel; the most recognizable cases are hot and bad pixels.
- quantum efficiency hysteresis, where the QE of a pixel depends upon its exposure history; very rare;

When working with echelle spectrograph, flat field correction is usually performed after the spectral extraction: both science and flat

field spectra are extracted and then the latter is divided by the former. This is the approach followed by the ESO Pipeline, for example. The idea is that the fraction of light lost due to imperfections on the CCD is the same in both spectra, but this assumption does not hold in all situations: in particular if there is even a slight difference in the central position of the calibration and science spectra in the cross-dispersion direction, the high-frequency variations will act in different way in the two spectra.

To correct for high-frequency variation in the spatial domain we tried a different approach: every month the ESO team takes a frame with the CCD completely illuminated by a the flat field lamp, these frames are called detector flat-field to distinguish them from the one obtained illuminating the fibres at the focal plane of the telescope. We assume that the difference in the QE between two adjacent pixels is constant with wavelength, i.e. the quantum efficiency dependency with wavelength of adjacent pixels follow similar curves, and that the amount of light received in the Detector Flat Field varies smoothly across several pixels, allowing us to determine a mean local value to be used as reference for pixel normalization.

In the following we describe the adopted procedure to obtain a pixel-to-pixel gain variation correction frame. Frames with exposure times less than 15 seconds have been rejected to avoid very low SNR cases, and those with counts exceeding 50K ADU have been rejected to avoid departures from linearity.

DATA PREPARATION Each frame is corrected for over-scan and dark if affected by the glow, and then normalized to median determined in a central box of 600 per 600 pixels. This is the best method for rescaling, since the lamp warms itself in a strange and unpredictable way, even over a homogeneous set of frames with the same exposure times.

DETERMINATION OF THE EXPECTED MEAN PIXEL VALUE We need to determine the expected mean value for a pixel in a given part of the frame, since we are also interest in determining the pixel-to-pixel variations in sensitivy. In other word we want to remove high spatial frequency variations.

- normalized frames are weight-averaged together, with two precautions:
  - flagged pixels are excluded from the average
  - if the number of available "good" pixels is less than 3, then also the bad pixels are used but the resulting pixel will be flagged. Hot pixels are now easily identifiable because the have normalized values larger the one (remember that we are using short exposures, so only the hottest pixel are identifiable)

- using the obtained frame, a local mean for each pixel is calculated. The median and dispersion are computed inside a box of size sample_size $= 8$, excluding hot pixels in a first run and re-including them if the previous run is unsuccessful.
- The error from the weighted mean is substituted with the σ, defined as the 68.27^{th} percentile of the distribution, of the surrounding pixels (closer than 3 pixels) around the median value determined in the previous step.
- we estimate again the local median, this time using a 2.5σ rejection algorithm. In some cases the clipping condition is relaxed until a sufficient number of pixel (at least 5) are collected for the median determination. This allows us to determine the mean value accurately even where the influence of the glow in the upper right edge of the CCD is strong.

DETERMINATION OF THE DETECTOR MASTER FLAT FRAME Now that the average expected value is known for each pixel, the normalized frames are reprocessed in order to determine a more accurate Master Frame. Every pixel is flagged accordingly to its deviation from the expected value by following this criteria, where $f_n^k(i,j)$ is the normalized value for the (i,j) pixel in the k-th frame, $f_l(i,j), \sigma_l(i,j)$ are respectively the local mean and its deviation determined in the previous step for the considered pixel, and mask is an array with initial values set to zero:

- if $f_n(i,j) < f_l(i,j) - 5 * \sigma_l(i,j)$ then $\text{mask}(i,j) = \text{mask}(i,j) - 1$
- if $f_n(i,j) > f_l(i,j) + 5 * \sigma_l(i,j)$ then $\text{mask}(i,j) = \text{mask}(i,j) + 2$

Such a high threshold is chosen to avoid an excessive masking of the frame when the σ is underestimated. The resulting masks for several sets have been verified by visual inspection, and all the pixels that show anomalous behavior respect to their neighbors are correctly reported. The difference in the assigned value between low and high pixels allows to disentangle between effects moving in opposite directions

After analyzing all the n frames, the mask array will contain all the information that we need to distinguish hot/bad pixels from the other:

$mask(i,j) = 0$	regular behavior in all the used frames;
$mask(i,j) = -n$	dead or low count pixel;
$mask(i,j) = 2 * n$	hot pixel;
$0 < mask(i,j) < -1$	anomalous ;
$0 < mask(i,j) < 4$	cosmic ray in one frame;

In the last two cases, the value of the Detector Master Flat for the considered pixel is determined excluding the measurements affected by the cosmic rays, and the output entry in the mask is set again to zero.

DETERMINATION OF THE PIXEL-TO-PIXEL VARIATION FRAME This frame is simply the division between the Detector Master Flat Field and the Local Mean Frame.

The Detector Master Flat Field, the Local Mean Frame and Pixel-to-Pixel Variation Frame are shown respectively from left to right in Figure 2.

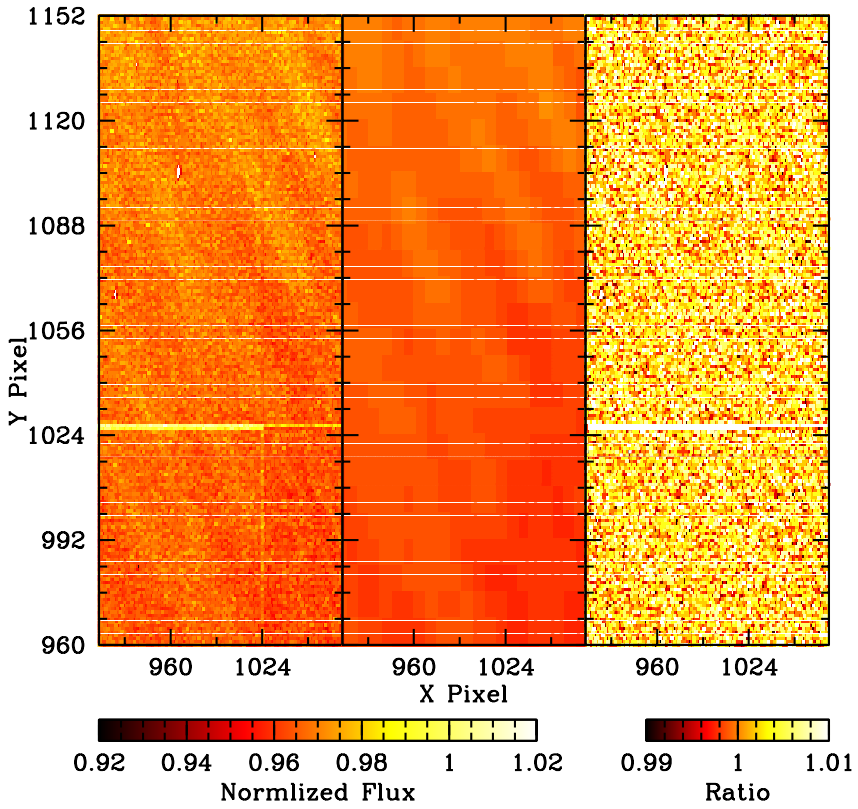

Figure 2: The Detector Master Flat Field, the Local Mean Frame and Pixel-to-Pixel Variation Frame are shown respectively from left to right

With the correction for pixel-gain variations, the corrections to be applied to calibration and science frames are completed. Fibre-Flat, Arc-Lamp and Science frames are ready for spectra extraction.

2.3 THE EMPIRICAL CROSS-DISPERSION SPREAD FUNCTION

The concept of *empirical Point Spread Function* (ePSF) was introduced first in Anderson and King [2000] with the name of *effective PSF* to

improve photometry and astrometry of stars on under-sampled CCD images from HUBBLE space telescope. The concept was extended to ground-based photometry and gained the actual name in Anderson et al. [2006].

The ePSF is simply a collection of values representing the expected fraction of light measured by a pixel as a function to the position of a star over the CCD.

The *pixel-phase* is defined as the fractional part of the position of the star in a given direction. The use of the word *phase* is justified by the fact that this value is always comprised between -0.5 and 0.5. The pixel containing the center of the line profile is taken as reference, so that when the center of star coincides with the center of the pixel, its pixel-phase is zero in each direction.

A stellar profile on the CCD is sampled in several points, all at one pixel of distance each other in the two directions. The value of the ePSF associated to a pixel is then given by the pixel phase plus the distance from the central pixel, i. e. the pixel containing the peak in the flux.

The ePSF approach is adapted to spectroscopy by reducing it to a mono-dimensional case, with the pixel-phase measured along the cross-dispersion direction.

2.3.1 *Advantages of the empirical approach*

As stated before, the empirical approach has been developed for astrometric purposes, although an improvement in the photometry has been obtained as well. Adopting this approach to the spectroscopy case has three main advantages:

- accurate position of the spectra in the CCD, at level of one hundredth of a pixel.
- intra-pixel sensitivity is automatically take in account, as much all the effect that could produce a variation of the cross-dispersion spread function (hereafter XSF) from a regular shape;
- no analytical model is needed: the empirical XSF can describe very well asymmetric and irregular shapes without additional effort.

In spectroscopy the actual position of the spectra in the CCD does not contain any useful scientific information, so one may be tempted to think that the empirical approach is not adding any information beyond existing implementations of optimal extraction algorithms. In reality, a better determination of the spectrum position results in an improved spectral extraction. There are some case where knowing the position of a fibre with a sub-pixel precision can be useful,e. g. to

determine fibre shifts between science frames and wavelength calibration ones, which might be taken also several hours after the observations. Also, this additional information can be used to analyze CCD defects like the recurring deformation in the pixel dimension due to the manufacturing process of the CCD [Murphy et al., 2012]

2.3.2 *Choosing the appropriate data to determine the eXSF*

In stellar photometry applications, the determination of the ePSF requires stringent condition on the stars that can be used to determine it. The stars must have a minimum number of counts in the central region, they must be isolated and in a sufficient number in each portion of the CCD for which the ePSF has to be determined; if not enough stars are present, spatial variability over the CCD must be sacrificed. This substantially explains why this technique has been so successful for Globular Clusters and dense Open Clusters.

When applying this technique to fibre-fed spectrographs these requirements become looser: such instruments are usually designed to have orders well spaced or at least with minimum cross-talk, and the Fibre Flat-Field (frames obtained illuminating the telescope in the observational setup with a feature-less lamp) have enough SNR to provide good sample points for the eXSF determination of each fibre.

2.3.3 *The recipe for a good eXSF*

Here we describe the process to determine an accurate eXSF for the considered dataset. While the values that we are providing are specific of the *HR9* GIRAFFE setting, the procedure is in principle applicable to any fibre-fed instrument and, with a little adaptation, also to *scrambled* slit echelle spectrograph. I took particular care in expressing all the possibile variable as external parameter, to easily adapt the code to different instruments.

As for the ePSF, the determination of the eXSF is an iterative process. We cannot derive a eXSF from a spectrum without prior knowledge of positions and fluxes of the slices, nor can we measure positions and fluxes without knowing the eXSF. In practice we will alternate between eXSF derivation and measurements, improving both with each iteration.

2.3.3.1 *Construction of the eXSF*

In this section we described the derivation of the XSF when the position and total flux of each *slice* is known.

Following Anderson and King [2000], we saved the eXSF as a sxsf_size $=$ 49 grid-points, with a subsampling factor $\texttt{textxsf\_subsample} = 3$ per pixel, which means a total of 7 pixels per side plus a centra pixel

are sampled Every pixel of each row in the cross-dispersion direction samples the eXSF at one point, provided that we know the cross-dispersion position and flux.

Obviously, all the values from a single row will sample the eXSF at a given pixel-phase, being all the samples at one pixel of distance from the next one. To properly determine the eXSF we need a set of sampling spanning different pixel-phases: while in photometry this is accomplished by using different stars in the field, by the fact that stars are not distributed in the sky following a geometrical pattern, in our case the position of the spectra in the cross-dispersion position will change smoothly with wavelength, due to the continuos nature of spectra. Fortunately echelle spectra are rarely straight line, due to the optical design of such instrument, so a good coverage of the pixel-phase is still guaranteed by taken enough rows of the spectra. We will come back to this matter in Section 2.3.4

The eXSF determination is an iterative process in which at each iteration the residuals between the sampling points and their actual expected values for the current model of the eXSF are determined. For each point of the grid, a weighted mean with 2.5σ rejection of the residuals is computed, and the model is adjusted according to the computed value. The input model is just a null model, while at every new determination of flux and position the previous model of the eXSF is used as starting point; a null model at each iteration works too, but it requires more computational time. Three constraints are applied to ensure a reliable determination of the eXSF:

SMOOTHNESS The shape of the XSF, although being unknown, is espected to be quite regular (e.g. without edges or sudden changes in the slope). This is true in general, with a few exception given by particular solutions used in the optics e.g. fragmented primary mirror, but it is especially true in our case since the light is coming out from the head of a fibre. We enforce this smoothing in the grid point of our eXSF using the Savitzky-Golay filtering, in the implementation of Press et al. [1992] instead of using a quadratic smoothing kernel like in the 2D case, since this solution is more suitable for 1D data.

CENTERING eXSF has to be centered on the center of the grid: to accomplish this, a parabola is fitted in the 7 grid points surrounding the peak; then the eXSF is resampled with cubic interpolation at the expected gridpoints locations. To avoid divergence we set a maximum offset of $1/4$ of pixel for each iteration.

NORMALIZATION Saturation is not expected to be a problem in a fibre-fed instrument since we are dealing only with selected targets, and we are not interested in measuring every source point in the field like in the case of photometry. The eXSF is normalized respect to the sum of the central 9 gridpoints with null

pixel-phase, and no effort is done to determine the shape of the eXSF when saturation occurs for the reasons cited above.

These three constraints are the main reasons why iterations are needed during the residual minimization process; after applying them, the model will be slightly different from the residual-corrected one, so a new iteration is needed in order to determine the best approximation of the true eXSF. Specifically, for each iteration except the very first one the centering is done before the eXFS adjustment, while the smoothing after it; normalization is performed at the end of the iterative cycle. Convergence is reached after a few iterations, we set a maximum limit of xsf_niter4 $= 7$ which ensures a good determination of the eXSF even with anomalous shapes (see Figure 3, where the shape is clearly notgaussian).

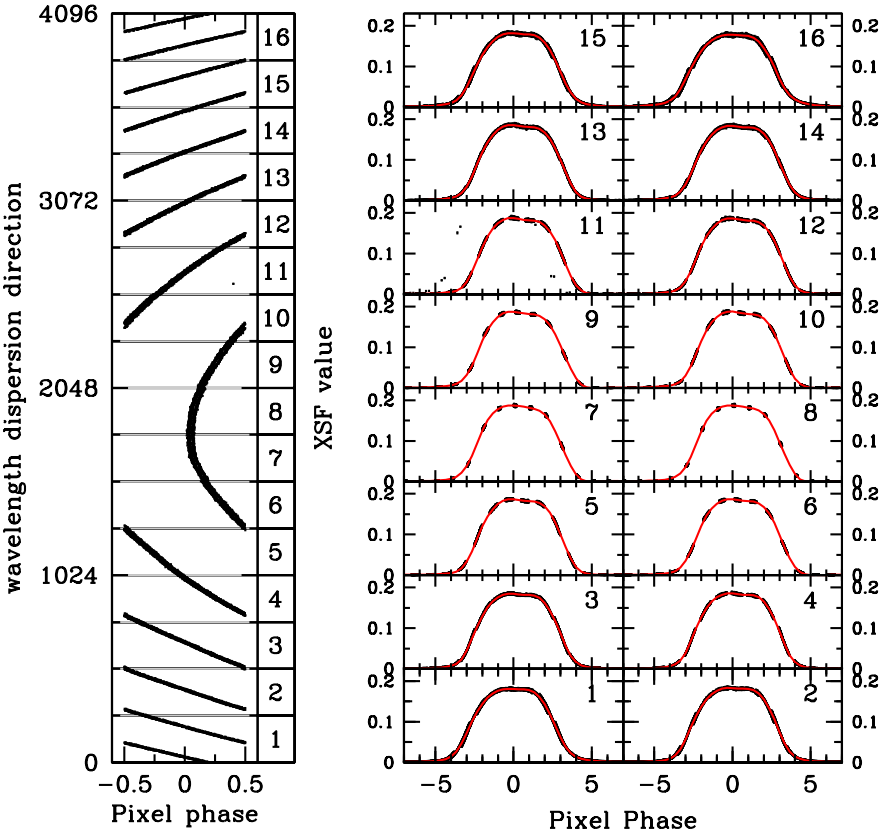

Figure 3: In the left panel, the pixel-phase of the selected fibre for each row of the CCD. In the right panel, the derived XSF (in red) and the sample points used for its derivation (in black) are displayed. The numbers identify the reference slice in the left panel.

Finally, we have to deal again with the iterative nature of the XSF, switching from flux and position determination, and shape improvement. Since we want a smooth convergence towards the best approximation of the eXSF, the output model that will be used to re-measure

fluxes and positions is actually a linear combination of the new determined model and the input one, with a ratio of 3 : 1. Greater robustness in the algorithm is then obtained sacrificing computational time.

2.3.3.2 *Measurement of Flux and Position*

Once we have a model for the Cross-Dispersion Spread Function, we can proceed to the determination of flux and position. Not differently from photometry, the determination of these two parameters is a least-squares process and the aim is to minimize this expression:

$$\chi^2(x_\lambda, f_\lambda) = \sum_i \frac{w_i \left[P_i - f_\lambda \psi_E(i - x_\lambda)\right]^2}{g P_i} \tag{2}$$

where the sum is over the assigned aperture of the fibre `fibre_width` and the symbols have the following meaning:

P_i	value of the *i*th pixel (in ADU),
$\psi_E(i - x_\lambda)$	$= \psi_i$, value of the eXSF at that offset
w_i	position weight function evaluated at i
g	ADU to electron conversion factor (gain)

The position weight function ensures that more importance is given to the central pixel (hence the name), where most of the information is contained:

$$\begin{aligned} w &= 1.0 \;\text{ if }\; |i - x_\lambda| \leqslant \text{xsf\_w1fit} \\ w &= 0.5 \;\text{ if }\; \text{xsf\_w1fit} < |i - x_\lambda| \geqslant \text{fibre\_width} \\ w &= 0.0 \;\text{ if }\; |i - x_\lambda| > \text{fibre\_width} \end{aligned} \tag{3}$$

The weight function is evaluated every time that a new determination of the center is available.

Since the position enters nonlinearly in Equation 2 through the eXSF, we solve alternatively for position and flux, using the new estimates at each iteration. Flux from aperture extraction and position from the fibre in the Fibre-Flat Field are good first approximations for a fast convergence of the algorithm. For a fixed position, the value of the flux that minimizes Equation 2 is easily solved (see equation:

$$f_\lambda = \sum_i \frac{w_i q_i P_i \psi_i}{w_i q_i \psi_i^2} \tag{4}$$

where q_i is an error weight function to take in account the error budget from data reduction and the poissonian error:

$q_i = P_i/(g\sigma_e^2 + P_i)$	first iteration
$q_i = 1.0/(g\sigma_e^2 + f_\lambda \psi_i)$	next iterations

The total variance of the flux is obtained following equation 8 in Table 1 of Horne [1986]:

$$\sigma_f^2 = \frac{\sum_i w_i \psi_o}{\sum_i w_i q_i \psi_i^2} \tag{5}$$

In the first iteration of the flux and position determination cycle the flux is still not very well constrained (i.e. a cosmic ray or a bad pixel not properly corrected can heavily influence the initial guess from aperture extraction), so the pixel are weighted according to their value. In the next iterations the flux is determined through the eXSF and such outliers are naturally excluded, so we can now use the expected value of the flux to model the Poissonian errors.

After a few iterations, when the value of the flux is converging, the expected value of the flux from each pixel is compared to the determined flux and the pixel get a zero weight if the difference from the two exceeds the uncertainty on the flux by 3 times:

$$|P_i - f_\lambda \psi_i| > 3\sigma_f \Longrightarrow w_i = 0 \tag{6}$$

After a value for flux has been determined, we solve Equation 2 for the position. The equation is not linear and cannot be solved analytically, since the eXSF itself is not available as an analytical expression. Basically our problem now is to determine the value of the position that minimizes the weighted sum of the square of the residuals $R_i = P_i - \psi_i$, and it can be easily accomplished with the Newton-Raphson method if the second derivative of the eXSF is available like in our case.

$$\begin{aligned} x_{\lambda,n+1} &= x_{\lambda,n} - \frac{\partial\chi^2(x_{\lambda,n}, f_\lambda)/\partial x_\lambda}{\partial^2\chi^2(x_{\lambda,n}, f_\lambda)/\partial x_\lambda^2} \\ &= x_{\lambda,n} + \frac{w_i q_i R_i \psi_i' f_\lambda}{w_i q_i f_\lambda \left[f_\lambda \psi_i'^2 - R_i \psi_i''\right]} \end{aligned} \tag{7}$$

The cycle is interrupted when the increment in the position is less then one part over a million of a pixel or a maximum number of `xsf_niter3` iterations have been performed.

The values of the pixel are intended to be already stray-light-subtracted.

2.3.4 *Chip Variability*

The PSF is not constant through the CCD and this is true also for spectrographs. In our implementation of the technique, the eXSF to be used in the science frames is determined using the Fibre Flat-Fields observed during the day, so no more than a few pixel of difference in the cross dispersion direction are expected. Along the dispersion direction, the eXSF is determined intervals of xsf_hslice $= 128\mathrm{pixel}$ of width, so a total of 32 eXSF are determined for each fibre; the espected XSF at any point is determined with a linear interpolation of the two closest eXSFs.

The number of rows to be used for the determination of a single eXSF is a compromise between the precision we want to reach in determining the eXSF itself and how good we want to trace the spatial

variability. To accurately determine the eXSF the sample points must span a good portion of the pixel-phase interval, i.e. the center of XSF must move inside a pixel along the dispersion axis. This is easily accomplished since the echelle spectra are never along straight lines (single rows), and a curvature is always present. The 128 pixel span has been chosen since it provides a good pixel-phase coverage even for the central fibre in the central part of the CCD, where the first derivative of the position of the fibres is at its minimum (see Figure 3).

2.3.5 *Fitting fibre position and flux*

When used in photometry, the ePSF technique must be coupled with a finding algorithm to look for stars, and the measurements must be iterated to resolve for faint or overlapping stars. In a fibre-fed spectrograph, the center of the fibre is known with an uncertainty of less than 1 pixel and the amount of cross-talking (the spectroscopic equivalent of overlapping in photometry) is already known since it is an instrumental design choice (i.e. spectra are not randomly distributed on the CCD). What really matters is to determine the amount of flux dispersed in the cross-dispersion direction; the advantage of a good determination of the center provides a better determination of the flux and can be useful to bound RV shifts with differences in postion between the ThAr Calibration Lamp and Science spectra.

2.4 COMPARISON WITH ESO PIPELINE

According to the GIRAFFE Pipeline User Manual, two optimal spectrum extraction methods beside the standard one (i.e. aperture) are offered. The first step, common to both extraction method, is to fit each cross-dispersion profile in the Fibre Flat-Fields with one of the following analytical expression (chosen by the user):

$$\begin{aligned} P(x) &= Ae^{-[(x-x_0)/\omega]^2} + B \\ P(x) &= Ae^{-|x-x_0|^\epsilon/\omega} + B \\ P(x) &= Ae^{-[|x-x_0|/\omega]^\epsilon} + B \end{aligned} \tag{8}$$

The profile parameters (A, B, ω, x_0) are then modeled using a Chebyshev polynomial as a function of the position in the cross-dispersion direction and stored for later use.

The user can choose between two extraction methods:

HORNE SPECTRUM EXTRACTION This is the implementation of the optimal extraction as described by Horne [1986], with the difference that the cross-dispersion profile is assumed to be one of the three given function in Equation 8. The method also applies a background subtraction.

OPTIMAL EXTRACTION This method extracts the fluxes of all the spectra simultaneously by fitting the following extraction model to the observed data in each wavelength bin x:

$$m(x,y) = B(x,y) + \sum_{j=0}^{N-1} f_j(c) P_j(x,y) \tag{9}$$

where $f_j(x)$ is the flux in x originated by the fibre j and $P_j(x,y)$ is the fibre profile, given by Equation 8 as in the Horne extraction. $B(x,y)$ is the background model, given by a Chebyshev polynomial of order l.

The second method has the advantage of treating correctly the cross-talk and at the same time the background determination. However, there are two drawbacks to this method. First, the parameters used for the analytical XSF are derived without taking into account the cross-talk, so artifacts in the XSF determination are propagated to the extracted spectra. Second, the scattered light correction is still necessary, although not required, since a low order polynomial may not be sufficient to treat the background in regions close to strong lines in the simultaneous ThAr spectra. The Pipeline User Manual also states that the second method is too slow for IFU and Argus mode, where the cross-talking is more accentuated than in the Medusa mode. Finally, both methods imply the knowledge of the cross-dispersion spread function: the provided analytical functions are not able to replicate asymmetric or anomalous shaped XSF, and the determination of position especially can be greatly influenced by this shortcoming.

In our treatment, spectra are extracted individually and no attempt to correct for cross-talking is done, since the fibres are well-spaced; the wings of the XSF can be contaminated by the surrounding fibres. This is the reason why we only use the central pixels to determine the flux, beyond being the ones with highest SNR. One of the advantages of the eXSF in fact is that the local shape around a given position is not influenced by the overall shape, i. e. cross-talking in the wings does not affect the pixel in the center, while the parameters of an analytical function are the result of a global fit.

Taking into account the cross-talk correction is feasible in principle, but to take advantage of the eXSF approach the Fibre Flat-Fields should be taken with adequate precautions to avoid contamination between the fibres used for the eEXSF, e. g. illuminating alternatively the odd and even fibres.

2.4.1 *Fibre positions*

With this technique we are able to determine the center of the fibre in the cross-dispersione direction with high accuracy. In Figure 4 the

measured position of the fibre in a single frame is interpolated as a function of the pixel position in the wavelength dispersion direction: the dispersion around the interpolated value is $\sigma = 0.0055 pixel$; when the values measured in the three Fibre-Flat frames are combined together, the dispersion is $\sigma = 0.0037 pixel$.

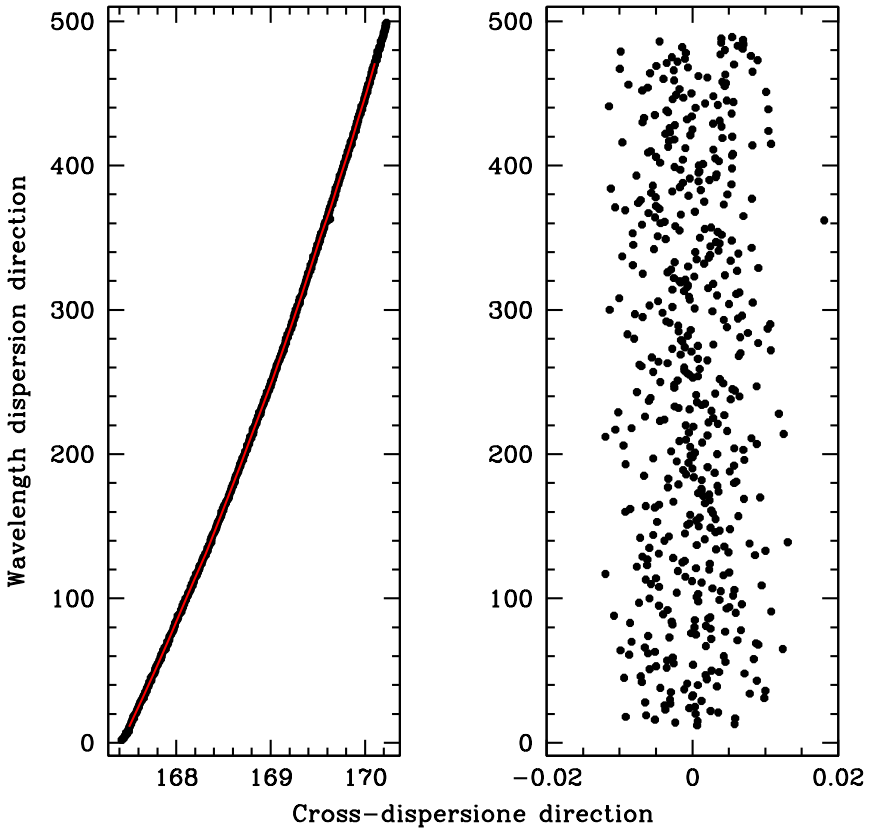

Figure 4: Precision of the fibre position measurement in a single Fibre-Flat frame. The right-hand panel show the difference between the measured position and a low-order polynomial that fits the data. The dispersion around the mean value is $\sigma = 0.0055 pixel$. When combining the measurements of three Fibre-Flats the dispersion goes down to $\sigma = 0.0037 pixel$.

A direct comparison with GIRAFFE data is not possible since the pipeline provides directly the result of the polynomial interpolation with the measured fibre position.

The precision reached in fibre position determination allows us to resolve the 512pixel manufacturing anomaly that affects modern CCD; while the existence of this anomaly is well known, its measurement requires dedicated instrumentation like the Laser Comb Frequency (Wilken et al. 2010, Murphy et al. 2012). The measurement of this pixel size variation is extremely useful for a more precise wavelength calibration in instruments dedicated to radial velocity surveys.

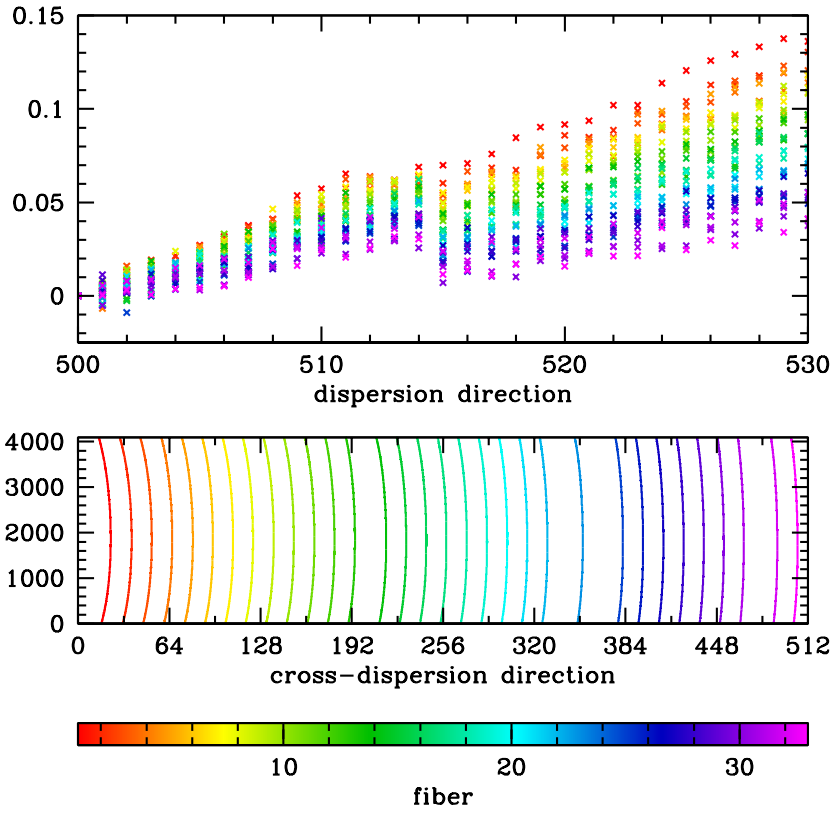

Figure 5: The different size of the $\simeq$ 512th pixel is clearly visible as a jump in the cross-dispersion position of the fibre. Individual fibres are registered using their values at $y = 500$ pixel in the dispersion direction. The first 32 fibres on the CCD are shown, color-coded according to their numer. The size of the jump is proportional to the inclination of the fibre at the considered point.

In figure Figure 5 the jump in the cross-dispersion direction is clearly visible. Several fibres are put together, using the value at $x = 500$ as registering point. The size of the jump is proportional to the inclination of the fibre, being more prominent for the most inclined fibres.

With our technique this measurement is virtually feasible for almost every fibrefed spectrograph. This technique makes use of standard calibration frames and no additional instrumentation is needed, so the gaps can be determined from archival data even for spectrographs that have been dismissed.

2.4.2 *Spectral extraction*

The original ePSF was developed for undersampled detector, where the intra-pixel variations can affect the measured flux, and for instrument where the analytical shape of the PSF is hard to determine. In our case the spectra are well sampled in the crossdispersion direction, the spectra are only slightly tilted so that the flux difference between neighbor pixels due to intra-pixel variations is highly reduced, and the fibre profiles are very close to a gaussian except for a few exceptions, so that the advantages in using the XSF for spectra extraction are highly reduced. Nevertheless an improvement is still visible in medium and low SNR spectra, as showed respectively. in Figure 6 and Figure 6.

For high SNR spectra there is no measurable difference between this techniques and the ESO implementation of the Horne algorithm, as expected in this case. We want to stress the fact that this approach is best when working on highly tilted spectra. An inclination of 20° respect to the vertical corresponds to a pixel-phase variation of 0.336pixel, which is enough to show significant flux loss. The more the spectra are tilted, the bigger will be difference in pixel phase between two adjacent points in the dispersion direction, and the bigger will be the difference of flux caused by the intra-pixel efficiency curve. In the worst scenario the shape of spectral lines can be severally affected.

2.5 CONCLUSIONS

In this chapter we have addressed two basic aspects in data reduction.

The first aspect concerns an in-depth analysis of the data pre-processing, specifically suited for the instrument. An improved algorithm for the Dark Master determination has been introduced, to deal with the low number of frames taken on a monthly basis and the variable glow in the upper side of the CCD that has been affecting the instrument fro several years. We also introduced a new algorithm for the correction of pixel-gain variations.

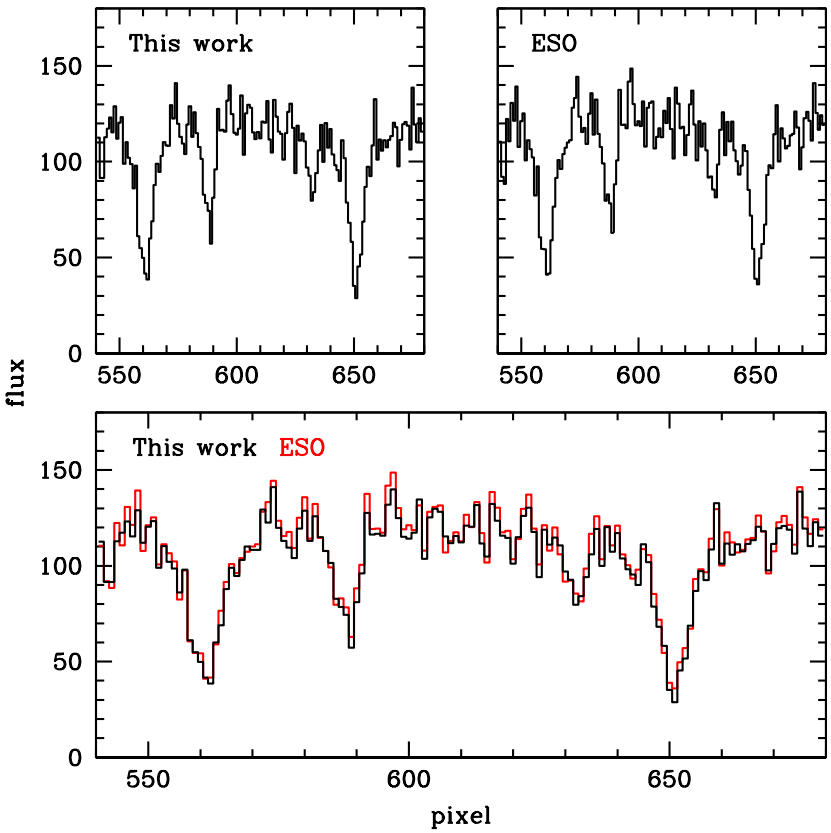

Figure 6: Comparison between ESO (red) and our pipeline (black) in the extraction of a medium SNR spectra.

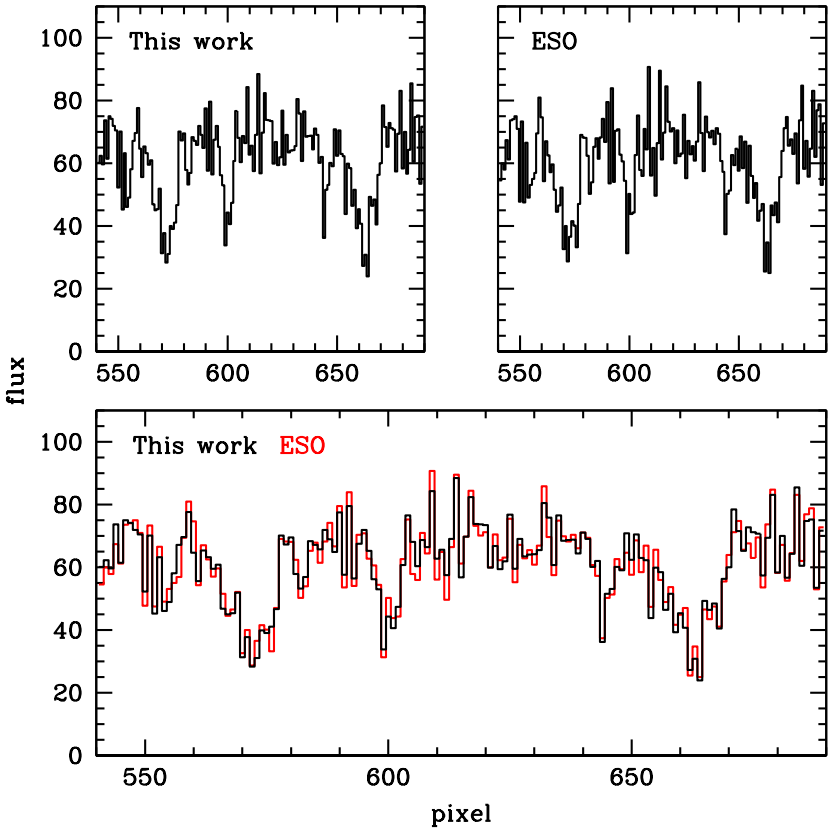

Figure 7: Comparison between ESO (red) and our pipeline (black) in the extraction of a low SNR spectra

In the second part of this work, the empirical Point Spread Function approach has been adapted to the spectroscopy case. This technique is well suited for asymmetric line profiles and automatically takes into account intra-pixel variations. The improvement in the determination of the spectral profile in the cross-dispersion direction results in a measurement of the position of the profile so precise that the $\simeq 512$th anomaly of the CCD is clearly visible as a jump in the position of the fibres. The spectra at our disposal are almost perpendicular to the cross-dispersion direction so intra-pixel loss of light between adjacent pixel in the dispersion direction is minimal, thus reducing the advantage of such approach. Despite of this, a precise knowledge of the spectral profile in the cross-correlation direction is a stringent - but often neglected - requirement for the Horne algorithm in the extraction of low SNR spectra. While the data at our disposal are not optimal to exploit all the advantages given by the ePSF technique, we still observe a decrease of the noise in low SNR spectra extracted with our tools with respect to the ESO pipeline.

3

RADIAL VELOCITIES

During the last decade our group has collected thousands of spectra using the GIRAFFE spectrograph at VLT to determine the geometrical distance and dynamical state of several Galactic Globular Clusters (GGCs) but ultimately the analysis has been hampered by systematics in data calibration and radial velocity measurements, see for example [Milone et al., 2006] and [Sommariva et al., 2009]. Well aware of these issues, we have focused our attention on wavelength calibration and radial velocity determination of very low SNR stellar spectra. Wavelength calibration is performed in two steps, where at the second step the results of the calibration of each individual night are collected together to selected the most stable spectral features in the calibration spectra. The drift in radial velocity that may occur between day-time calibration and science observations is modeled as a function of the position of the spectra in the CCD, and not only as a single value in function of the average position in the cross-dispersion direction as done by the ESO pipeline. We finally applied the technique to a dataset of 7250 spectra for 2771 stars of the GCC M4 (NGC 6121) in the wavelength region $5145 - 5360$Å obtained with GIRAFFE to analyze the internal dynamics of the cluster. Several HST orbits have been allocated to determine accurate photometry and proper motions of stars in this cluster, and the combination of these data with the results of this chapter will provide a deeper insight into the three-dimensional dynamics of the cluster.

3.1 DATA DESCRIPTION AND REDUCTION

The spectra for our project were originally obtained by Sommariva et al. [2009] for an investigation of the internal velocity dispersion of M4 and a search for spectroscopic binaries. The data were gathered with the GIRAFFE medium-high resolution spectrograph fed by the VLT Fibre Large Array Multi Element Spectrograph (FLAMES; [Pasquini et al., 2000]) in MEDUSA multi-fiber mode. In this configuration one can obtain single-order spectra for 132 objects (target stars and sky) in each integration.

The GIRAFFE HR9 setup was chosen, which yielded spectroscopic spectrographic dispersion of 0.05 Å/pixel, measured 4-pixel resolving power $R \equiv \lambda/\Delta\lambda = 25{,}800$, in the wavelength range $5143\text{Å} < \lambda 5356\text{Å}$. This setting has been chosen for the presence of a large number of iron lines and very strong absorption lines,i. e. the Mg-b triplet.

The M4 target stars were selected from an astrometric and photometric catalog based on Wide Field Imager (WFI) data from the ESO/MPIA 2.2m telescope, The original selection criteria was that each star had no neighbors with $V_{neighbor} < V_{target} + 2.5$ within an angular distance of 1.2 arcsec, given the fiber radius of 0.6 arcsec [Sommariva et al., 2009].

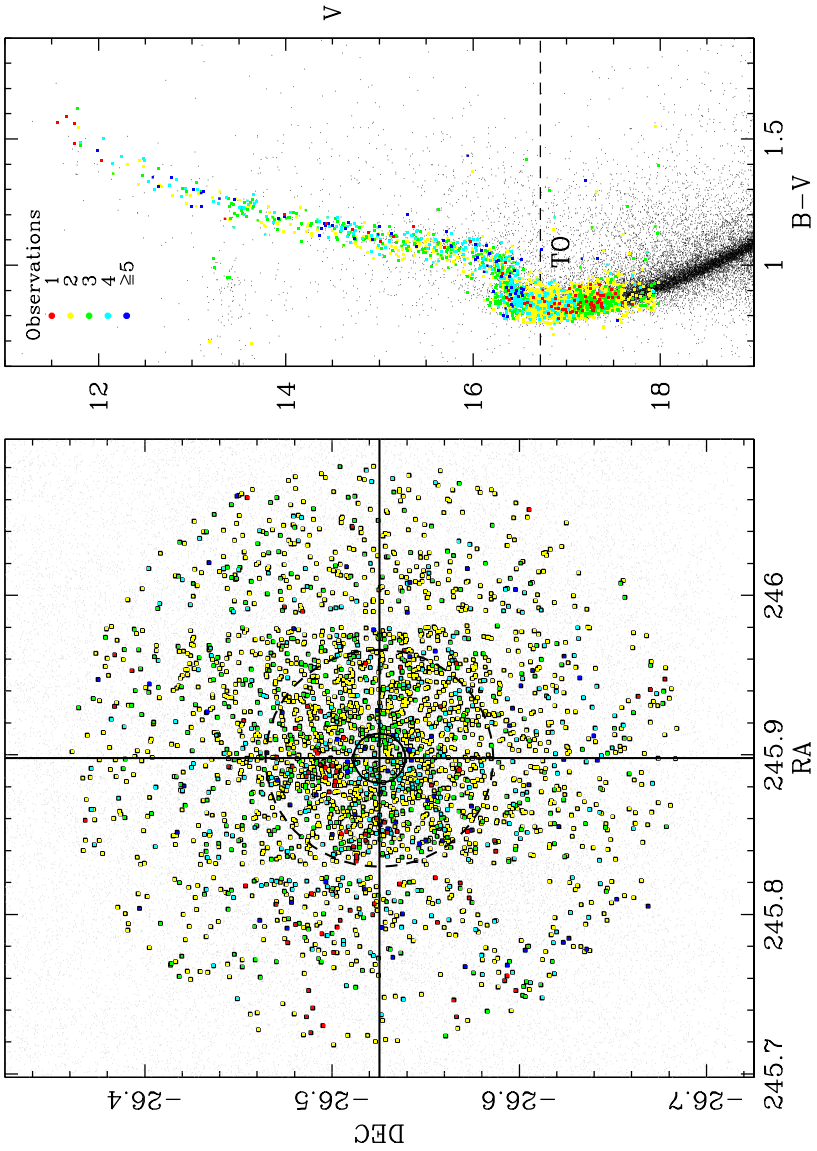

Figure 8: Spatial and photometric properties of M4 target stars. Stars are color coded according to the number of available exposures.

The spatial and photometric properties of our target stars are illustrated in Figure 8. The left-hand panel of the figure contains a cluster map with the positions of our targets. The cluster core radius (where the surface luminosity has decreased by a factor of two from the center) and the half-light radius are also drawn in this panel, in-

dicating that essentially all of M4 has been sampled spectroscopically. The right-hand panel shows a color-magnitude diagram in B and V for the stars. The limiting magnitude of the targets, $V \lesssim 17.5$, was set by the requirement that a single M4 integration had signal-to-noise $SNR > 10$ for each star in the GIRAFFE plate.

A total of 2771 stars covering color-magnitude diagram positions from the upper red-giant branch to about one magnitude fainter than the main-sequence (MS) turnoff (TO) luminosity were observed spanning a temporal period of six years, including 306 new spectra obtained in 2009 and targeting MS stars already observed in the previous epochs. Since these spectra originally were gathered to determine the M4 velocity dispersion and to assess the binary star fraction, nearly all stars were observed at least twice, and three or more spectra were obtained for nearly 40% of the sample. We summarize the observation numbers in Table 1; a total of 7250 individual spectra have been used in our study.

N. of Observations	N. Of Stars	Total Spectra
1	83	83
2	1722	3444
3	501	1503
4	321	1284
5	72	360
6	20	120
7	1	7
8	27	216
9	10	90
10	12	120
11	1	11
12	1	12
Total	2771	7250

Table 1: Spectroscopic Observation Statistics

Sommariva et al. [2009] used the ESO Giraffe standard reduction pipeline [Blecha et al., 2000] and the Ancillary Data Analysis Software [Royer et al., 2002] to reduce the M4 data from raw CCD exposures to wavelength-calibrated 1D spectra, apply the drift correction in the wavelength scale and determine the radial velocities of the observed stars. A systematic offset of 150m/s in the radial velocity zero point is observed in the five simultaneous calibration fibres between 2003 and 2006 data. Changes in environmental conditions inside the spectrograph (i. e. air pressure and temperature) in the interval of time between observations and calibration exposures are automatically taken into account by the ESO pipeline, so the only ex-

pected variation should be a random distribution independent from the epoch as a consequence of the precision of this correction. The dispersion inside a single epoch is at list ten time smaller then the offset (Figure 6 of Sommariva et al. [2009]), and this seems to imply a problem in the wavelength dispersion solution. In the next sections our different approach to wavelength calibration and drift correction is described.

3.2 GIRAFFE WAVELENGTH CALIBRATION

ESO provides a pipeline for GIRAFFE as a part of the *VLT Data Flow System* (DFS). A detailed description of the pipeline can be found at the ESO website (Science Users Information > Data Processing and Tools > Science Software > Instrument Pipelines), here we provide a brief description of the wavelength calibration subroutine to highlight the differences with the technique we used (i. e. *Direct calibration*) in the next sections.

Calibration frames used the dispersion solution determination are obtained by illuminating the entrance of the fibres with a hollow cathode Thorium-Argon lamp (hereafter simply ThAr lamp) using the same setup as for the stellar spectra.

3.2.1 *ESO Pipeline calibration*

A complete GIRAFFE dispersion solution consists of an optical model, a polynomial fit of the optical model residuals, and the correction of individual fibre offsets.

The solution is determined iteratively, i. e. at each iteration the previous model is updated until the measured emission lines have a negligible radial velocity respect to the tabulated values. A previous dispersion solution is needed for a rapid computation, otherwise a new dispersion solution can be computed from scratch using the average optical model that comes along with the pipeline.

At each determination, a set of unsaturated Th-Ar emission lines is selected according to the instrumental setup of the images. The position of these lines on the CCD is determined using an analytical optical model. The residuals between these positions and the measured ones are then modeled with a 2D Chebyshev polynomial. At this stage the slit geometry can be computed again by cross-correlating the ThAr lamp spectrum with a mask of selected calibration lines.

3.2.2 *Direct calibration*

A more classical, direct approach has been followed in this work. For each individual ThAr spectrum, a selected list of emission lines with

known wavelength are measured and then a B-spline is fitted to determine the wavelength dispersion solution.

The algorithm presented here is designed to automatically perform those steps that in other tools such IRAF require a human interaction. Although our reference wavelength range and resolution are the ones defined by the HR9 setting of GIRAFFE, we want to keep our algorithm as general as possible for a rapid implementation on different wavelength ranges or instruments.

3.2.2.1 *Line list of Thorium and Argon lines*

Several Th-Ar line lists are available in literature [Murphy et al., 2007, Lovis and Pepe, 2007], which lines have been selected according to their stability and strength from high resolution spectra. When dealing with medium resolution spectroscopy however blending with close or faint lines is becoming important: the center of the lines can be shifted of several hundredth of angstroms, thus leading to a less precise determination of the wavelength dispersion even with excellent data.

To take into account line blending, a new line list is created using the Th-Ar reference spectrum obtained with CES [Enard, 1982]. This spectrum has been obtained with a resolving power of $R = 200,000$, which is almost nine time the resolving power of GIRAFFE in the HR9 setting; the spectrum has been degraded to match the resolution of our spectra; we will refer to this smoothed spectrum as the *reference* one. All the calibration will be relative to the reference system given by this spectrum.

Lines have been automatically identified using the first derivative of the spectrum. To separate strong lines from small features and weak emission lines, only points with null derivative at the expected line central wavelength and inflection points on the side that have a minimum value of 200 flux units per Ångstrom have been selected.

Figure 9 shows an example of the line identification process, with the smoothed spectrum in the upper panel and its first derivative in the lower one. In our wavelength range 178 lines have been identified. Flux selection has been performed using the flux measurements from the ThAr lamp spectrum during the dispersion solution determination.

3.2.3 *First approximation of the wavelength dispersion solution*

A first rough approximation of the dispersion solution is needed to allow the algorithm to determine the solution.

A given number of the brightest lines (ten lines in our case) are selected from the reference ThAr spectrum in the wavelength range of interest. The number of line must be chosen in a way that the included lines form a subset with mean flux that is at least a few times

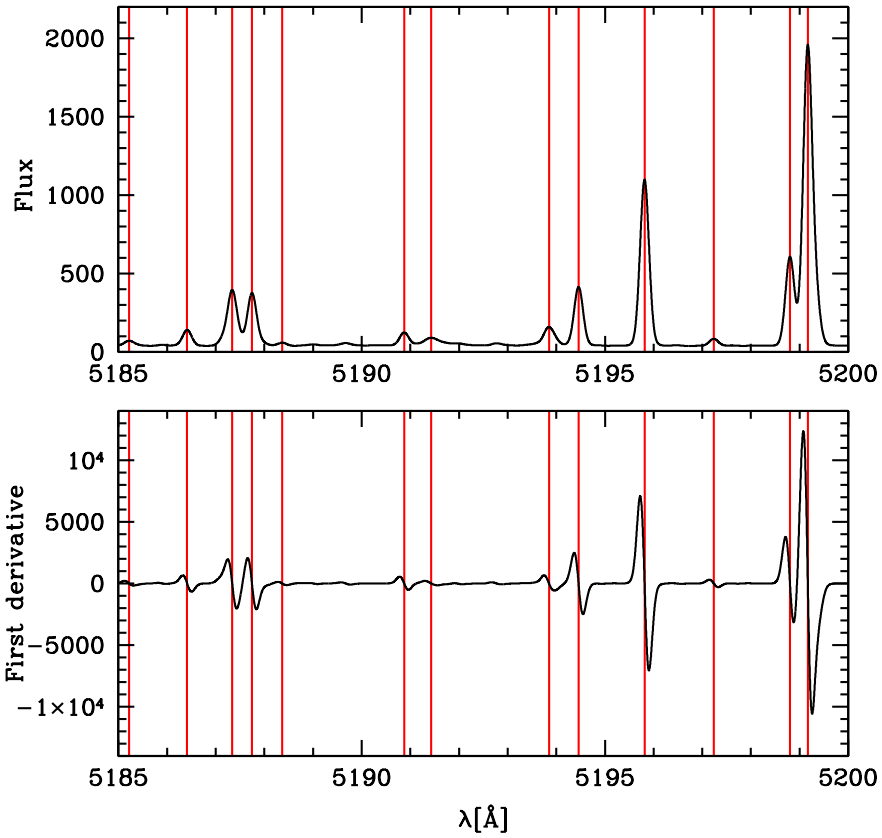

Figure 9: An example of the line identification, with the smoothed spectrum in the upper panel and its first derivative in the lower one. Very weak lines will be escluded according to their flux measurements in the Thorium-Argon lamp spectrum.

the flux of the brightest lines not included in the subset; in this way the selection in the observed spectra will not be affected by differences in the line flux ratios that may be present since we are comparing spectra obtained with lamps made by different manufacturers. A very rough pixel position for each line must be provided too: a single value is good enough for all the fibers, despite their shift in the cross-dispersion direction. The wavelength values are used as initial guess for a gaussian fit, so a precision of one half of the FWHM is good enough. This selection can be easily performed by eye looking directly at the reference spectrum, and must be done only once for a given wavelength range or instrument setup. This is the only human interaction required by the algorithm.

A gaussian fit is performed on both the pixel space for the observed line and in the wavelength space for the reference one. The values are sorted in increasing order, so that the values in the pixel space have now an associated wavelength. The only precaution to be taken at this step is to check that no strong lines are present close to the border of the wavelength range (e. g. 10Å around the border), otherwise an error in the association of lines can occur.

Before proceeding to the first approximate determination of the wavelength solution, we need an additional step to ensure that cosmic rays or uncorrected hot pixels have not been identified as spectral lines. By eye inspection we noticed that spectral lines have a total shift of around 100 pixels across the fibers, with the maximum shift respect to the first fiber around the center of the frame, so we set this threshold to 50 pixels. This shift is a consequence of the optical design of the instrument and it varies smoothly across the sensor, so for each wavelength value we perform a low-degree polynomial interpolation of the pixel values versus their fiber number. The fit is performed a second time after the removal of outlier pixel values. The approximate wavelength solution is finally obtained interpolating the wavelengths values versus the corresponding pixel values with a low order polynomial (degree=3)

3.2.4 *Wavelength dispersion solution for a single frame*

Now that an approximate solution is available, the determination of the final dispersion solution is an iterative process where the resulting precision is increased at each passage. These are the steps followed at each iteration:

- at each iteration more faint lines are measured in the calibration frame and associated with lines in the reference spectrum;
- lines with FWHM higher then a given threshold, nominally twice the instrumental FWHMat that wavelength, are excluded from the list;

- the dispersion solution is computed again, and the order of the polynomial can be increased according to the number of added lines.
- Lines that have an interpolated wavelength that differs from the reference value over a given threshold are rejected and not included in the next iterations; the threshold is initially very large and it is rescaled at each iteration.

In the last two iterations, the lines are weighted according to their brightness and their width in both the calibration spectra and the reference one. After several tests Equation 10 was found to provide the best combination of these parameter.

$$w_i^{-1} = \sqrt{\sigma_{cal,i}^2 + \sigma_{ref,i}^2} + 500/\sqrt{amp_i} \tag{10}$$

In the last two iterations, the polynomial function is substituted with a B-Spline (EFC algorithm from SLATEC).

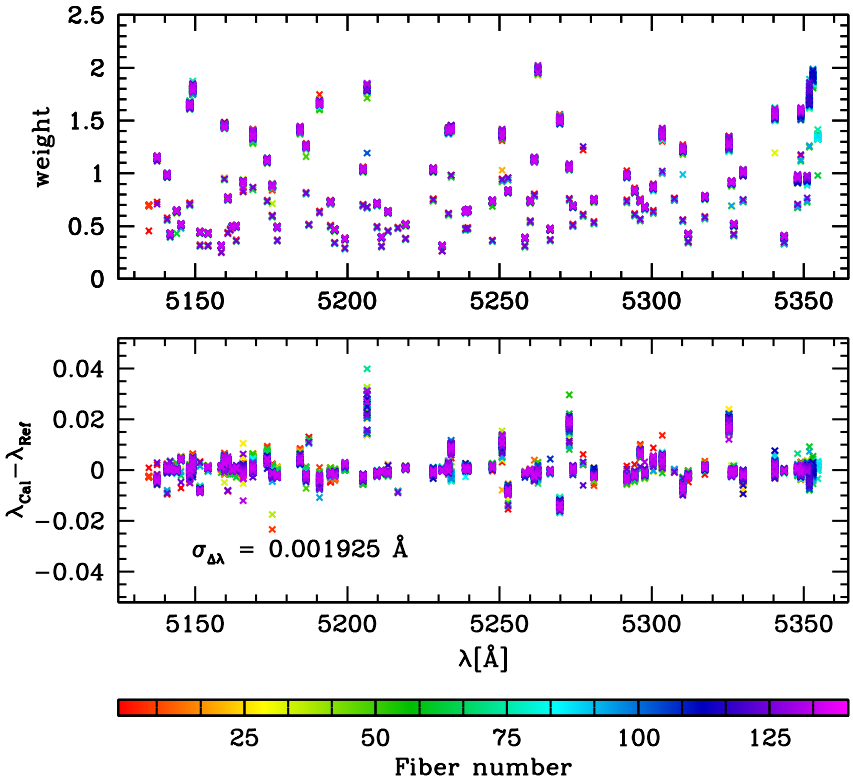

Figure 10: In the upper panel, the weight associated to each line. In the lower panel, the difference between the wavelength determined using the dispersion solution for the measured center of the line, and the wavelength associated using Equation 10

The Standard Deviation of the calibration of a single fiber is $\sigma = 0.00193$Å, which correspond to $\simeq 1/20^{th}$ of a pixel. This procedure is repeated individually for each fiber.

3.2.5 *Overall calibration*

After the wavelength calibration of a frame, the lines used for the dispersion solution determination are stored. Our dataset span a large temporale range, from almost the beginning of the scientific observations of GIRAFFE in 2003 until 2009. The information on emission ThAr lines from all these frames is collected togheter to identify a stable set of lines and improve the dispersion solution determination. Only the lines that have been successfully measured on 90% of the fibers are retained; the only exception is represented by those lines that are closer than 10Å to the edge of the spectrum, for the same reasons explained in Section 3.2.3. The weight associated to each line is the mean of the weights calculated in each calibration fiber. A total of 114 lines has been selected (2).

Table 2: Final list of lines and weights used for the dispersion solution determination

n	Wavelength [Å]	Weight
1	5134.74566	8.776
2	5136.12025	26.557
3	5137.47752	15.334
4	5140.76424	13.070
5	5141.78381	5.382
6	5143.91369	8.399
7	5145.30472	6.584
8	5148.21225	19.641
9	5149.21024	21.811
10	5151.61036	5.564
11	5154.24336	5.420
12	5157.45124	26.060
13	5158.60425	3.818
14	5159.59109	19.307
15	5160.7159	9.8942
16	5161.54024	4.598
17	5162.28824	6.315
18	5163.45824	6.360
19	5164.40224	23.736
20	5165.76825	12.135
21	5166.65224	23.825
22	5168.92424	18.375

23	5170.24825	23.561
24	5170.24825	25.057
25	5173.67824	14.890
26	5175.32625	11.583
27	5176.96425	6.212
28	5178.46025	21.545
29	5180.71982	22.175
30	5182.52425	15.943
31	5184.45624	19.374
32	5186.41225	16.782
33	5187.34024	5.066
34	5190.87425	19.680
35	5194.45825	9.513
36	5195.81424	5.877
37	5197.23625	21.669
38	5199.17024	4.668
39	5202.02312	21.904
40	5203.84753	19.515
41	5205.15225	13.871
42	5206.51225	19.620
43	5207.79067	24.472
44	5209.72524	9.314
45	5211.23087	4.933
46	5213.35889	8.251
47	5216.61235	6.105
48	5219.11214	6.626
49	5220.90312	8.875
50	5226.53226	24.091
51	5228.22914	14.010
52	5228.99845	22.407
53	5231.15993	3.836
54	5233.22626	18.568
55	5234.12225	19.372
56	5237.91226	28.873
57	5238.81424	8.417
58	5239.55023	8.574
59	5241.85626	30.608
60	5243.75424	24.384
61	5247.65424	9.685
62	5250.87424	18.822
63	5252.80426	11.150
64	5254.25424	12.568
65	5258.36026	4.827

66	5260.10461	9.866
67	5261.47626	15.428
68	5262.60549	23.083
69	5264.80026	19.409
70	5266.71024	6.025
71	5269.78424	19.385
72	5272.93423	14.107
73	5274.11826	9.093
74	5277.49675	13.711
75	5281.06977	9.834
76	5282.40097	23.019
77	5283.69025	24.195
78	5286.88954	15.986
79	5289.90476	21.810
80	5291.82021	13.201
81	5294.39806	11.068
82	5295.08657	17.948
83	5296.28228	9.744
84	5297.75750	8.974
85	5300.52426	11.599
86	5301.40424	13.648
87	5303.46885	18.513
88	5305.69226	16.192
89	5307.46605	14.933
90	5309.61426	20.676
91	5310.26900	16.136
92	5312.00226	5.200
93	5315.22424	19.243
94	5317.49426	10.140
95	5320.77270	21.339
96	5322.90024	20.316
97	5325.40787	17.428
98	5326.27241	11.946
99	5326.97626	6.434
100	5329.39426	10.258
101	5330.07825	13.466
102	5333.39026	24.130
103	5337.01251	24.140
104	5338.26797	23.143
105	5340.50478	19.221
106	5343.58172	4.772
107	5345.31539	22.528
108	5346.38055	20.157

109	5347.97374	12.361
110	5349.00534	19.482
111	5351.12889	12.436
112	5351.84424	21.831
113	5353.02077	22.987
114	5354.60553	17.678

The dispersion solution is then computed again, using the same function of the last iteration in Section 3.2.4 but with the improved line list.

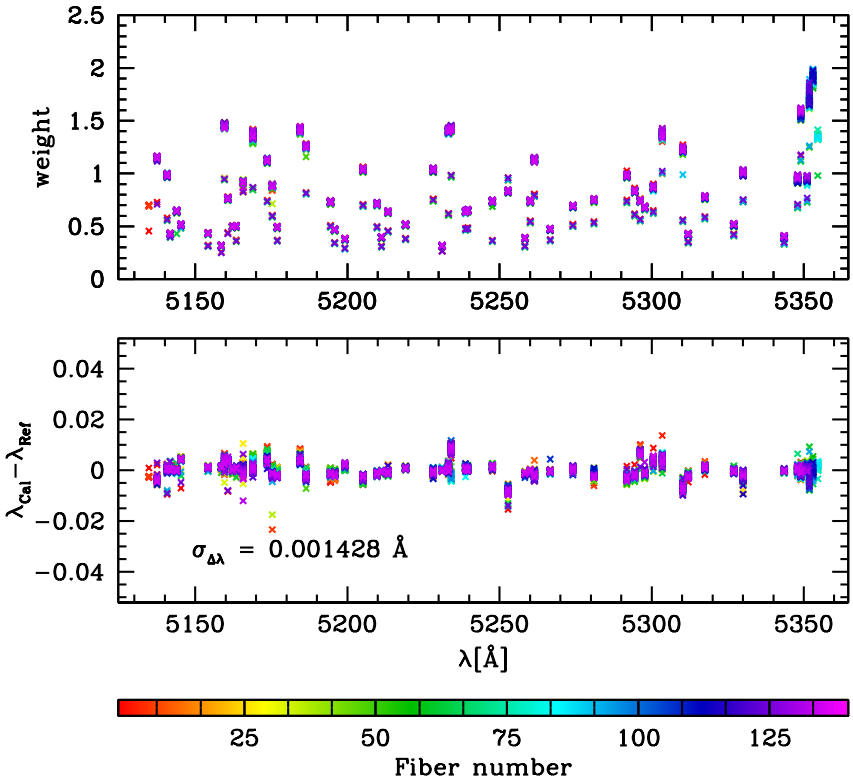

Figure 11: As in Figure 10, but using the refined line list.

Using the same set of lines and weights to compute the dispersion solution results in a more homogeneous calibration of the entire dataset, while using the more stable lines reduce the scatter around the computed dispersion. In Figure 11, the resulting overall calibration for the same frame in Figure 10 is shown. A further reduction of the dispersion from $\sigma = 0.00193$Å to $\sigma = 0.00143$Å is obtained, given an average pixel size of $\simeq 0.45$Å/pixel.

3.3 SIMCAL CALIBRATION

A simultaneous calibration lamp (SimCal) provides a reference system simultaneous to the observations to take in account drifts of the spectrum in the wavelength space that may occur during the time between the ThAr lamp calibration frame and observations.

GIRAFFE has 5 fibers available for ThAr SimCal, homogeneously distributed between the fibers. In ESO pipeline, drifts are corrected by cross-correlating the SimCals with the provided numerical mask; the resulting offsets are linearly interpolated across the CCD in the cross-dispersion direction to determine the wavelength offset to be applied for each fiber to the dispersion solution obtained using the day-time calibration frame. Two apparently reasonable assumptions are made here: the drift is constant in the velocity space across the wavelength range of the considered order, and that the overall shift is linear in the spatial direction of the CCD.

We first check if these two assumptions are correct. In order to do so, the calibration frames from two different nights are taken; emission line positions for each fibers are measured in both frames and the difference in pixel space is taken.

Conversion of differences from pixel space Δx_i^{pixel} into radial velocity space ΔRV_i for the i_{th} line is given by:

$$\Delta RV_i = \Delta x_i^{pixel} \delta\lambda(i)\ c/\lambda_i \tag{11}$$

with λ_i given by the day-time wavelength dispersion solution, $\delta\lambda_i$ is the size of the pixel in the wavelength space at the position of the i_{th} line.

Figure 12 and Figure 13 shows the difference between the calibration frames of two nights. Different nights combinations show similar results, being the only difference the amount of overall radial velocity shift. It appears clear that both previous assumptions do not hold: shift is not constant for a single fiber, and it's not a linear function. Probably this differential shift happens also during the night, even if in a less extended way. It is clear that correcting for a single value of the radial velocity will not provide a good correction for drifts, specially if the instrument is not stabilized (thermally controlled, vacuum) as in our case.

A 2D polynomial fit (Chebyshev polynomials of the first kind) of degree $(2, 2)$ has been chosen to model the shift in Radial Velocity as a function of pixel coordinates.

During science observations only five fibers out of 136 are available for simultaneous calibration (fibers 1,32,63,94 and 125). To check if the polynomial fit is still working with a reduced number of fibres, we performed two fits of the measured RV differences for two given nights (first panel of Figure 14): the first one using all the available fibers (second panel of Figure 14), and the second one using only the

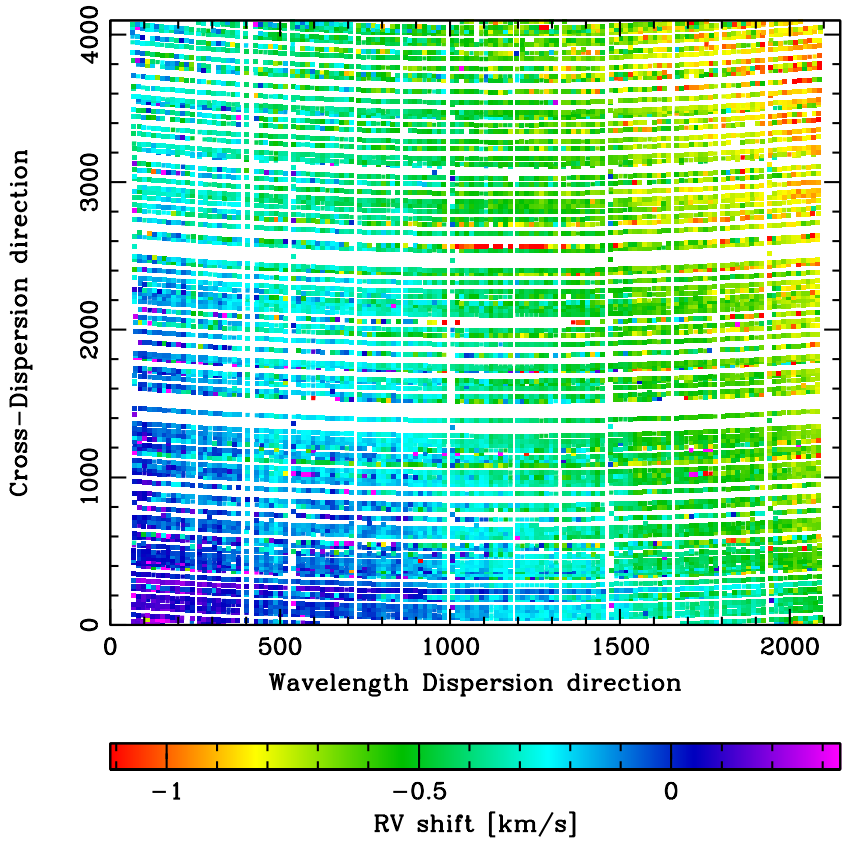

Figure 12: Radial velocity shift of ThAr emission lines between two nights over the CCD. The RV shifts are color-coded.

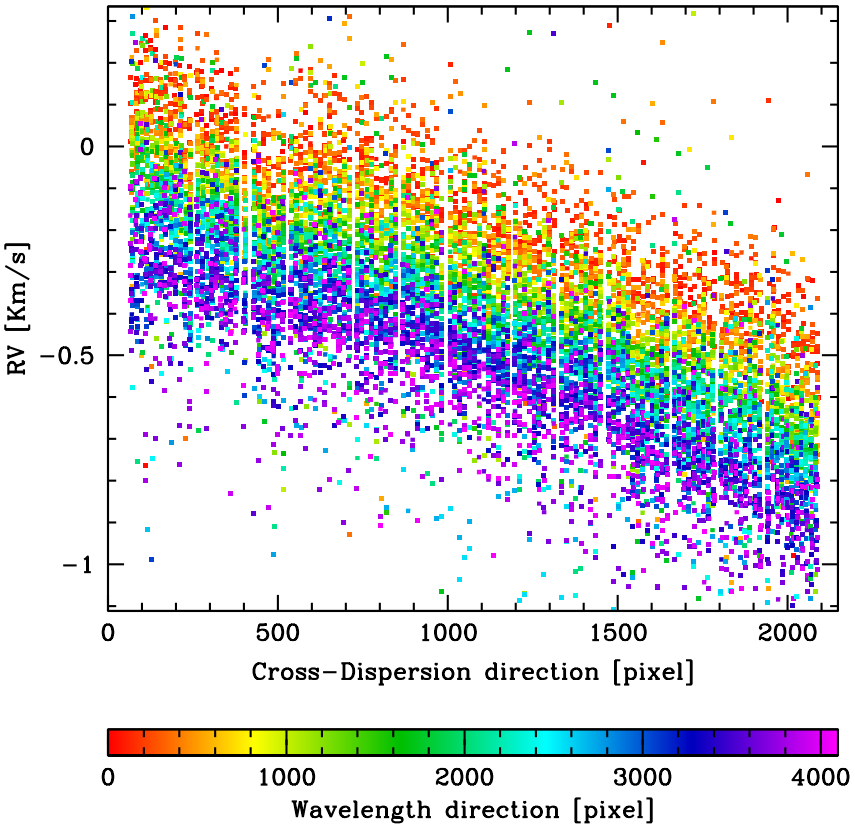

Figure 13: Radial velocity shift along the cross-dipersion direction. The position along the wavelength dispersion direction is color-coded.

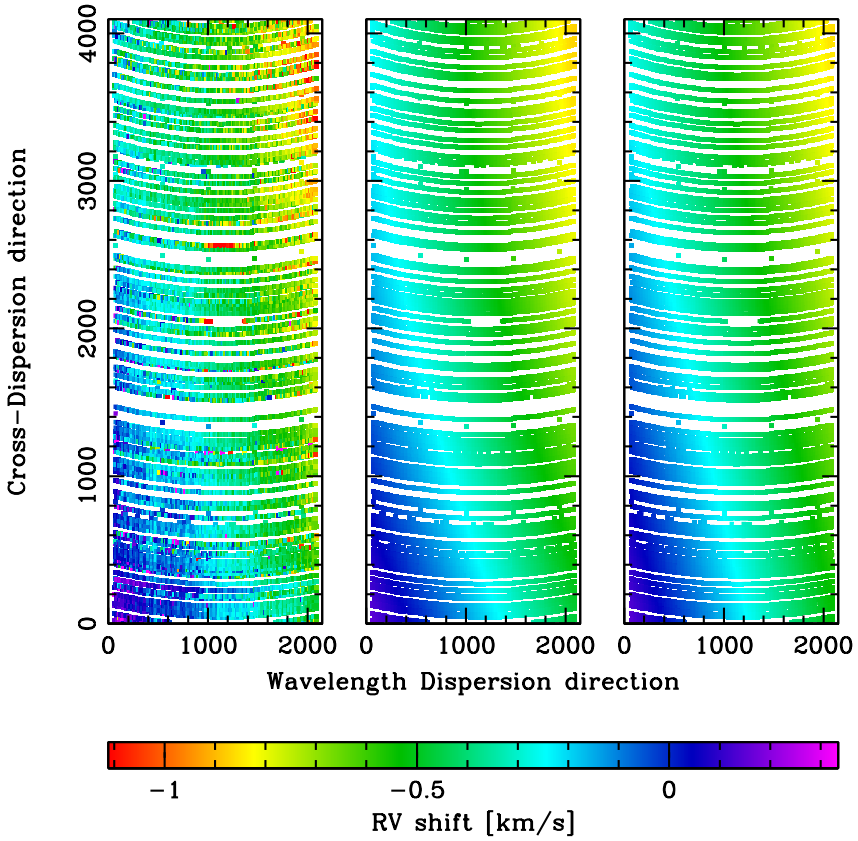

Figure 14: First panel: the observed RV shifts between the Th-Ar calibration frames of two different nights. Second panel: 2-D polynomial fit using all the available fibers in the frame. Third panel: the same fit is performed using only the Simultaneous Calibration fibers. Radial Velocity shifts are color-coded.

SimCals fibers (third panel of Figure 14), in both cases using the same function to fit the measured differences.

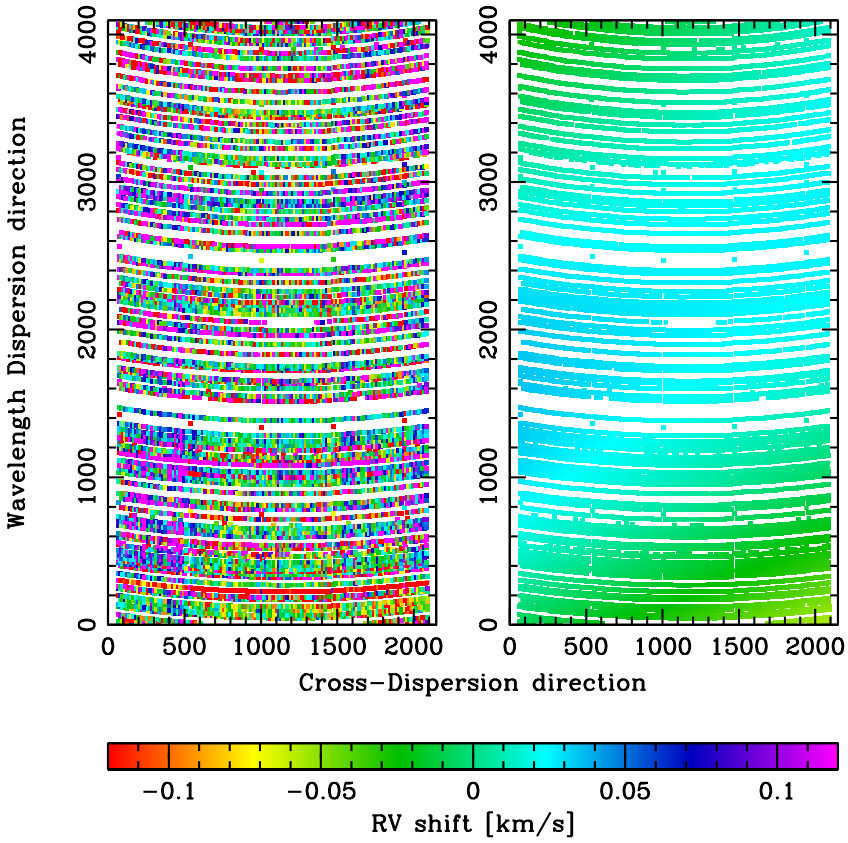

Figure 15: First panel: differences between the SimCals drift model and the observed radial velocity shift. Second panel: differences between the SimCals drift model and the same model using all the calibration fibers.

In the example shown in Figure 15 , the difference between the drift model using only SimCal fibers and measured line positions (left panel) has a mean of $\overline{\Delta_{RV}} < 10 m/s$ and standard deviation $\sigma \simeq 125 m/s$ after $5-\sigma$ clipping. The difference with the model obtained using all the fibers (right panel of Figure 15) shows a systematic of $\overline{\Delta_{RV}} < 10 m/s$ and a $\sigma \simeq 20 m/s$ (after $5-\sigma$ clipping). Considering that we are using only 5 fibers out of 136 to redetermine the drift across the whole detector, this is a reasonable result.

3.4 RADIAL VELOCITY DETERMINATION

In the next section the three most important techniques in radial velocity determination are described: *Classical* Cross-Correlation Function, Numerical CCF and Synthetic Template Matching. The step for

the determination of the radial velocity for each exposure is then described in detail.

3.4.1 *Classical Cross Correlation function*

This is the technique ideated by Tonry and Davis [1979] and implemented in the RVSAO package [Kurtz and Mink, 1998] The observed spectrum is rebinned into a linear scale, continuum normalized and then the Fourier power spectrum (FPS) is computed. The Fourier power spectrum of a pre-selected template spectrum is computed too, and the two FPS are cross-correlated to determine the radial velocity shift. The main disadvantage of this technique is the rebinning process (required for the FPS computation) that can introduce correlated noise, in case of very low SNR spectra

3.4.2 *Binary Mask*

This is the technique introduced by Baranne et al. [1996] and improved by Pepe et al. [2002], also known as *CORAVEL*. The observed spectrum is cross correlated in the wavelength space with a numerical mask with non-negative values at the central position of the line, and zero otherwise. The value of the mask, or *weight*, is taken using a synthetic spectrum and it takes into account the depth of the line and its width, in order to optimize the extraction of the doppler shift information. The cross correlation function is constructed by shifting the mask as a function of the Doppler velocity, and integrating the product of the mask with the observed spectrum. This technique has proven its reliability in the discovery of several exoplanet, but achieves the best results when a huge number of lines is available; see an example in Pepe et al. [2004].

3.4.3 *Synthetic Template matching*

A synthetic spectrum to be used as rest-frame reference is generated using the atmospheric stellar parameters from photometry. The observed spectrum is shifted in the RV space and for every shift V_r the χ^2 of the difference between the observed and the synthetic spectrum is computed. The value of V_r that gives the lowest χ^2 is taken as the radial velocity of the star.

The sky spectrum is removed from the observed one and the result is normalized to the continuum function of the stellar flux as determined in Section 3.4.4. Instead of a χ^2 function, we minimize the *fitness function* defined in Equation 12, where $f_\star$ is the observed stellar spectrum (after the steps described above) and f_{syn} is the syn-

thetic one, and the sum is over the n_{pixel} data points of the observed spectrum.

$$FIT(V_r) = \sum_{i}^{n_{pixel}} \left(\frac{f_\star(i) - f_{syn}(i, -V_r)}{1.2 + f_{syn}(i, -V_r)} \right)^2 mask(i) \tag{12}$$

For each determination of the fitness function the synthesis is shifted of $-V_r$ in the radial velocity space (in the opposite direction since V_r is the velocity associated to the star) and then rebinned into the wavelength scale of $f_\star$, thus avoiding the more problematic rebinning of the noisier observed spectrum. This is the reason why the dependance on V_r has been included in the synthetic spectrum $f_{syn}(i, -V_r)$.

The denominator of Equation 12 differs significantly from the one of a standard χ^2 function. We did not include the dependance on the continuum in FIT determination because this shape is changing with time and changes in instrumentation (e. g. the change of CCD in May 2008), and it could have introduced a bias depending on the epoch of observations. We instead decided to divide ny the synthetic spectrum to give more weight to spectral lines, while the constant has been added to avoid an excessive weight for the deepest lines.

Pixels affected by cosmic rays or CCD defects and spectral ranges contaminated by Th-Ar emission from close SimCals fibres are flagged with a zero value in mask, and unitary value otherwise.

We compare the Numerical CCF and the Synthetic Template Matching CCF in Figure 16. The reduced number of lines due to the small wavelength range of the spectrum and the low SNR of our spectra make the second technique our chosen one.

3.4.4 *Continuum normalization and Sky flux determination*

Continuum normalization is one of the most daunting tasks in stellar spectroscopy: it usually performed manually by interpolation of lines-free wavelength intervals identified by eye, or automatically when the EW of the line has to be determined and the selected line is quite isolated, since only the knowledge of the local continuum level is required.

Our dataset present several major complications. Spectral range is very short, $\Delta\lambda = 214$Å, and very crowded, so there are no line-free region available to set a continuum level; in fact, the HR9 setting has been chosen for the richness in this spectral window. The wings of the Magnesium-I triplet (5167.33Å, 5172.70Å, 5183.62Å) are affecting nearly one quarter of the spectrum and from its initial part, so that the continuum must be extrapolated from the last three quarters of the spectrum . Moreover, many lines are blended or very close making the determination of local continuum very difficult.

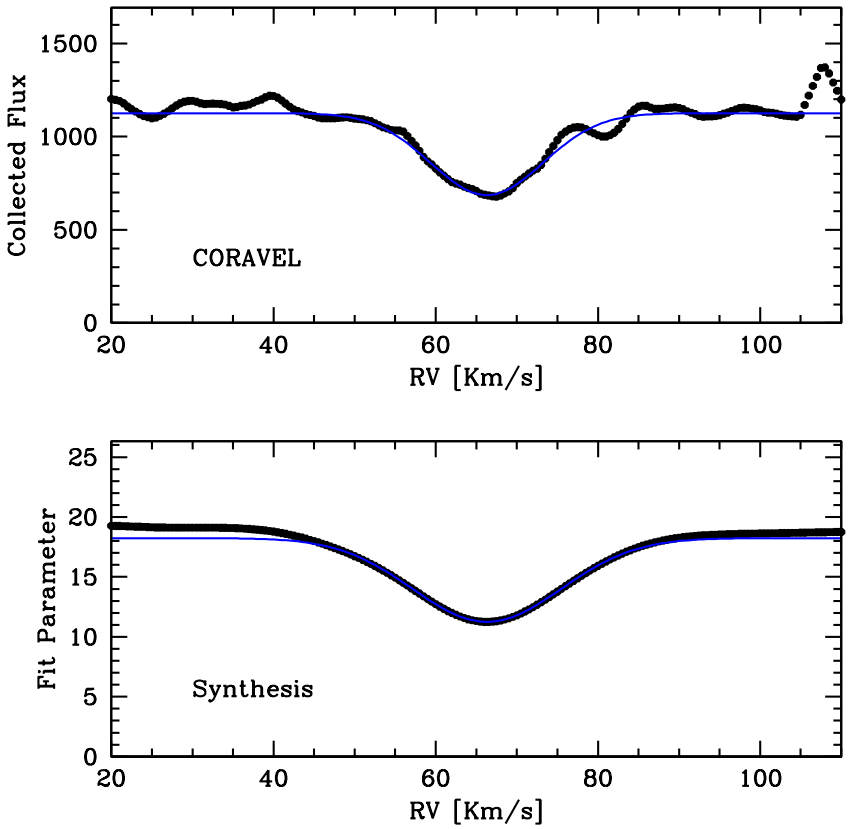

Figure 16: Comparison between the Numerical CCF (upper panel) and Synthetic Template Matching (lower panel). The blue line represents a Gaussian fit.

To determine the spectrum continuum, we used a technique similar to the one introduced in Section 4.1.2. Briefly, the continuum level and sky contribute are determined by using a spectral synthesis, obtained using the photometric stellar parameters and average metallicity of the cluster, and the Solar Flux Atlas Kurucz et al. [1984] as a template sky spectrum, both normalized to unity. Photometric atmospheric parameters for M4 have been derived using the calibrations from Ramírez and Meléndez [2005] and Casagrande et al. [2010] as described in Section Section 4.2; the latest determinations for reddening and distance from Hendricks et al. [2012] have been used, i. e. $E(B-V) = 0.37 \pm 0.01$, $R_V = 3.62$ and $D_{cluster} = 1.80$kpc. Syntheses are generated using the current version of the Local Thermodynamic Equilibrium (LTE) code MOOG [Sneden, 1973], and Kurucz [1992] atmosphere models calculated with the New Opacity Function Distribution (ODFNEW, Castelli and Kurucz 2004). More details can be found in Section Section 4.3.

To apply this technique, we need to know the radial velocity of the star. At this stage high precision in RV is not required: a single pixel has an average size of 2.5km/s, so an error of several hundreds of meter still permits a good fit of the synthesis over the observed spectra, given the fact that our goal at this stage is to determine the continuum level and not to derive accurate stellar parameters.

The radial velocity is derived using a numerical CCF [Baranne et al., 1996] with line positions and weight derived from the same synthetic spectrum that will be used to determined the continuum, following the prescriptions in Pepe et al. [2002].

$$CCF(V_r) = \sum^{i} CCF^i(V_r) = \sum^{i} \int_{\lambda^i_{V_r}-\delta\lambda/2}^{\lambda^i_{V_r}+\delta\lambda/2} f(\lambda) c^i d\lambda \tag{13}$$

As can be seen from Equation 13, the value of the CCF at a given radial velocity point V_r is proportional to the sum of the stellar flux collected by each *hole* i in the numerical mask, rescaled for the weight c_i. $\lambda^i_{V_r}$ is the wavelength of the line i shifted for the radial velocity of the CCF and $\delta\lambda$ is the size of the hole in the numerical mask.

The numerical CCF allows us to determine the star/sky flux ratio. CCF is computed for smoothed synthetic spectrum (CCF_{synth}) and sky spectrum (CCF_{sky}) without any further correction and using the same numerical mask. The cross-correlation function of the observed spectrum CCF_{star} is then fitted with a linear combination of CCF_{synth} and CCF_{sky}, being the linear coefficients the fractions of stellar and sky fluxes.

The upper panel of Figure 17 shows the CCF_{synth} (red) CCF_{sky} (blue) derived from normalized spectrum; the sky spectrum has a lower continuum and a deeper CCF because of the higher metallicity and lower temperature respect to the observed star. In the lower panel,

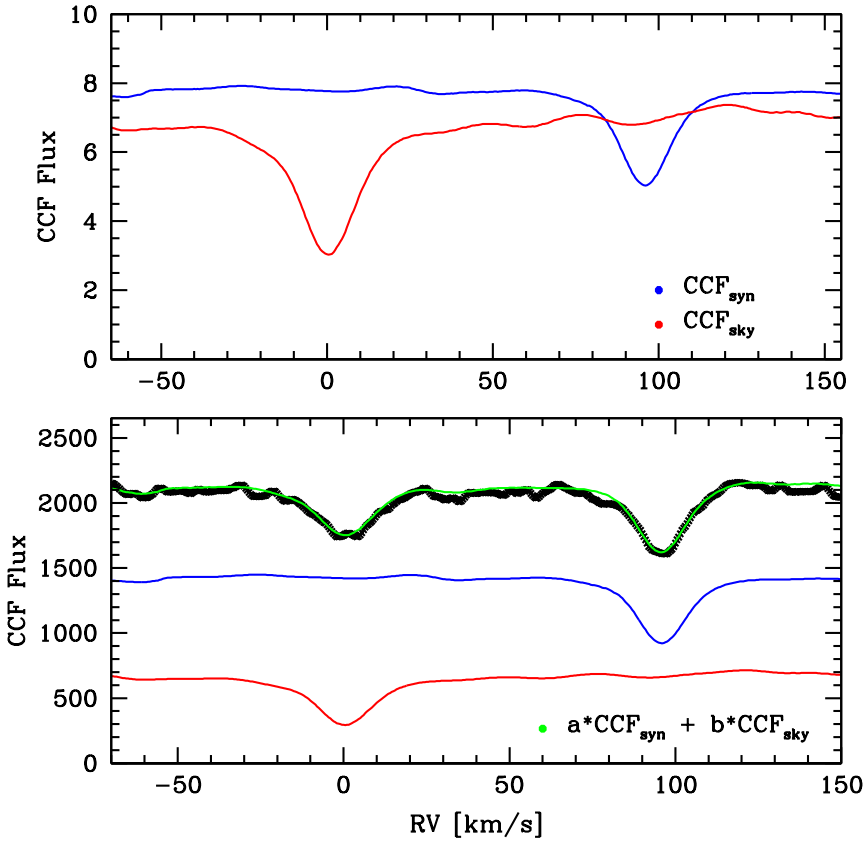

Figure 17: In the top panel, the CCF from the normalized spectrum of the synthesis (blue line) and the sky (red) are shown. In the bottom panel, $\mathrm{CCF}_{\mathrm{syn}}$ and $\mathrm{CCF}_{\mathrm{sky}}$ are rescaled according to the measured flux to show the relative influence, and their sum (green) is compared with the observed CCF (black).

the two CCFs are rescaled according to the measured flux to show the relative influence, and their sum (green line) is compared with the observed CCF (black points)

Note here that this technique works best when the observed RV of the star (including Earth motion contribute) differs significantly from the RV of the sky, i. e. $|V_r^{star}| > 20$km/s.

Radial velocities shifts due to the intrinsic velocity of the star and the Earth motion are applied to the stellar synthetic spectrum. Heliocentric and Barycentric corrections have been determined using the FORTRAN subroutine by Stumpff [1980].

Both solar and synthetic spectra are rebinned into the wavelength scale of the science spectrum (keeping constant the flux for wavelength unit) and degraded to match the spectral resolution of the instrument, and finally the instrumental effects are applied to make the comparison with the observed spectrum possible. A normalized linear combination of the two modified spectra, with the sky/flux ratio as coefficients, is multiplied by a polynomial function: its coefficients are determined by fitting the observed spectrum. The partial continuum levels of star and sky are obtained by substituting the initial synthesis and the sky spectra respectively with unitary spectrum and applying all the cited steps, with the exception of the fit of the polynomial function which coefficients are now known. The overall continuum level is the linear combination of the two partial continuums.

When sky subtraction and continuum normalization play a decisive role, e. g. in the determination of the stellar atmosphere parameters, more refined values for the star/sky ratio and continuum coefficients can be obtained as described in section Section 4.1. More initial conditions are required, e. g. the radial velocity of the star must be precisely known, thus justifying the introduction of the simplified approach shown here.

Figure 18 shows the continuum determination for a $V_{mag} = 16.4$ star, with photometrically derived $T_{eff} = 6000$K, $\log g = 4.0$dex, $[Fe/H] = -1.15$; the spectra has been taken on 2006 September 04, $\simeq 5$ days before full moon and at a distance of $\simeq 40$deg. This technique is well suited for spectra with a short spectral range, low SNR , strong lines and strong sky contribute. All these effects combined together make almost impossible to find suitable line-free continuum zones to normalize the spectra. However, as we show in the lower panel of Figure 18, this technique works well even when the spectrum is strongly affected by the Magnesium triplet from both star and sky.

For faint stars observed several days after new moon, the sky can contribute up to half of the observed spectra, so it must be taken in account. Otherwise the additional lines would lower the continuum level determined using only the stellar synthesis. Our spectral range

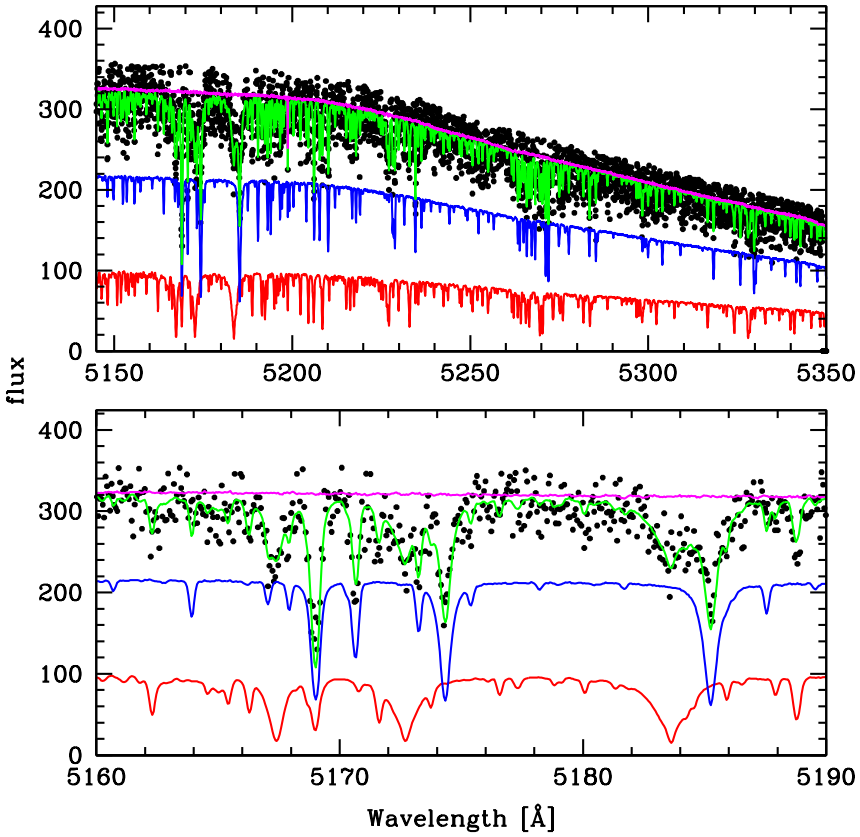

Figure 18: Upper panel: the contribute of star (blue) and sky (red), their combination (green) and the derived normalization level compared to the observed spectrum (black) ($V_{mag} = 16.4$). The magenta line represents the continuum determination. Lower panel: the Magnesium triplet region.

(5140Å $< \lambda <$ 5360Å) is free from atmospheric absorption lines, so no correction is required for this potential issue.

3.4.5 *Final RV determination*

Some of the observations were taken close to full moon, and the sky lines are severely affecting the observed spectrum, as can be seen in Figure 18. Sky lines are affecting CCF too, so the sky spectrum (red line in Figure 18) has been removed from the observed one before computing the RV of the star.

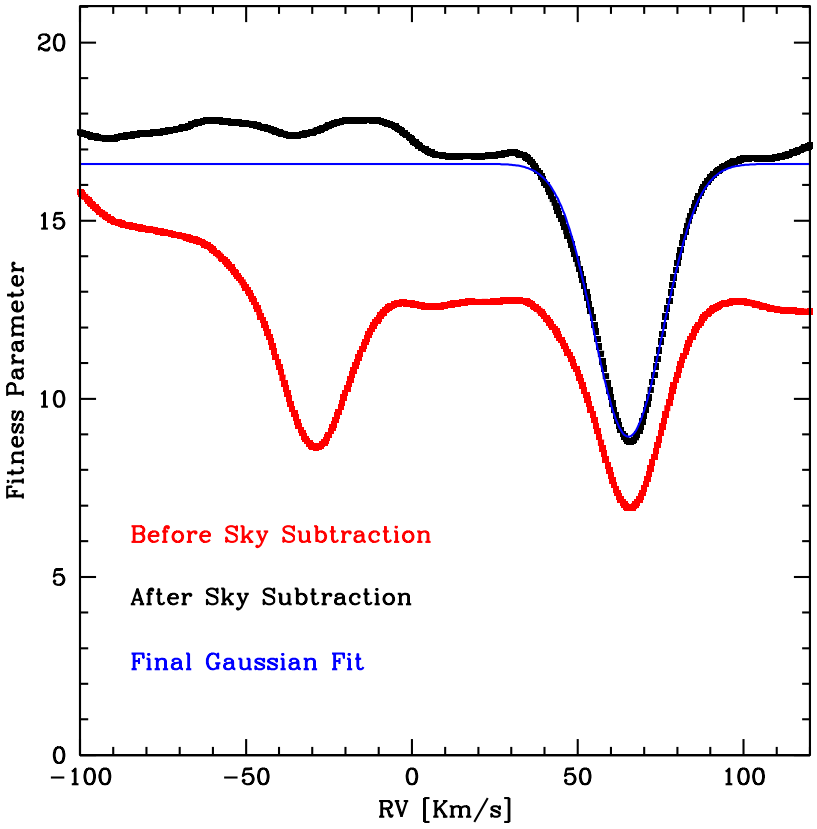

Figure 19: CCF for the same observations in the previous sections, before (red) and after (black) sky subtraction. Blue line represents the gaussian fit to determine the RV of the star

Figure 19 shows the CCF before and after the removal of the sky spectrum. The spectrum has been normalized using the continuum associated to the combined sky+synthesis in the first case, and the continuum associated to the synthesis alone in the second case (Figure 17)

After sky subtraction and normalization, the radial velocity is finally determined with a gaussian fit of the resulting CCF, determined as described in Section 3.4.3.

3.5 RESULTS AND DISCUSSION

The error in the radial velocity associated to a single exposure is a function of the spectral type of the star, the characteristics of the instrument, the radial velocity determination technique and the SNR of the spectrum.

Instead of trying to determine a formal error of a single exposure, we determine an empirical error using the several radial velocity measurements available for each star.

The weighted mean velocity of a star V_i is determined using Equation 14, where i is the index of the star, j is the index referring to individual exposures and n is the total number of observations available for the given star.

$$\overline{V_i} = \frac{\Sigma_{j=1}^{n} w_{i,j} v_{i,j}}{\Sigma_{j=1}^{n} w_{i,j}} \tag{14}$$

We use as weight $w_{i,j}$ the inverse of the minimum of the *fitness function* (Figure 19) after normalizing it for the number of used pixels.

Following Sommariva et al. [2009], the unbiased estimator for the population variance is determined with

$$(\sigma_{V_i})^2 = \frac{\Sigma_{j=1}^{n} w_{i,j}}{(\Sigma_{j=1}^{n} w_{i,j})^2 - \Sigma_{j=1}^{n} w_{i,j}^2} \, \Sigma_{j=1}^{n} w_{i,j} (v_{i,j} - V_i)^2 \tag{15}$$

It is interesting to compare the distribution of the radial velocity dispersion associated to each star before and after the SimCals correction for three different intervals in de-reddened V magnitude (Figure 20). Going from the lower panel to the upper one, we can see how the correction for instrumental drifts in the radial velocity becomes increasingly more important for brighter stars, i. e. with higher SNR. Although radial velocity measurements for faint stars are mostly dominated by photon noise, the improvement in the σ_{V_r} is still noticeable.

The average sigma $\langle \sigma_{V_r} \rangle^{bin}$ for a magnitude bin is computed as the 67.5th percentile of σ_{V_i} distribution. This value is computed for several magnitude bins and the resulting values are interpolated to determine the estimated error on the radial velocity as a function of magnitude $\langle \sigma_{V_r} \rangle (V)$ (the dependence on magnitude is dropped in the next sections).

In Figure 21 our derivation of $\langle \sigma_{V_r} \rangle$ (red line) is compared with the one in Sommariva et al. [2009] (red thick line in Fig. 10 of their paper, blue line in Figure 21). At high SNR (i. e. $V < 13$) the error is dominated by the instrumental stability, the goodness of the SimCals correction and the precision of the RV determination technique, for a total uncertainty of $\simeq 100 m/s$ (< 0.0018Å, or < 0.02pixel). We regard this as the practical precision limit in radial velocity that can be obtained from these GIRAFFE data.

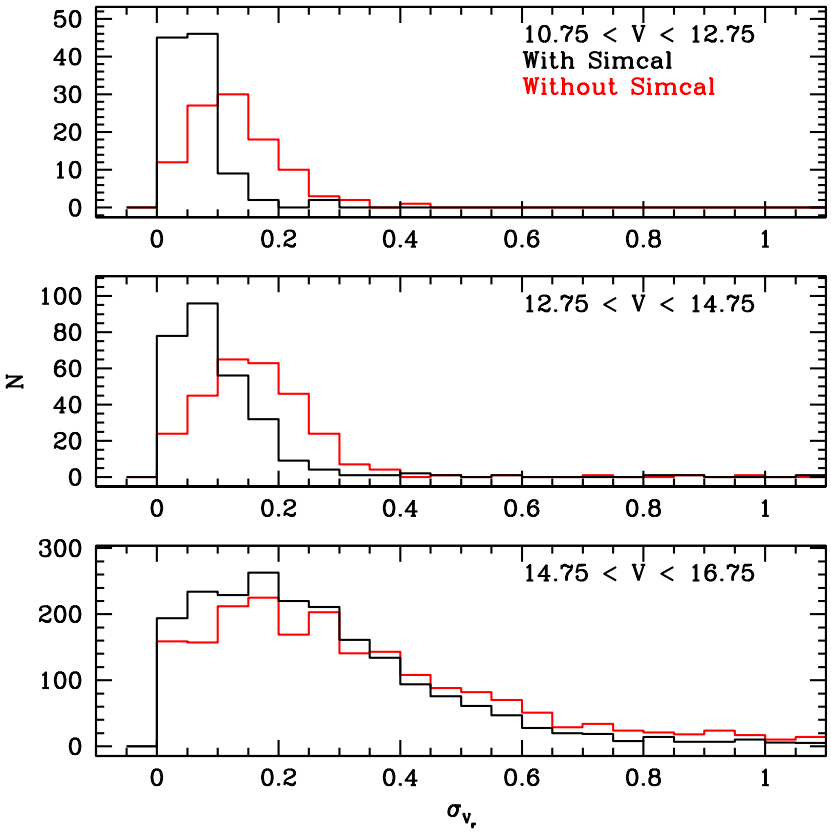

Figure 20: Distribution of the radial velocity dispersion σ_v associated to each star before (red lines) and after (black lines) the SimCals correction. Three ranges in de-reddened V magnitude are considered.

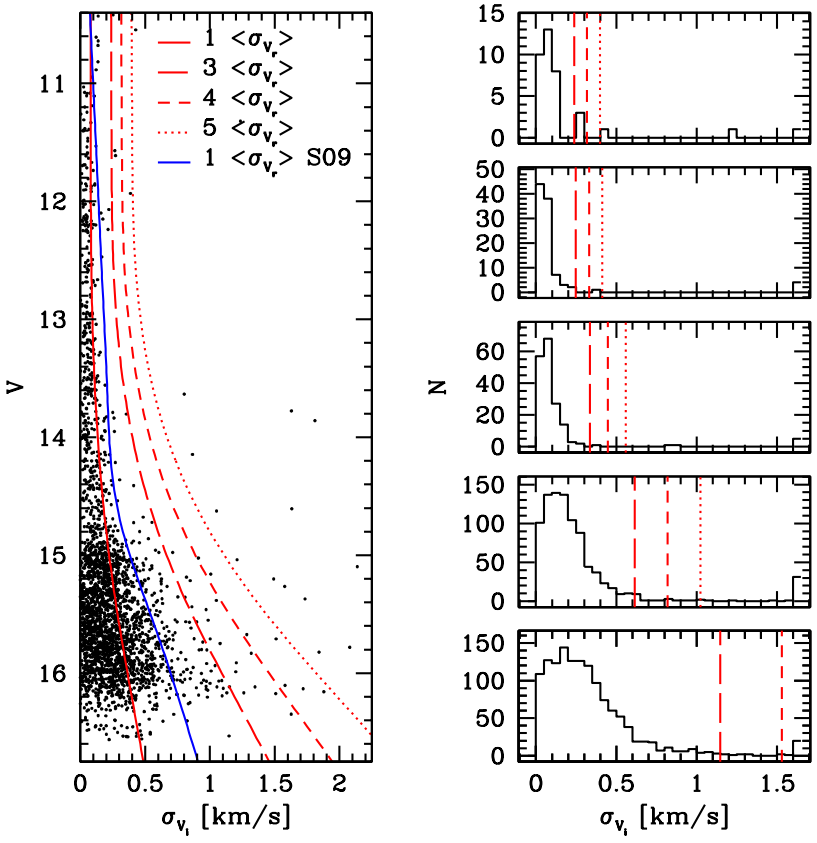

Figure 21: The unbiased weighted radial velocity dispersion determined for stars with multi-epoch measurements. The estimated error on the radial velocity as a function of magnitude $\langle\sigma_{V_r}\rangle(V)$ from our analysis (red thick line) is compared with the one obtained by Sommariva et al. [2009] (blu line) using the same dataset. Stars the satisfy the condition $\sigma_{V_i} > 3 * \langle\sigma_{V_r}\rangle$ are considered binary candidates.

At this magnitude the effect of the sky in RV determination is negligible even for observations close to the full moon. At lower SNR the quality of the spectra becomes the main component in the precision of the RV determination.

The combination of an improved wavelength calibration and a careful determination of radial velocity result in an increased precision of a factor of two in the radial velocity determination of a single star.

Stars that satisfy the condition $\sigma_{V_i} > 3 * \langle \sigma_{V_r} \rangle$ are considered binary candidates. We find 30 candidates with $V < 14.9$ over 464 targets, and 66 candidates with $V > 14.9$ on a sample of 1990 stars, for a total of 96 binary candidates. A total of 39 more candidates are found compared to the previous analyses by Sommariva et al. [2009], using the same dataset and the same statistical approach. We did not attempt to determine the binary fraction because of the incompleteness of the survey.

3.5.1 *Mean radial velocity and velocity dispersion*

The average velocity of the cluster has been computed for four magnitude bins. Stars exceeding five time the estimated error $\langle \sigma_{V_r} \rangle$ evaluated at their magnitude are excluded from the set, while we retain stars with only one RV measurement by assigning them $\sigma_{V_i} = \langle \sigma_{V_r} \rangle$. Outliers in the CMD and obvious not-members (i. e. RV lower than 55km/s or higher than 87km/s) are excluded as well, leading to a total of 2570 stars out of the 2771 of the initial sample. The weighted radial velocity mean for the cluster, the associated error, the dispersion of the distribution and the number of stars for each magnitude bin are reported in Table 3 and displayed in Figure 22. There is no significant trend in the cluster mean radial velocity with magnitude.

N	V range	$\langle V_c \rangle$	$\sigma_{\langle V_c \rangle}$	σ_{V_c}
125	$10.2 \leqslant V < 12.9$	71.022	0.331	3.687
180	$12.9 \leqslant V < 14.2$	70.921	0.317	4.241
1054	$14.2 \leqslant V < 15.6$	70.784	0.123	3.979
1211	$15.6 \leqslant V < 17.0$	70.989	0.108	3.778
2570	$10.2 \leqslant V < 17.0$	70.970	0.078	3.936

Table 3: Radial Velocity Means

The average radial velocity is $\langle V_c \rangle = 70.896 \pm 0.077$km/s with a cluster dispersion of $\sigma_{V_c} = 4.037$km/s.

Our measured radial velocity is in agreement with previous studies. Peterson et al. [1995] reported a radial velocity mean of $\langle v \rangle = 70.9 \pm 0.6$km/s using 200 giant stars. Côté and Fischer [1996] reported a value of 70.3 ± 0.7km/s from the analysis 33 turn-off dwarf stars. Our analysis is in disagreement with the one performed by Sommariva et al. [2009], which determined a mean radial velocity of

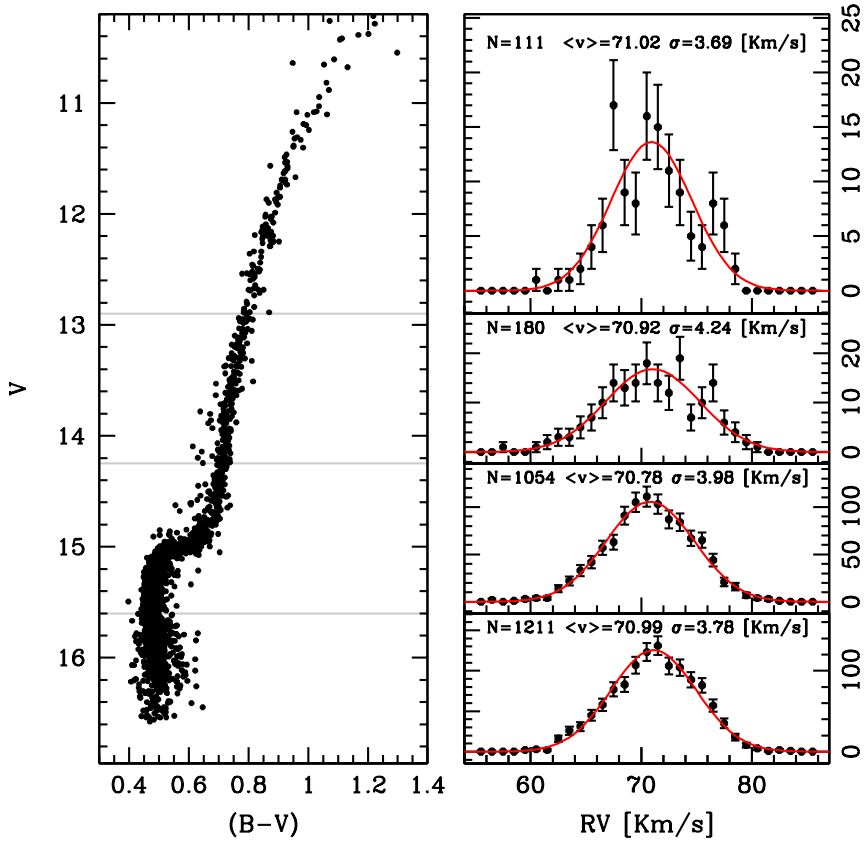

Figure 22: The weighted radial velocity mean of the cluster, the dispersion of the distribution and the number of stars for four magnitude bins, delimited with a grey line in the Color-Magnitude Diagram. There is no significant trend of the mean velocity of the cluster with magnitude.

70.27 ± 0.19km/s using our same dataset. The main differences are in the wavelength calibration and in their use of a numerical mask. Unfortunately no radial velocity standards are present in our dataset so for the moment it is not possible to assest the origin of this systematic offset.

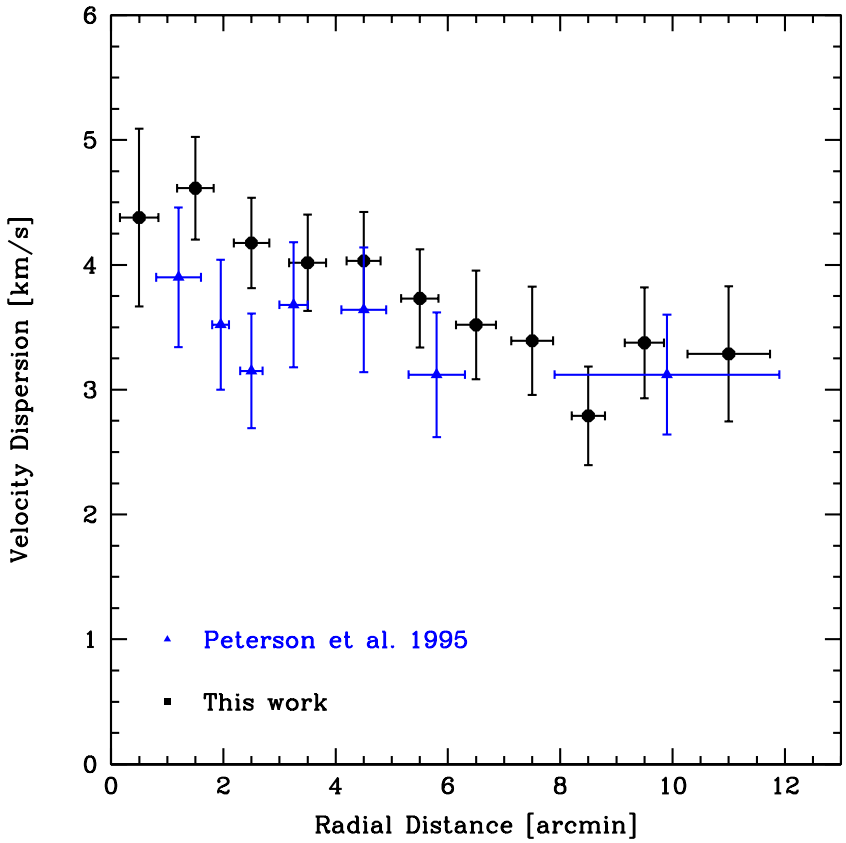

Figure 23: The radial velocity dispersion is plotted as a function of distance from the cluster center (black squares). Data from Peterson et al. [1995] is displayed for comparison. Our data show a systematic higher dispersion close to the center and a more pronounced decrease outward.

The dependence of the radial velocity dispersion with the distance from the cluster center is analyzed by dividing our sample in 11 radial annuli and calculating the weighted velocity dispersion in each bin, as displayed in Figure 23. To take into account the broadening due to the observational uncertainty, for each bin the 67.5th percentile of the error distribution has been computed and the resulting value ($\simeq 0.3$km/s) quadratically subtracted from the cluster dispersion. The error on the mean has been quadratically subtracted as well. For each bin this the vertical error bar represents the error in the mean of each group, while the horizontal error bar is determined by the 67.5th percentile of the distribution of distance respect to the center of the bin.

Compared with the radial profile in Peterson et al. [1995], our data show a systematically higher dispersion close to the center. This trend is still present with the same amplitude when considering stars with at least 3 individual radial velocities, when applying a more stringent selection on RV errors, i. e. $\sigma_{V_i} \leqslant 3 * \langle \sigma_{V_r} \rangle$ instead of $\sigma_{V_i} \leqslant 5 * \langle \sigma_{V_r} \rangle$ or when selecting the stars for magnitude intervals. The radial distribution in Peterson et al. [1995] was determined using 182 members with a RV precision of $\simeq$ 1km/s for individual star, while our sample comprises a larger number of stars (2570 members) and the higher precision for a single star RV determination (a few hundreds of meter per second).

3.5.2 *Cluster Rotation*

Hints of rotation with an amplitude of -0.9 ± 0.4km/s at a position angle of 100° for stars farther than 4′ from the cluster center have been found by Peterson et al. [1995]. We use our extended catalogue to confirm this result. We have divided our stars in two groups, the *inner sample* containing stars within 5′ from the center of the cluster, and the *outer sample* containing all the remaining stars ($d > 5'$). Radial velocities against position angle for the two groups have been plotted respectively in Figure 24 and Figure 25.

Position angle has been divided in bins of 30° for the stars in the inner sample and 24° for stars in the outer sample. The weighted mean and the associated error for each bin have been determined in the same way as described in Section 3.5.1, and are shown in the lower panels in Figure 24 and Figure 25

For the outer sample, a sinusoidal fit with amplitude -0.70 ± 0.2km/s and phase of $120 \pm 10°$ (blu line in Figure 25) results in dispersion of $\sigma = 0.32$km/s, which goes down to $\sigma = 0.19$km/s if the two outliers at 174° and 318° are excluded. In comparison, a linear fist results in a dispersion of $\sigma = 0.47$km/s.

No signs of rotation in the inner part of the globular cluster was found in Peterson et al. [1995]. In our inner sample, a sinusoidal fit with amplitude -0.40 ± 0.2km/s and the same phase of the outer sample (blue line in Figure 24) results in a decrease to $\sigma = 0.26$km/s from $\sigma = 0.33$km/s obtained from a linear fit. If the two outermost points at 135° and 225° are excluded, the dispersion goes down to $\sigma = 0.14$km/s. According to Peterson et al. [1995], such rotation is negligible for the dynamics of the cluster.

It is important to notice that the values of the amplitudes varies significantly ($\pm$0.1km/s) with the size of the bins. Quoted errors are a rough indication of the expected errors obtained by analyzing variations of σ when changing the sinusoidal parameters. Rotation of the cluster from radial velocities and proper motions will the subject of a detailed study in the next future.

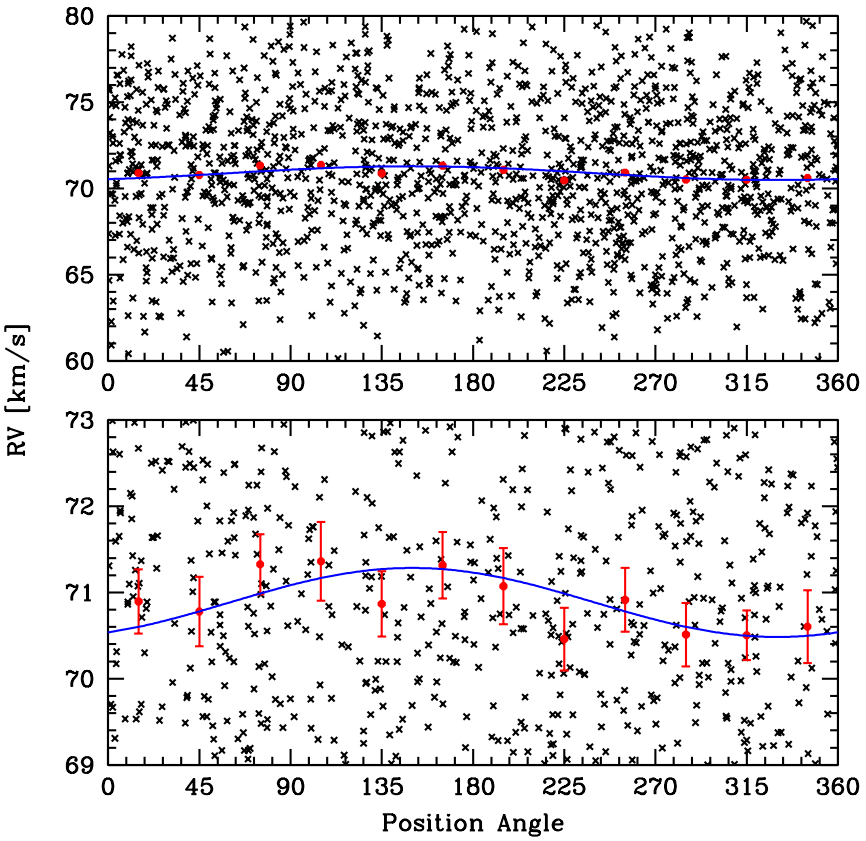

Figure 24: Radial velocity as a function of Position Angle for star closer than 5′ to the cluster center. In the upper panel all stars are shown, and in the lower panel we have zoomed vertically to display only the range 69km/s $\geqslant V_i \leqslant$ 73km/s. Position angle has been divided in bins of 30°, red points represent the mean radial velocity of the stars included in each bin. A sinusoidal fit with amplitude $-0,4$km/s and phase 120° minimize the dispersion around the red points.

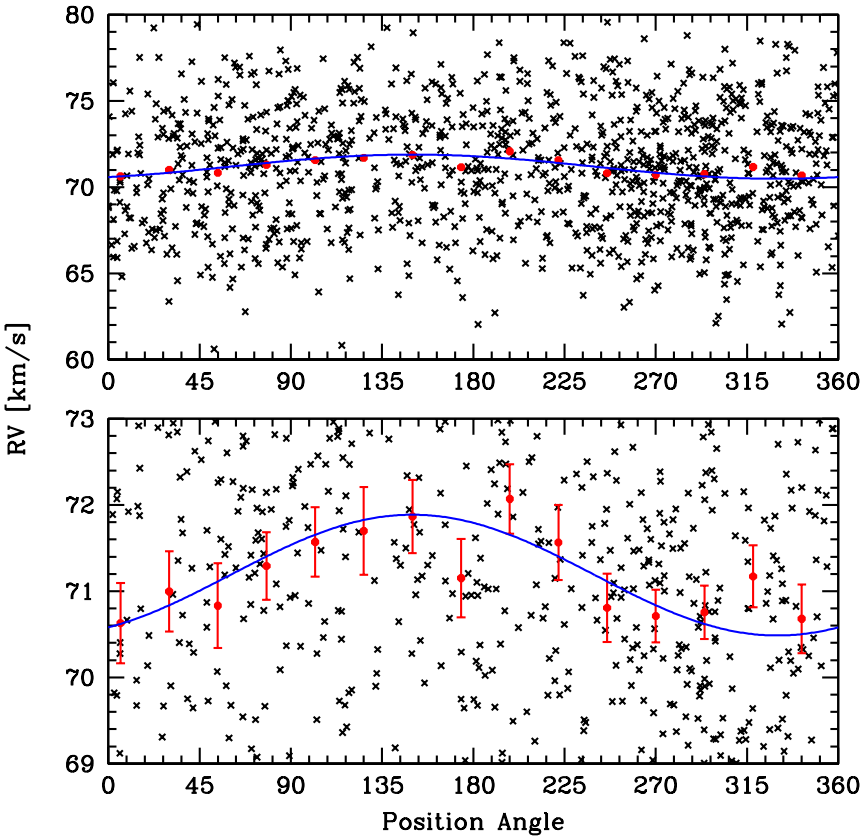

Figure 25: Same as Figure 24 but for stars at $d > 5'$. Stars has been divided in bins of 24°. A sinusoidal fit with amplitude $-0,7$km/s and phase 120° minimize the dispersion around the red points.

3.6 CONCLUSIONS

We have presented a new approach for the determination of accurate wavelength dispersion solutions for not-stabilized instruments. We have developed an algorithm for improved corrections for radial velocity shifts using simultaneous calibration fibres. Several precautions have to be taken when dealing with low SNR spectra covering a short wavelength range and strong absorption features, in particular during sky spectrum subtraction and continuum normalization. Together with a careful choice of the radial velocity technique that better suites the characteristics of our spectra, we have been able to significantly increase the precision of the radial velocities respect to existing tools. Algorithms have been optimized to require only a minimal user interaction, and they have been tested on a large dataset of M4 with already state-of-the-art radial velocity measurements in order to assest the improvement in precision. A total of 7250 individual spectra for 2771 stars, gathered with the GIRAFFE spectrograph at VLT during a multi-epoch campaign, have been analyzed. The derived cluster radial velocity is $\langle v \rangle = 70.896 \pm 0.077$km/s, in agreement with previous measurements in literature but with an offset of $\Delta v = 0.63$km/s with the previous value derived using the same dataset, probably due to the different radial velocity determination technique. We increased the number of binary candidates to 96, 30 giants and 66 sub-giants and dwarfs, almost doubling the number found by the previous analysis using the same dataset and the same statistical approach. An average radial velocity dispersion of 4.5km/s within 3 arcsec from the center of cluster and steadingly decreasing outward has been found, in contrast with the nearly constant value of 3.5km/s from Peterson et al. [1995]. Finally a preliminary study on the cluster rotation has been conducted, confirming a rotation of amplitude -0.70 ± 0.2km/s and phase of $120 \pm 10^\circ$ for stars farther than $5r$; a weaker sign of rotation (amplitude -0.40 ± 0.2)km/s) seems to be present in the inner sample too, but a deeper and more detailed analysis is required.

4

STELLAR ATMOSPHERE PARAMETERS FOR M 4

Galactic globular clusters contain $10^4 - 10^6$ stars, yet until the 1970's they generally were considered to be simple systems with members born at one epoch from chemically homogeneous gas. Their actual complexity has been revealed as photometric and spectroscopic data have steadily increased in quality and quantity. We now know that in most (all?) globular clusters (GCs) there are large star-to-star variations in all light elements that are susceptible to proton-capture fusion reactions (Li, C, N, O, Na, Mg, Al). Several clusters have internal variations in the heavy neutron-capture elements (e. g. , Y, Zr, Ba, La, Eu). The large number of papers that have been contributed to these conclusions have been discussed in several major reviews, e. g. , Kraft [1979], Freeman and Norris [1981], Kraft [1994], Gratton et al. [2004].

Recently attention has been drawn to the discovery of multiple color-magnitude sequences in several GCs, for example NGC 2808 [Piotto et al., 2007], NGC 1851 [Milone et al., 2009], and 47 Tuc [Anderson et al., 2009]. These split evolutionary sequences have been interpreted in followup spectroscopic studies to be due to metallicity variations (e. g. , M22, Marino et al. 2011) and helium abundance variations (e. g. , ω Cen, Piotto et al. 2005; NGC 2808, Bragaglia et al. 2010). An increasingly popular interpretation of the metallicity and abundance ratio inhomogeneities is in terms of multiple stellar generations with individual clusters (e. g. , Carretta et al. 2010). But many of the claimed abundance variations are not much larger than the analytical uncertainties, so it is essential to attack these questions with large spectroscopic samples of high quality.

M 4 (NGC 6121) is a bright, well-populated, mildly metal-poor Galactic GC. It is usually regarded as the closest GC to the Sun, e. g. , 1.8 kpc [Dixon and Longmore, 1993], 1.7 kpc [Peterson et al., 1995] and 2.2 kpc [Harris, 1996][1]. The closeness of M 4 makes its upper red-giant branch (RGB) accessible to high-resolution spectroscopy with medium-large telescopes: at the red-giant tip, $V \approx 11.8$, and at the horizontal branch (HB) $V \approx 13.5$. However, M 4 lies between the Sun and Galactic center and consequently suffers significant interstellar extinction, $A_V \sim 1.2$ mag. Additionally, this extinction is variable over the cluster face (Cudworth and Rees 1990, Ivans et al. 1999, Marino et al. 2008), and so transformations of colors and magnitudes into luminosities and temperatures are not trivial for M 4. So far, M 4 has not been identified as a multiple-generation GC.

1 http://physwww.physics.mcmaster.ca/ harris/mwgc.dat

Comprehensive high-resolution spectroscopic investigations of atmospheric parameters, metallicities, and abundance ratios in M 4 red giants have been conducted for several decades by several author groups, including Geisler [1984], Gratton et al. [1986], Brown and Wallerstein [1992], Drake et al. [1994], Ivans et al. [1999], and Marino et al. [2008]. The metallicity of M 4 now appears to be well determined at [Fe/H] ≈ -1.15 with a range from individual studies of $\approx \pm 0.15$ (Kraft and Ivans 2003, Gratton et al. 2004, and references therein). But this cluster has also been studied extensively in an attempt to understand its CN bi-modality (Norris 1981, Smith and Briley 2005 and references therein). Therefore there have been several high-resolutions spectroscopy studies that have focused on star-to-star the behavior of more limited element groups: Na, Al, and O, Drake et al. [1992]; Li, D'Orazi and Marino [2010]; Na in HB stars, Marino et al. [2011]; K, Takeda et al. [2009]; neutron-capture elements, Yong et al. [2008a,b].

During the last years several clusters have been the target of a spectroscopy survey with the Very Large Telescope (VLT) to study their internal velocity dispersion conducted by our group ($71.D-0205(A)$ and $72.D-0742(A)$ ESO Proposal, P.I. Piotto). These spectra have been taken specifically to measure radial velocities and at first glance they may seem not suitable for a chemical analysis: the wavelength range is short ($\leqslant$ 214Å), the spectra are noisy (SNR$\geqslant$ 10) and many lines are blended ($R \simeq 24000$). These spectra still contain much more information than just velocities, and given their number (several thousands of spectra for cluster) and the targeted stars (from the Upper Red Giant Branch down to the Main Sequence) we decided to try new approach to see what else about M 4 might be extracted. Here is a partial list of questions we hoped to answer with these data.

- Can temperatures and gravities derived from spectra be mapped reliably onto the color-magnitude quantities? Are the scatters in these relationships be understood in terms of the uncertainties in both sets of quantities?
- Is the derived metallicity the same at the RGB tip and the MS? Again, is the scatter compatible with the uncertainties?
- Can efficient, physically defensible, and sensitive automated algorithms be developed for these data that have general applicability to other cluster data sets?
- Is it possible to get interesting abundance ratios from these spectra?

In this work, we will focus on the parameter atmosphere determination of M 4 stars.

4.1 OBSERVATIONS AN DATA REDUCTION

The spectra analyzed in this work have been originally obtained by Sommariva et al. [2009] for an investigation of the internal velocity dispersion and binary fraction of M 4. The data were gathered with the VLT Fibre Large ArrayMulti Element Spectrograph (FLAMES; Pasquini et al. [2000]) employed with the GIRAFFE medium-high resolution spectrograph in MEDUSA multi-fiber mode. In this configuration one can obtain single-order spectra for 132 objects (target stars and sky) in each integration.

The GIRAFFE HR9 setup has been chosen, which yield spectroscopic dispersion of 0.05 Å/pixel, measured 4-pixel resolving power $R \equiv \lambda/\Delta\lambda \simeq 25.800$, in the wavelength range $5143\text{Å} < \lambda < 5356\text{Å}$. This spectral region is characterized by a large number of lines, many of them blended, and the presence of the strong MgI B triplet.

The M 4 target stars have been selected from an astrometric and photometric catalog based on Wide Field Imager (WFI) data from the ESO/MPIA 2.2m telescope [Sommariva et al., 2009]. The original selection criteria was that each star had no neighbors with $V_{neighbor} < V_{target} + 2.5$ within an angular distance of 1.2arcsec, being the fiber radius 0.6arcsec. Being a radial velocity survey in search of spectroscopic binaries, almost all the stars have multi-epoch observations.

For our analysis, a photometric catalog of NGC6121 has been kindly provided by Prof. Stetson (private communication) as part of a comprehensive catalog of homogeneous photometry for Milky Way clusters and dwarf spheroidal. The basic description of that project is reported in Stetson et al. [1998]. Photometry for over 70,000 M 4 stars in broadband U, B, V, R, and I was obtained in a multi-year effort involving many telescope/instrument combinations, as summarized in D'Antona et al. [2009]. Nearly all of the stars have both B and V measurements, about 76% also have R measurements, but only about 43% were observed in the I passband. Photometry of a subset of nearly 800 M 4 stars has been made publicly available as part of Stetson's "Photometric Standard Fields"[2], and most of those stars are included in our large catalog. The magnitudes given in the public database are consistent with those that we are using.

Spatial properties of our target stars are illustrated in Figure 26, where the targets are color-coded according to the number of exposures available for a given stars. The cluster core radius (where the surface luminosity has decreased by a factor of two from the center) and the half-light radius are also drawn in this panel, indicating that essentially all of M 4 has been sampled spectroscopically. All the observed stars are inside the tidal radius of the cluster. Photometric characteristics of the stars are shown in Figure 27, here again color-coded depending on the number of observations. The two panels shows the

2 http://www1.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/community/STETSON/standards/

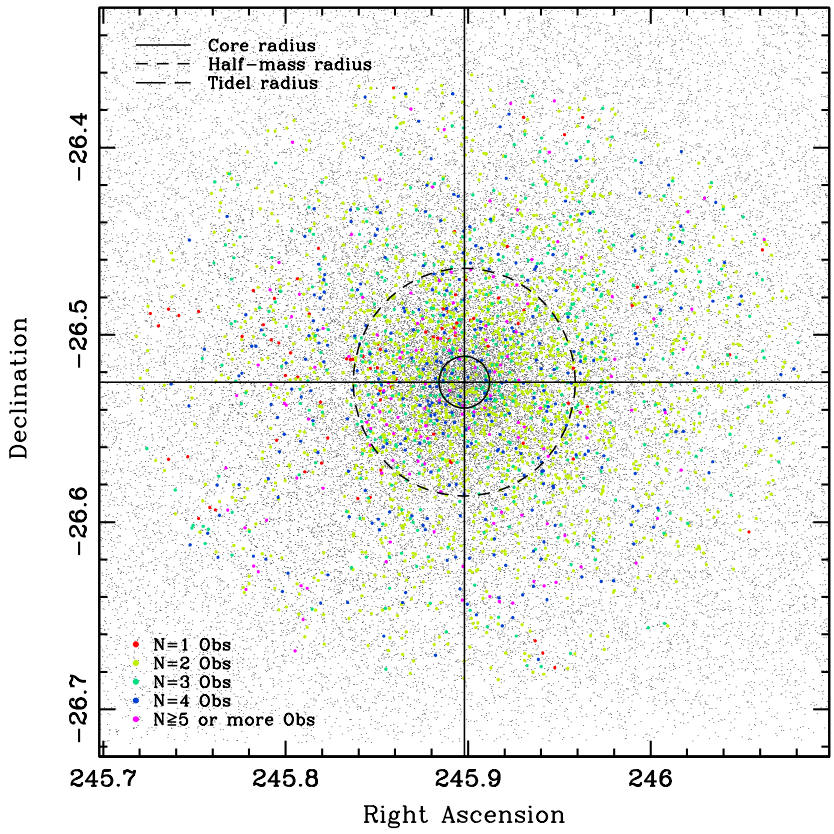

Figure 26: The spatial distribution of our targets. The whole sample is inside the Tidal Radius of the cluster. Targeted stars are color coded according to the number of available exposures.

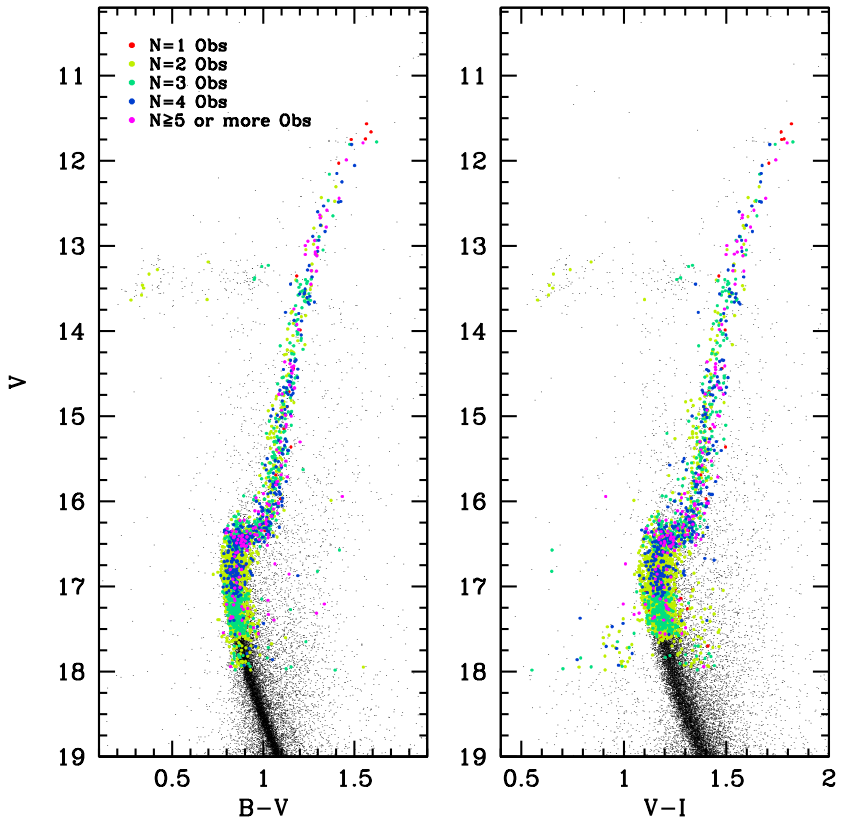

Figure 27: Photometric characteristics of our target stars, superimposed to the Color-Magnitude Diagram of the cluster. Targeted stars are color coded according to the number of available exposures.

color-magnitude diagram in $V/B-V$ (right-hand panel) and $V/V-I$ (left-hand panel) before any correction for reddening. The limiting magnitude of the targets, $V \lesssim 17.5$, has been set by the requirement that a single M 4 integration had a Signal to Noise Ratio SNR > 10 for each star in a single observation. Although Horizontal Branch stars are present in our dataset, they have been excluded from the analysis.

A total of 2771 stars covering color-magnitude diagram positions from the upper red-giant branch to about one magnitude fainter than the main-sequence (MS) turnoff (TO) luminosity have been observed. Observations span a temporal period of six years, and we refer to Sommariva et al. [2009] for a description of the observation strategy. We added 306 new spectra obtained in 2009 and targeting MS stars already observed in the previous epochs. Since these spectra originally were gathered to determine the M 4 velocity dispersion and to assess the binary-star fraction, nearly all stars have been observed at least twice, and three or more spectra have been obtained for nearly 40% of the sample. We summarize the observation numbers in Table 1; a total of 7250 individual spectra have been used in our study.

Sommariva et al. [2009] used the ESO Giraffe standard reduction pipeline [Blecha et al., 2000] to reduce the M 4 data from raw CCD exposures to final extracted, flat-fielded, wavelength-calibrated spectra. We have decided to revisit the reductions, paying particular attention to spectral extraction and wavelength calibration, as shown in the previous chapters.

Degradation of the SNR with magnitude is shown in Figure 28 by taking four example spectra of stars with different magnitudes. These spectra have been chosen among the best exposed during observing runs without moon illumination. Other spectra at the same magnitude can have lower SNR depending on exposure times and observational conditions.

The chosen lower limit of SNR > 10 is sufficient for a good determination of a star's radial velocity. However, atmospheric stellar parameters determination with such a low relative flux level requires very careful handling of the spectra. After many numerical experiments with our data, we adopted a new approach to handle the spectrum preparations. Specifically, after extraction and wavelength calibration of the raw spectra, we decided not to perform any subsequent data corrections. We made this choice in order to avoid introduction of any false correlation into the spectrum noise. All the steps that usually lead to sky-corrected, rest-frame, continuum-normalized stellar spectra were instead performed in the opposite direction: deliberate degredation of the *perfect* synthetic spectra to best match the real extracted spectra. This procedure is explained in details in Section 4.4.

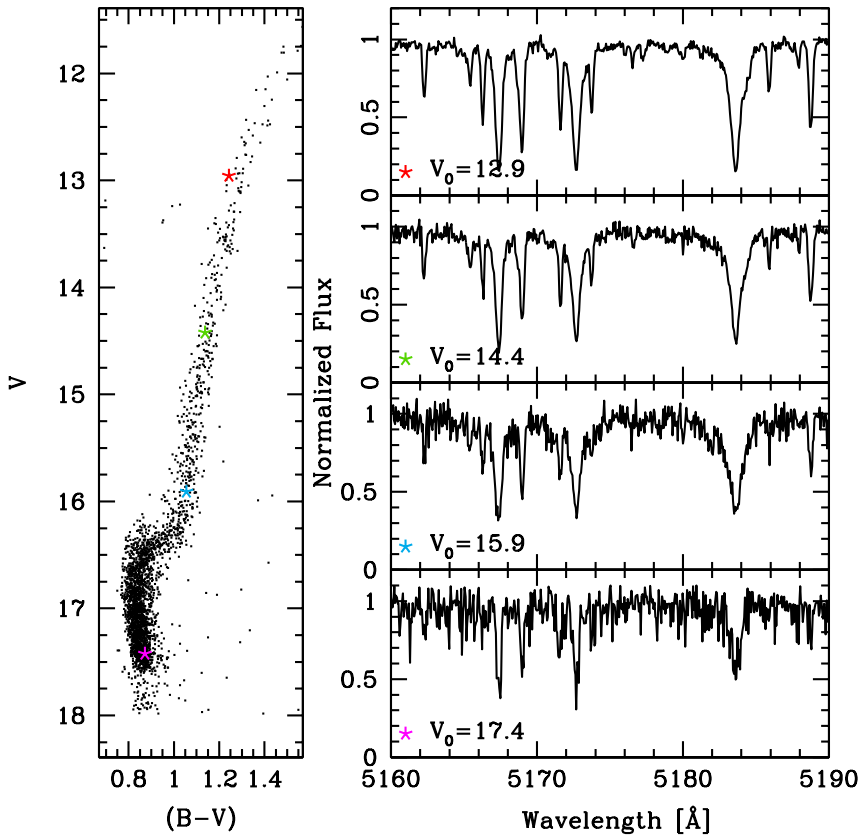

Figure 28: Degradation of the SNR with magnitude is shown by taking four example spectra of stars with different magnitudes. These spectra have been chosen among the best exposed during observing runs with good atmospheric conditions and without moon illumination.

4.1.1 *Sky flux determination*

For every GIRAFFE exposure on the M 4 targets, about 5-10 fibers have been allocated to empty sky regions. Sky subtraction is usually performed by coadding the spectra from these fibers and subtracting the result from the observed spectrum. There are several drawbacks when doing so:

- These spectra are intrinsically very noisy, having the sky a maximum brightness of $\simeq$ 17.5mag/arcsec^2 at 30° from full Moon. Even when coadded, the resulting spectra is still very noisy, then increasing significantly the noise of the sky-subtracted spectra especially in those cases where star and sky fluxes are comparable.
- Each fiber has a different transmission efficiency, the error in its determination is propagated when the sky spectra have to be coadded, and when the coadded spectrum is subtracted from the star spectra.
- Sky spectra must be rebinned to a common wavelength scale to be coadded, and then the coadded spectrum or the star spectra must be rebinned to the wavelength scale of the other spectra. In any case, the rebinning of a low SNR spectrum can introduce correlated noise and should be avoided.
- several fibers must be allocated for the sky spectra, reducing the efficiency of a multiplexing instrument such as GIRAFFE.
- even if the allocated sky fibers are scattered around the field of view, the coadded spectrum wil not be sensitive to any spatial variability of the sky and the individual spectra are too noisy and their number too low to determine it over the $\simeq$ 500arcmin^2 covered by a MEDUSA plate.

Since many of the observing runs occurred in gray/bright times, the sky spectrum constitute a significant (and constant) part of the flux. An example is given in Figure 29 where observed spectrum of a relatively faint program star ($V = 17.73$) is decomposed into the stellar (blue line) and the sky (red line) contributions. The spectrum in the figure was taken on 2006 September 04, $\simeq$ 5 days before full moon and at a distance of $\simeq$ 40deg. Analyzing the coadded sky spectrum we noticed that the sky light was composed of two distinct parts: one resembling the solar spectrum, due manly to the presence of the moon, and another featureless *gray* one (yellow line). This gray feature has been first noticed in sky spectra taken during dark time, i. e. without moon illumination, and confirmed to be necessary when trying to fit the coadded sky spectra from moon illuminated nights with the solar spectrum.

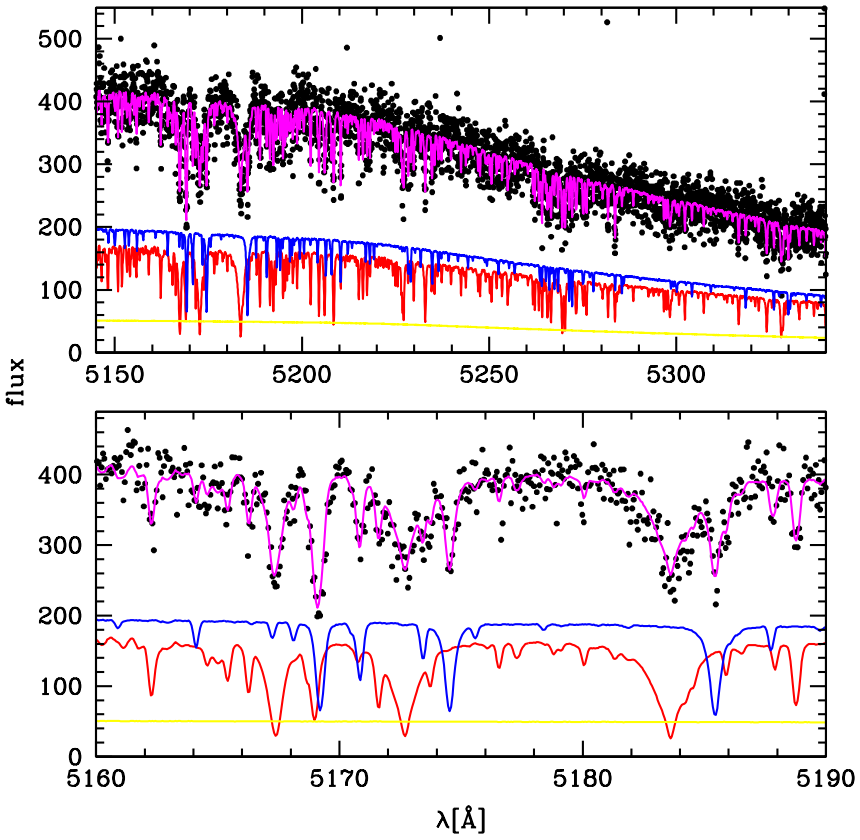

Figure 29: Observed spectrum of a star with $V = 17.73$ and a few days before to the full moon. Synthetic sky spectrum (red line), stellar spectrum (blu line) and a grey continuum (yellow line) are combined (magenta line) and compared to the observed stellar spectrum (black points). The lower panel shows the Mg b triplet region, to highlight the contribution of the sky spectrum to the observed total star+sky spectrum.

There is clear evidence of sky contamination in this star's low SNR spectrum, especially in the strong Mg b lines (lower panel). We remark that in our spectral range no atmospheric absorption lines are present, so no correction is needed.

The inclusion of the sky in the correct proportion with respect to the stellar flux is then a fundamental part for a good synthesis match of the data. To determine the sky flux level, we proceed in the following way. The extracted and wavelength-calibrated sky spectra are corrected for fiber transmission and flat-field normalization. The amount of sky flux for unit wavelength is determined by integration the sky spectra over 2 Å intervals; this value has been chosen as a good compromise between reducing the noise and preserving the main spectral features of the spectra.

In order to account for the loss of flux due the spectral lines, the same calculation is performed over the same wavelength intervals on a normalized Kurucz solar spectrum smoothed accordingly to the spectral resolution of the instrument.

The values determined from the observed sky are fitted with a linear combination of two functions: a low-order polynomial multiplied for the values from the solar model, to take into account the moon illumination, and a constant for the gray continuum. The resulting functions represent the continuum flux per unit wavelength as a function of the wavelength respectively of the moon illumination and the gray continuum level. The former function is split in two parts, a constant representing the average flux of the sky, and a normalized function for the shape. We will refer to them as *flux constant* and *continuum function* respectively. Since the gray continuum is a constant, its continuum function is simply a unitary constant value.

The coefficients of the functions as derived above represent only the first step in the sky flux derivation. During the continuum determination process, described in the next sections, the flux constants are refined using the sky lines in the spectrum as a proxy, to take into account the sky variability and fiber transmission differences.

Precautions to ensure that the extra absorptions in faint stellar spectrum are truly sky contributions and not accidental inclusion of some faint non-cluster member star in the fiber apertures have been taken. Not only the mean sky radial velocities are consistent with a rest velocity system, but their fluxes positively correlate with lunar phase. We have also rejected individual sky spectra that had anomalously high flux (indicative of accidental pointing at a faint star). This has convinced us of the real nature of the sky spectra.

4.1.2 *Continuum Normalization*

Continuum rectification has proved to be challenging due to the restricted total wavelength range (5150–5360 Å) of our spectra and to

the presence of the strong Mg b absorption lines in the first 50 Å of the spectra. The GIRAFFE HR9 setting has been deliberately chosen for the original Sommariva et al. [2009] study due the line richness of this spectral region. In addition to the Mg b lines there are many strong FeI lines, which makes this wavelength domain particularly attractive for radial velocity studies. But this wealth of absorption lines creates many wavelength intervals of overlapping spectral features, so defining an adequate number of points for a *global* continuum normalization is a very difficult task. Local normalization, usually determined by visual inspection when measuring equivalent widths, is also unsuitable for two reasons: the large number of spectra in our sample; and the subjectivity in manually selecting the continuum (which probably would have introduced a systematic bias in our lower SNR spectra). Spectral synthesis requires good normalization in the whole wavelength range of interest, and while is still feasible to manually adjust the continuum to match the synthesis when the number of stars is low, for our very large sample we have to find an alternative way.

A first estimate of the continuum stellar flux per unit wavelength is determined in a similar way as done for the sky flux. The sky and grey continuum contributions are removed from the observed spectra. Then for each star we adopt a synthetic spectrum with photometric atmosphere parameters (Section 4.2) and metallicity $[\mathrm{Fe/H}] = -1.15$ to calculate the *flux constant* and *continuum function* relative to the stellar component of the spectrum, following the same steps described in Section 4.1.1 for the sky continuum.

The synthetic stellar spectra and the solar spectra are degraded to match the spectral resolution of the instrument and together with the unitary function for the gray continuum they are rebinned into the wavelength scale of the science spectrum, keeping constant the flux per wavelength unit. All spectra are then multiplied for the corresponding continuum flux function and the resulting spectra are linearly combined, using the flux constants as coefficients. The instrumental effects (e. g. multiplication for flat-field) are finally applied to make the comparison with the observed spectrum possible.

The flux constants are finally determined by fitting the resulting model spectrum to the observed one by χ^2 minimization (Equation 19), in the same way as we do to determine the atmosphere stellar parameters (Section 4.5). Small adjustments in the continuum functions also are allowed, to deal with spectral regions affected by the wings of strong spectral lines, e. g. in the proximity of the MG b triplet.

A binary mask is included in the χ^2 determination to exclude points too sensitive to variations of Temperature, Gravity, Micro-Turbulence and Metallicity. This mask is derived by taking two boundary values for each stellar parameter (i. e. the photometric value $\pm$ its estimated error) and flagging those points that show a percentage variation

over a fixed threshold. A detailed explanation is given in Section 4.5. Points that have normalized flux lower than 0.8 in the stellar synthesis are excluded too, while the lines in the sky spectrum are retained to allow a better determination of the sky flux level.

The flux continuum functions and the flux constants are adjusted as described above for each set of stellar atmosphere parameters used during the χ^2 determination, so that the derived values will not be influenced by the initial choice in the photometric parameters for the continuum determination. A good starting point however allows a quick recalibration of this function.

Radial velocities shifts due to the intrinsic velocity of the star and the Earth motion are applied to the stellar synthetic spectrum. No rebinning of the observed spectra is necessary, and no noise is introduced since all the operations are performed on noise-free high-resolution synthetic spectra.

4.2 STELLAR ATMOSPHERIC PARAMETERS FROM PHOTOMETRY

In every minimization algorithm, a good starting point always results in a faster determination of the variable of interest. Spectral synthesis is a very lengthy process by itself and with four parameters to be determined by χ^2 minimization and several thousand of stars to be analyzed, this task can be excessively time-consuming. Greater speed in this process can be provided through photometrically derived stellar parameters, but some precautions are necessary.

M 4 is known to have both a large mean reddening $E(B-V) = 0.34$ [Harris, 1996] and a large differential reddening across its surface. We use the latest determination derived by Hendricks et al. [2012] and listed in Table 4

Distance	1.80 ± 0.05kpc
A_v	1.39 ± 0.01
$(M-m)_0$	11.18 ± 0.06
E_{B-V}	0.37 ± 0.01
$E_(V-I)$	0.53 ± 0.01
R_v	3.63 ± 0.07

Table 4: Reddening parameters for M4 [Hendricks et al., 2012]

We have corrected the cluster mean $E_(B-V)$ for differential reddening using the technique described in Milone et al. [2012]. The reddening direction in the CMD is defined as $\theta = \mathrm{atan}(A_V/E(B-V))$. A fiducial line along the RGB and MS is determined and then for every star the median shift of the surrounding stars respect to the fiducial line in the reddening direction is taken as indicative of the amount of the differential reddening, taking advantage of the fact that stars in a local spatial region of M 4 are all shifted by the same amount. Finally,

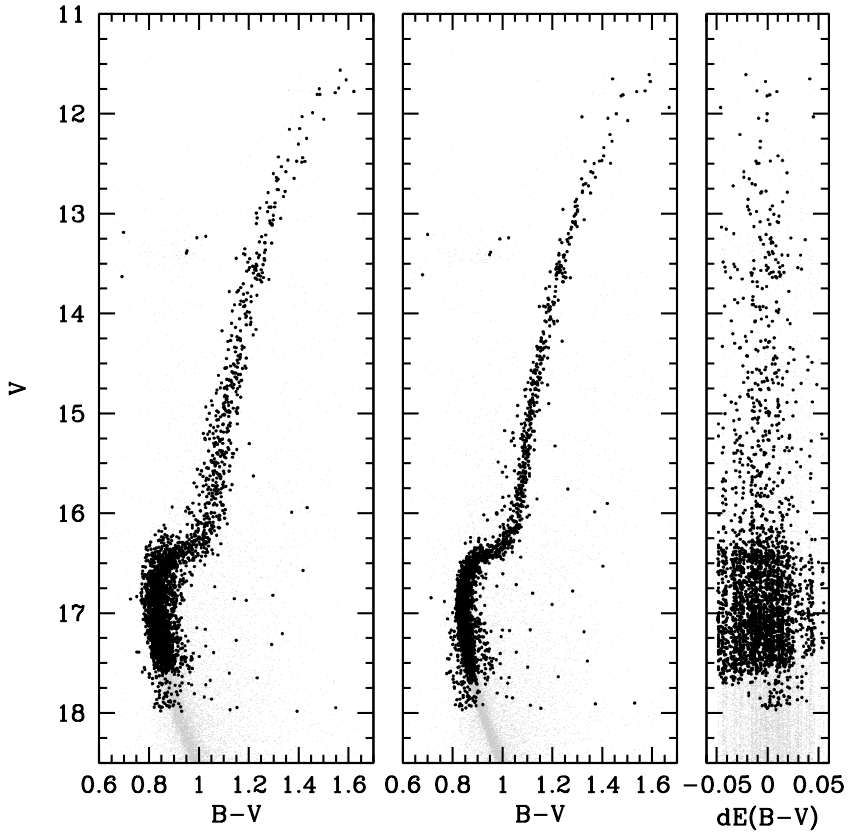

Figure 30: The $V/(B-V)$ diagram before (first panel) and after (second panel) the differential reddening correction, with the target stars highlighted. The third panel shows the amount of the correction. Reddening correction has not been applied.

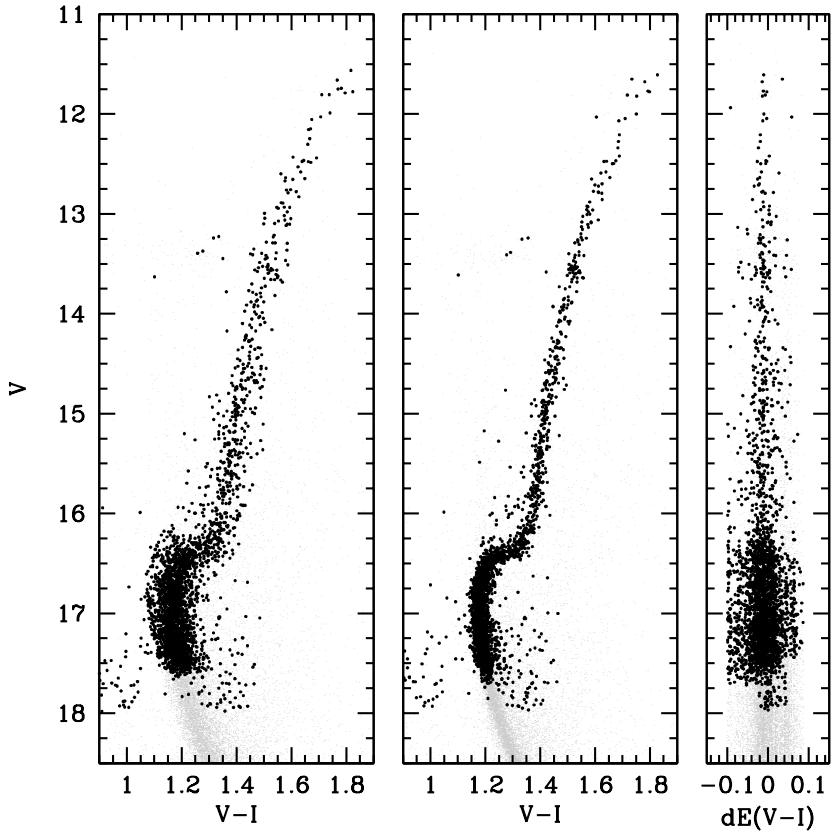

Figure 31: As in Figure 30, for the $V/(V-I)$ plane

we have adopted the values cited above to convert $dE(B-V)$ to dA_V values. In the same way the $V/(V-I)$ diagram has been corrected for differential reddening, to obtain a second color index whose correction is independent from the first one. Although one photometric band is in common, the two reddening corrections can still be considered independent since we are correcting an effect in color. The differential reddening corrected CMDs in the two color combinations for our target stars are shown in Figure 30 and Figure 30, along with the amount of correction as a function of magnitude. The overall cluster reddening correction (Table 4) has not been applied here yet.

The two indexes have been used to obtain two estimates of the (photometric) stellar atmosphere parameters, using the calibrations described in the next subsections. Although very precise, we must keep in mind that these value will be used only as input values for our spectroscopic determination.

4.2.1 *Effective Temperatures*

For every star two independent values for effective temperatures are obtained from $(B-V)$ and $(V-I)$ colors using the relations from Ramírez and Meléndez [2005] for giant stars and Casagrande et al. [2010] for subgiants and dwarfs stars. The switch between the two relations has been set at $V^{switch} = 16.1$

The differences in derived effective temperature using the two colors are shown in Figure 32. The upper panels show a direct comparison between the derived temperatures using the two colors, for giants stars on the left-hand and dwarfs on the right-hand. The threshold of the switch has been chosen by looking where the two temperatures begin to differs systematically, i. e. the points starts to deviate from the red lines. The lower panel shows the difference of the color-derived temperatures respect to their mean, which is the value that will be used to derive the other stellar parameters. A gap around $T_{eff} \simeq 5200K$ is clearly visible:Casagrande et al. [2010] report a systematic difference of $+ \simeq 100K$ with the calibration of Ramírez and Meléndez [2005] for dwarf stars, and probably such difference is affecting also the calibrations for giant stars. We did not attempt to correct this difference, letting the parameter determination algorithm solve it.

It is interesting to see the dependence of temperature on small variations of colors, as given by the calibration. The variations in temperature as a function of $d(E-B)$ and $d(V-I)$ can be used as the correction in temperature due to the differential reddening correction, or as the error associated to Temperature for a given error in color. It is interesting to notice that the $(V-I)$ color has a larger error than $(B-V)$ for a given star, but usually the reddening correction is more reliable when using passbands close to the infrared. An aver-

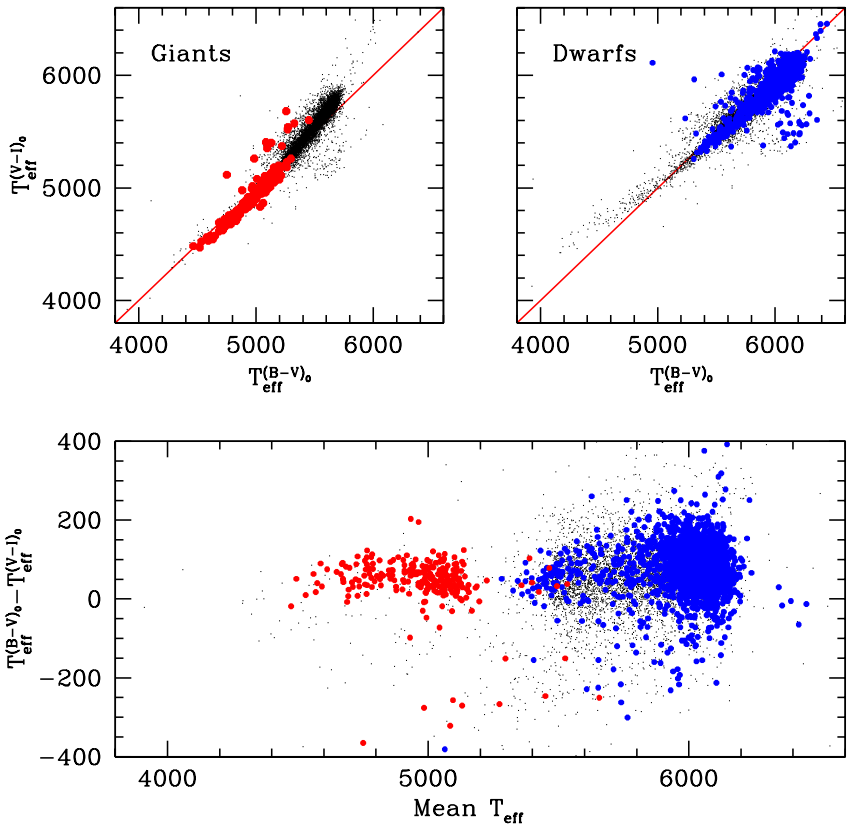

Figure 32: Differences in derived effective temperature using the colors $(B-V)_0$ and $(V-I)_0$. The upper panels show the temperature differences when using the calibration for giant (left-hand panel) and dwarf (right-hand panel) stars. Switching the calibration at $V = 16.1$ provides the lower scatter in the difference when considering giant (red points) and dwarf (blue points) stars together (lower panel).

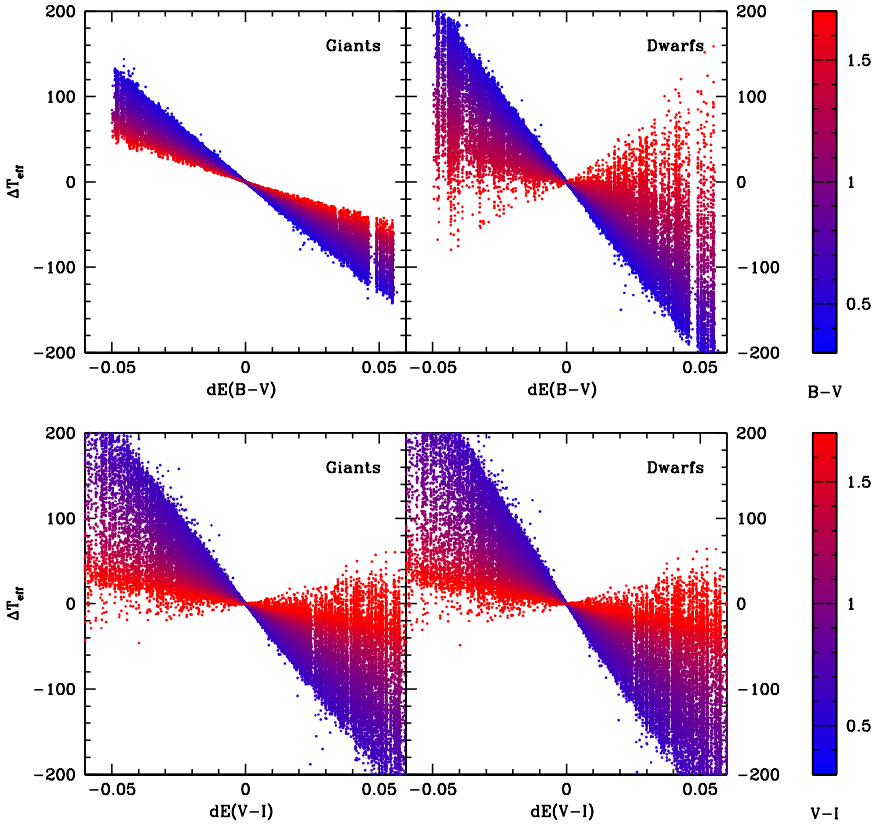

Figure 33: The variations in temperature as a function of $d(E-B)$ and $d(V-I)$ can be used as the corrections in temperature due to the differential reddening correction, or as the error associated to Temperature for a given error in color.

age error of $T_{eff} = 150K$ has been estimated for photometric derived temperatures.

4.2.2 *Surface Gravities*

Photometric gravities are calculated from the classical relation:

$$\log g^{\star} = \log g^{\odot} + 4 \log \frac{T_{eff}{}^{\star}}{T_{eff}{}^{\odot}} + 0.4(M_{V,*} + BC - M_{bol,\odot}) + \log \frac{M_*}{M_{\odot}} \quad (16)$$

where we have adopted a mean distance of 1.8kpc to relate the reddening-corrected relative magnitude to the absolute one in the relation. A mean mass of $M_{\star} = 0.85 M_{\odot}$ has been assumed. For every star the bolometric correction is obtained interpolating the table from Houdashelt et al. [2000] for the values of temperature and gravity of the star and iterating over $\log g_*$ until convergence is reached. The solar values $T_{\odot}$=5770K, $\log g_{\odot}$=4.44, $M_{bol\odot}$=4.75 are adopted.

4.2.3 *Micro-turbulent Velocities*

Initial values of micro-turbulent velocities for stars cooler than 5000K are derived using the relation from Marino et al. [2008]:

$$\xi_t = -0.254 * \log g + 1.930 \quad (17)$$

This relationship has been derived from giants stars in M 4, and it may not be suitable for dwarf stars. For dwarf stars we initially used the relation from Gratton et al. [1996]:

$$\xi_t = 1.190 * 10^{-3} * T_{eff} - 0.900 * \log g - 2.00 \quad (18)$$

The two calibrations have similar values at $\simeq$ 5100K, or $\log g =$ 3.1dex

As discussed later, several problems have been encountered while determining the micro-turbulent velocities of dwarf stars. We have finally decided to use the velocities given by Equation 17, given the fact that these values are used only for in the first iteration of stellar atmosphere parameters determination.

4.3 SPECTRAL SYNTHESIS

The first ingredient for a good spectral synthesis is a list of spectral lines as accurate and complete as possible. All the atomic lines in the spectral range 5140Å − 5360Å have been retrieved from the VALD-2 database [Kupka et al., 2000]. The listed values for the oscillator strength $\log gf$ of each line are usually the results of laboratory

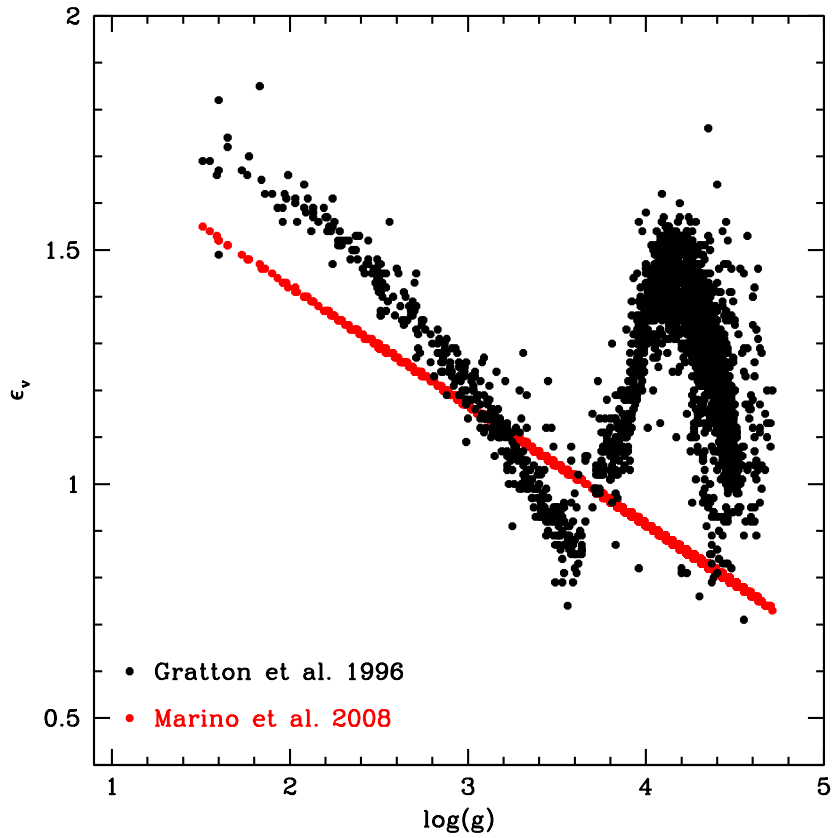

Figure 34: Calibration of micro-turbulence as a function of gravity from Gratton et al. [1996] (black points) and Marino et al. [2008] (red points) for our target stars. The first calibration is also dependent on temperature, hence the spread in the plot.

experiments or semi-empirical estimates. Calculated log gf's may be subject to large uncertainties, which fortunately can be corrected in numeric experiments with astrophysical spectra.

We have used the Solar Atlas by Kurucz et al. [1984] and the Atlas of Arcturus Spectrum by Hinkle et al. [2000] to improve the transition probabilities in our line list. For the Sun we have assumed the standard atmospheric parameters T_{eff}= 5777K , log g= 4.44, $\xi_t = 1.5$km/s , [Fe/H]= 0.00. For Arcturus there is no general consensus on the atmospheric parameters. We have decided to use the values listed by Massarotti et al. [2008] which are T_{eff}= 4325K , log g= 2.1, $\xi_t = 2.0$km/s , [Fe/H]= -0.60. The model atmospheres are generated by using an interpolator of Kurucz [1992] models calculated with the New Opacity Function Distribution (ODFNEW, Castelli and Kurucz 2004). The synthesis are generated using the current version of the Local Thermodynamic Equilibrium (LTE) code MOOG [Sneden, 1973]. Oscillator strength values have been modified to let the corresponding lines match iteratively the solar spectrum and the Arcturus one. When it was not possible to fine a satisfying value to match both spectra, the priority has been given to the Solar one, given the uncertainty in the stellar parameters and chemical composition of Arcturus. In some case it has been necessary to add some atomic lines with arbitrary parameters in order to reproduce the solar spectra in places where there is no known atomic or molecular identification. These *fake* lines have been chosen to be Fe I transitions with excitation energies of 3.5eV and gf-values adjusted to match the observed spectral features.

Figure 35 shows the improvement in the synthesis from before (red line) and after (blue lines) the refinement of log gf values. For the Sun the improved line list causes a reduction of the dispersion of the synthesis respect to observed points from $\sigma = 0.0507$ to $\sigma = 0.0147$, while for Arcturus we have a reduction from $\sigma = 0.0688$ to $\sigma = 0.0336$, as determined by MOOG.

The synthesis for M 4 stars have been generated using the Kurucz [1992] α−enhanced model atmospheres here again with the New Opacity Function Distribution (AODFNEW, Castelli and Kurucz 2004). The interpolator program has been modified so that the synthesis program automatically set the abundance of the α elements to the ones measured by Ivans et al. [1999], nominally $\log \epsilon(\mathrm{O}) = 7.68$, $\log \epsilon(\mathrm{Mg}) = 6.52$, $\log \epsilon(\mathrm{Si}) = 6.60$, $\log \epsilon(\mathrm{Ca}) = 5.12$, $\log \epsilon(\mathrm{Ti}) = 3.79$ for [Fe/H] $= -1.15$, with those values scaled accordingly with [Fe/H].

Since many stars in our sample share similar temperatures, gravities and micro-turbulent velocities, to speed up the determination of the atmosphere parameters we have created a grid of synthetic spectra with a spacing of 100K in temperature, 0.1 in gravity, 0.1km/s in micro-turbulent velocity and 0.05 in metallicity. The grid is built on the photometric atmospheric parameters derived for all the stars in the sample. For a specific set of parameters, the synthesis is computed

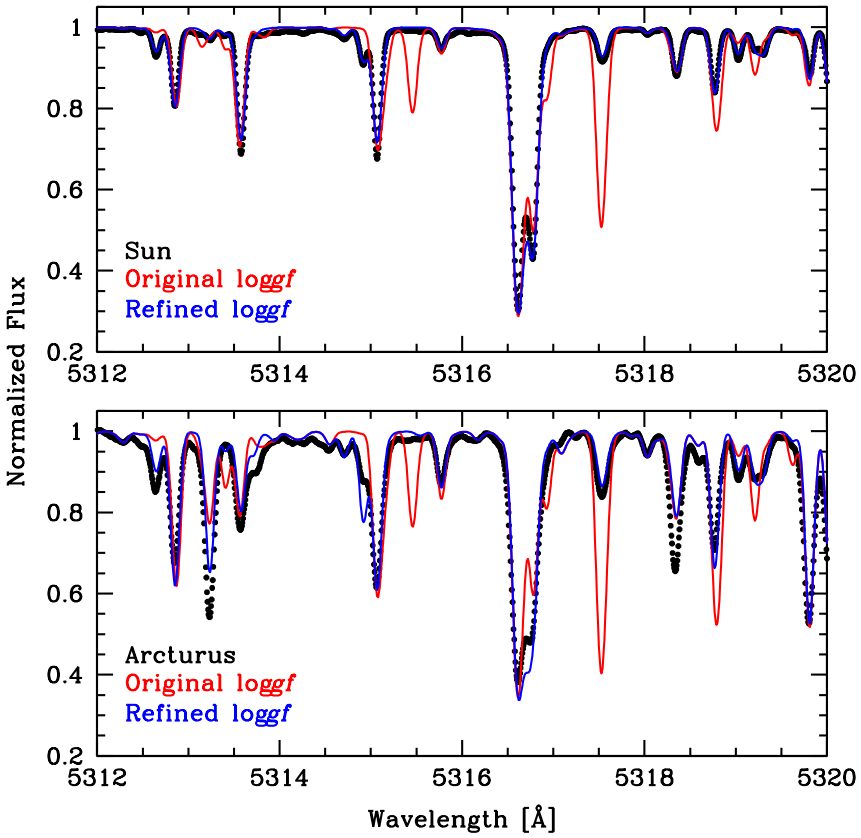

Figure 35: The improvement on spectral synthesis from before (red line) and after (blu line) the refinement of the log gf values in the input line list, compared with high SNR spectrum of the Sun (upper panel) and Arcturus (lower panel).

by linear interpolation of the syntheses with the closest parameters within the spacings of the grid.

4.4 COMPARING THE OBSERVED SPECTRA WITH THE SYNTHESIS

When determining the atmosphere parameters or the chemical abundances by the mean of synthesis, usually the observed stellar flux is processed in order to be comparable with the synthesis: sky contribute is subtracted, spectra are corrected by flat-fielding, normalized and finally rebinned in a linear wavelength scale. Any step in this procedure will produce an increase or a correlation in the noise that is usually negligible when working on spectra with a good SNR, but that can lead to a loss in precision in the atmospheric parameter determination when the SNR is very low like in our case.

We have decided to take the opposite direction to avoid any manipulation of the observed spectra. All the necessary steps to compare the theoretical spectra with the observed one are performed on the syntheses. In the following, we use the term *exposure* to refer to a single observation for a given stars. For every exposure of the star under study the synthesis to be compared is subject to these transformations:

- the synthesis is moved in the radial velocity space to match the measured total RV shift of the exposure (i. e. star heliocentric motion plus the Earth motion);
- both synthesis and solar spectra are degraded to the resolution of the instrument, determined from the Th-Ar exposures, using a Gaussian kernel;
- synthesis, solar spectra and gray continuum are rebinned in the wavelength scale of the exposure, keeping constant the amount of flux for unitary wavelength;
- the rebinned synthesis and sky spectra are multiplied for their continuum function, and together with the rebinned gray continuum they are linearly combined using as coefficient the corresponding flux constants (see Section 4.1.2 and Section 4.1.1).
- the resulting spectrum is multiplied for the flat-field relative to that exposure to include instrumental effects;

Star in Globular Clusters are known to be slow rotators, so the effect of rapid rotation in the synthesis before the second step are deliberately to not taken in account; nevertheless it would be very easy to incorporate this effect in the analysis, as well as determining the best fitting value during the χ^2 minimization. We call the result of these operations *processed synthesis* to distinguish it from the synthesis generated by MOOG. Since every exposure differs from the others due

to different observational conditions, variations in the instrument response, moon illumination etc. this process must be performed for every exposure of the star, and any exposure can be compared only with the relative processed synthesis.

4.5 STELLAR PARAMETER AND ABUNDANCE ANALYSIS

Suppose that we have a star with n_{obs} exposures, each one with n_{pixel} flux point $F(n,i)$ for the exposure n at the pixel i and associated wavelength value. For each set of atmosphere parameters we can calculate the χ^2 as given in Equation 19:

$$\chi^2 = \frac{\sum_{n=1}^{n_{obs}} \sum_{i=1}^{n_{pixels}} \frac{(F(n,i)-S_o(n,i))^2}{S_o(n,i)} p(n,i) w(n,i)}{\sum_{n=1}^{n_{obs}} \sum_{i=1}^{n_{pixels}} p(n,i) w(n,i)} \tag{19}$$

where $p(n,i)$ is a binary mask to include only points sensitive to the atmospheric parameter of interest, $w(n,i)$ is a weight function and $S_o(n,i)$ is the value of the processed synthesis, each of them relative to exposure n at the pixel i. S_o is a function not only of the stellar parameters, but also of the polynomial coefficients for the flux continuum, which are however determined univocally for each set of stellar parameters. The final set of atmosphere parameters will be the one that minimizes the χ^2.

ATMOSPHERE PARAMETER MASKS Although there are no lines affected only by a single parameter, every line reacts differently to a change of one of them. To determine which pixels are more sensitive to a specific parameter respect to the others, we created several masks applying sequentially the following scheme for each stellar parameter:

- two extreme values based roughly on the expected photometric errors are chosen for the atmospheric parameter under analysis.
- Two synthesis with the extreme values of the considered parameter and the others kept fixed to the photometric ones are calculated;
- When the difference between the two synthesis is more than 5% (10% for metallicity), the binary mask has a positive value, otherwise is null.

The range used for masks creation are $\Delta T_{eff} = \pm 150K$, $\Delta \log g = \pm 0.5$, $\xi_t = \pm 0.5$ and $[Fe/H] = \pm 0.5$.

The metallicity mask is actually used only to exclude from the continumm the points that are affected by a change in the overall metallicity (i. e. spectral lines). Although the continuum is being determined

at every synthesis comparison, we want to minimize as much as possible the influence of spectral features. The continuum mask has positive value only when all the other masks have negative.

An additional mask is created using the FeI line from the line-list provided by Ivans et al. [2006] and Sousa et al. [2011]: these lines have been selected for EW measurements in high resolution spectra (respectively HIRES and HARPS), and their astronomical $\log gf$ have been carefully determined. This mask is the one used used in the determination of the $[Fe/H]$ value.

Since the phometric values values of the atmospheric parameters are reasonably close to the ones finally derived in our analyses, we do not expect any significant change between the mask obtained with the photometric value and the one that we could obtain if the real atmosphere parameters were already known. It is still possible however to run the program a second time using a new approximation for the input stellar parameter, e. g. using new values for the interstellar reddening, since photometric values are calculated on the fly.

WEIGHT FUNCTION The weight function $w(n, i)$ is basically a correction factor to give more importance to the stellar lines with respect to the sky ones in the χ^2 determination. For every atmospheric parameter two extreme values are selected around the photometric value (nominally, $\Delta T_{eff} = \pm 200K$, $\Delta \log g = \pm 0.5$, $\Delta \xi_t = \pm 0.5$), a synthesis is generated for every combination of the parameters, and for every wavelength point in the synthesis the lower and higher values from all the syntheses are saved. Hereafter we refer to these two spectra simply as *minimum* and *maximum* spectra.

The two spectra are traslated to the observational space as explained earlier, including the addition of the sky background. We define a selection strip by adding the expected poissonian noise multiplie by two to the processed maximum spectrum, and subtracting the poissonian noise multiplied by three to the processed minimum one . The poissonian noise is computed using the overall (sky and star) continuum level. All the points outside the strip are excluded from the χ^2 computation, as well as data points flagged as unusable during spectral extraction, by giving them nell weight.

With this procedure, we are able to safely exclude undetected lines, anomalous pixel values (e. g. due to anomalous measurement of flux or inaccurate flat-field correction) and cosmic rays that survived to the optimal extraction or have not been flagged properly.

The weight is determined using Equation 20:

$$w(i, n) = \sqrt{2 - (S_{syn}^{max}(i, n) - S_{syn}^{min}(i, n))/2} \tag{20}$$

where S_{syn}^{max} and S_{syn}^{min} are respectively the maximum and the minimum spectra before being processed. With the weight defined in this way, more importance is given to the line belonging only to the stellar

spectrum while the global fit is preserved, resulting in an improved atmosphere parameter determination.

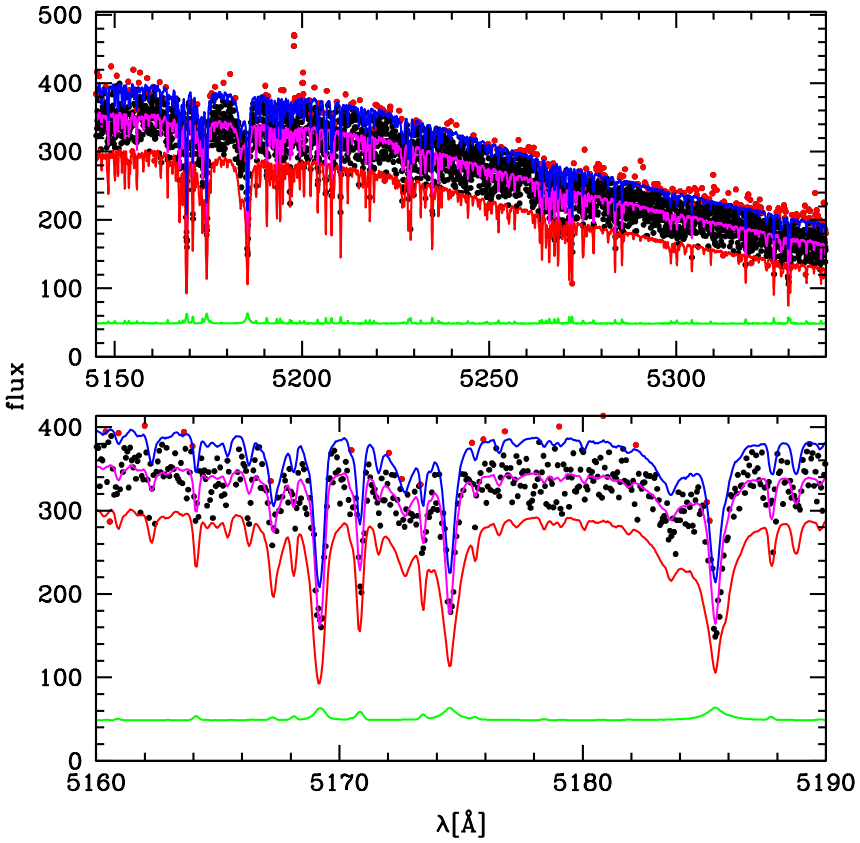

Figure 36: The same spectrum in Figure 29 is shown with the upper boundary (blue line) and the lower one (red) highlighted. Points that have been excluded from the χ^2 determination are red flagged. The green line represents the weight function $w(i, n)$, multiplied by a factor of 50.

In Figure 36 the selection strip is shown. Points excluded from the χ^2 determination have been red flagged. Green line shows the weight function resulting from Equation 20, multiplied of a factor of 50.

We decided to perform this selection before the χ^2 determination and not during the synthesis-observation comparison, i. e. using the photometric values as reference instead of the actual synthesis as a proxy, since the inclusion or esclusion of different points or a changing the value of the weights during the χ^2 determination can affect the final results.

χ^2 MINIMIZATION AND ATMOSPHERE PARAMETERS DETERMINATION Due to the complex interdependence of the atmosphere parameters, we have opted for a simple but robust method for the χ^2. Each iteration is characterized by a set of *input* values of T_{eff}, $\log g$, ξ_t

and [Fe/H], being in the first iteration the photometric value. Each parameter is varied individually while the other are kept fixed to the input value, and the value that minimize the χ^2 is saved in the *output* set. A new set is generated by adding to each of the parameters of the input set half of the difference between output and input set parameters, and the resulting values are taken as the new input value for the next iteration. The cycle is broken when the difference input and output set parameters are < 15K for temperature and < 0.05dex for the other parameters, with all these conditions to be satisfied simultaneously. If convergence is not reached, determination is considered unsuccessful and the star is dropped.

At every iteration of the algorithm a new set of atmosphere parameters is generated, and a synthesis with these parameters must be obtained to proceed with the χ^2 minimization. The synthesis is calculated by linear interpolation of the pre-existing grid of spectra introduced in Section 4.3. If the required syntheses are not included in the grid, they are generated at run-time and included in the grid. With this expedient only the synthesis needed by the program are created, optimizing the coverage of the parameters space against the time spent to create such grid.

4.6 DISCUSSIONS

The techniques described above are applied to determine the stellar atmosphere parameters T_{eff}, $\log g$, ξ_t and [Fe/H] of our 2771. Of these stars, 2570 have reliable radial velocities, as shown in Chapter 3, and only for 1910 of them have spectra of sufficient quality to determine the atmosphere parameters.

The color-magnitude of the stars with successful determination is shown in Figure 37 (left-hand panel). Stars for which the algorithm failed in finding the atmosphere parameters are marked as red. The trend of derived metallicity is shown in the right-hand panel. For stars with $V_0 < 14.7$ we have an average value for metallicity of $[Fe/H] = -1.068 \pm 0.001$ dex and a dispersion around the mean of $\sigma = 0.025$ dex. This value, derived using 322 stars along the RGB, is in good agreement with $[Fe/H] = -1.07 \pm 0.01$ dex determined by Marino et al. [2008] from 105 high-resolution UVES spectra using Equivalent Widths measurements down to $V_0 = 12.6$. This result also proves the correctness of our approach in continuum determination.

For stars fainter than $V_0 = 14.7$ the dispersion of the derived metallicity increases fairly abruptly to $\sigma = 0.18$ dex. This behavior can be better understood by analyzing the trend of the other stella parameters with color (Figure 38)

It can be seen that at higher magnitudes, i. e. at lower SNR, the derivation of micro-turbulent velocity becomes unreliable, with a huge dispersion that is not easily explainable from a physical point of view.

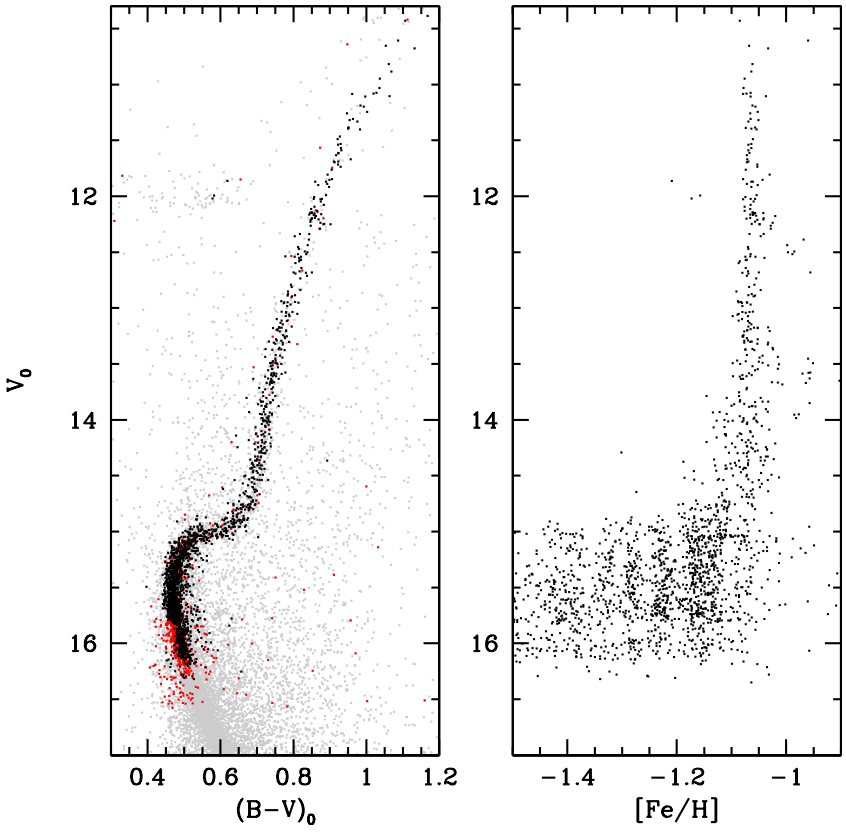

Figure 37: Metallicity as a function of V_0 magnitude is shown here. The average metallicity for RGB stars, i. e. stars with magnitude $V_0 < 14.7$ [Fe/H] $= -1.162 \pm 0.002$ dex, is significantly lower than RGB stars. Assuming lower values for the micro-turbulence does not reduce significantly this difference.

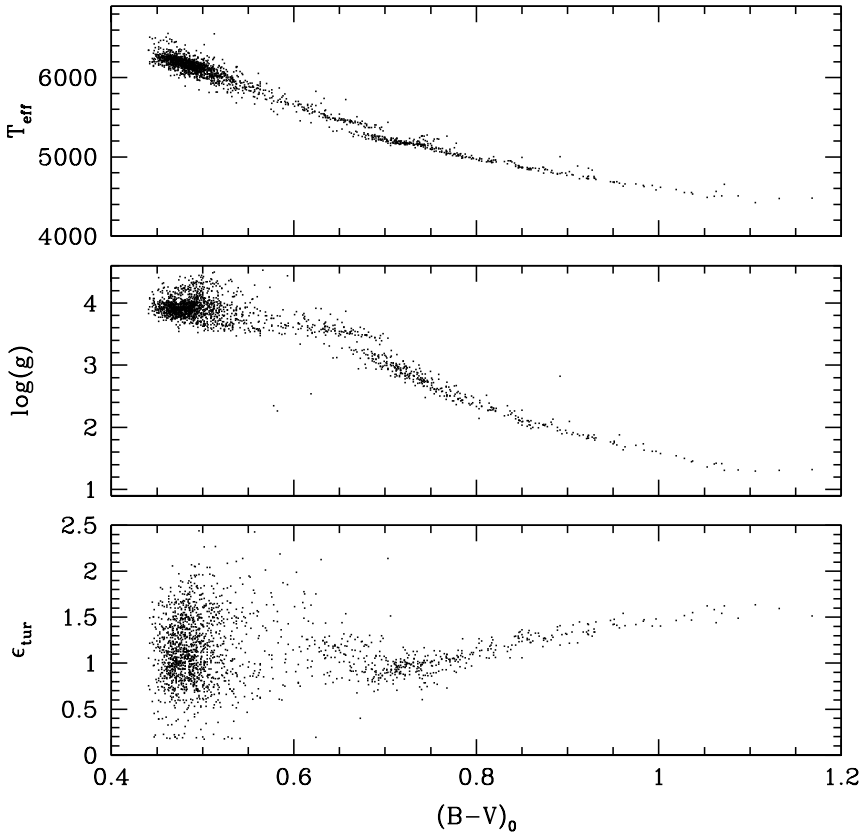

Figure 38: The trend of effective temperature (upper plot), gravity (middle plot) and micro-turbulent velocity (lower plot) as a function of color $(B-V)_0$. The jump at $(B-V)_0 \simeq 0.70$ is probably due to the unreliable determination of ξ_t , which in turn is affecting the other parameters during the χ^2 minimization

More probably, the intrinsic difficulty of deriving the micro-turbulent velocity due to the lack of strong lines for stars in the Main Sequence and Sub giant Branch coupled with the low SNR is causing a reduction of independent data points, i. e. points that are more sensitive to an atmosphere stellar parameters than the others. The fitting procedure is not able to break the degeneracy between micro-turbulent velocity and the other parameters, settling down to a set of values dependent on their intial values and the path followed by the χ^2 minimization algorithm. This effect can be seen also in the jump in T_{eff} in conjunction with the point where ξ_t stops following a defined trend at $(B-V)_0 \simeq 0.70$. Indeed that point corresponds to the points where the SGB join the RGB.

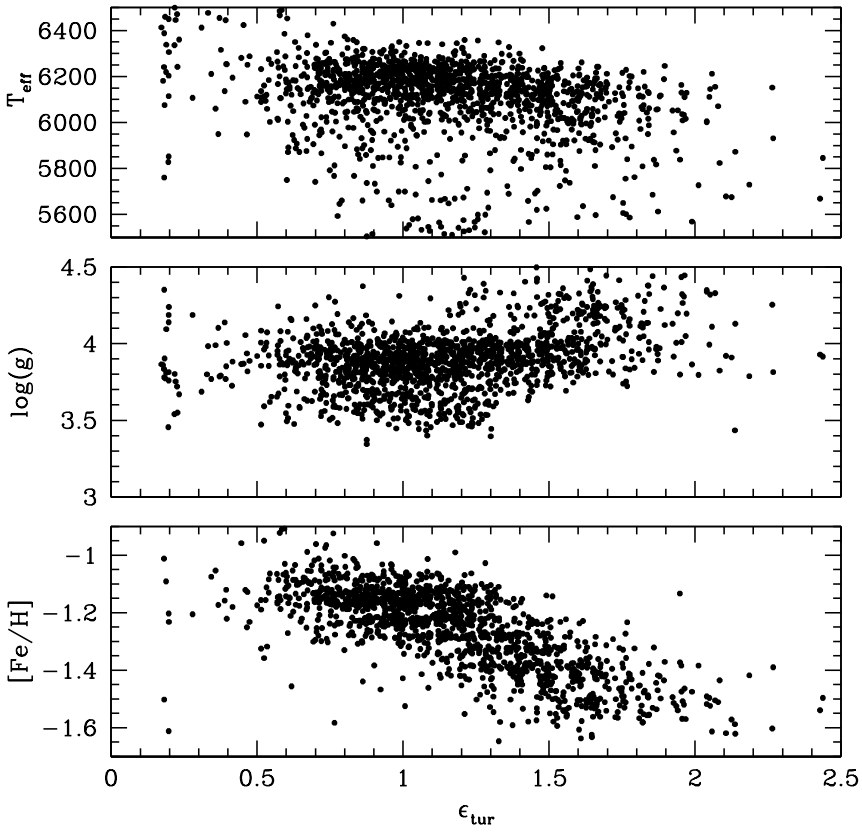

Figure 39: A strong trend in metallicity with micro-turbulent velocity is present fro Main Sequence and Red Giant Branch

The incorrect derivation of stellar parameters affects the determination of metallicity: Figure 39 shows the clear correlation of [Fe/H] with the micro-turbulence velocity. Unfortunately this effect is amplified by the lack of a good calibration of ξ_t as a function of colors and magnitude, so that the derivation of this value is made even more difficult.

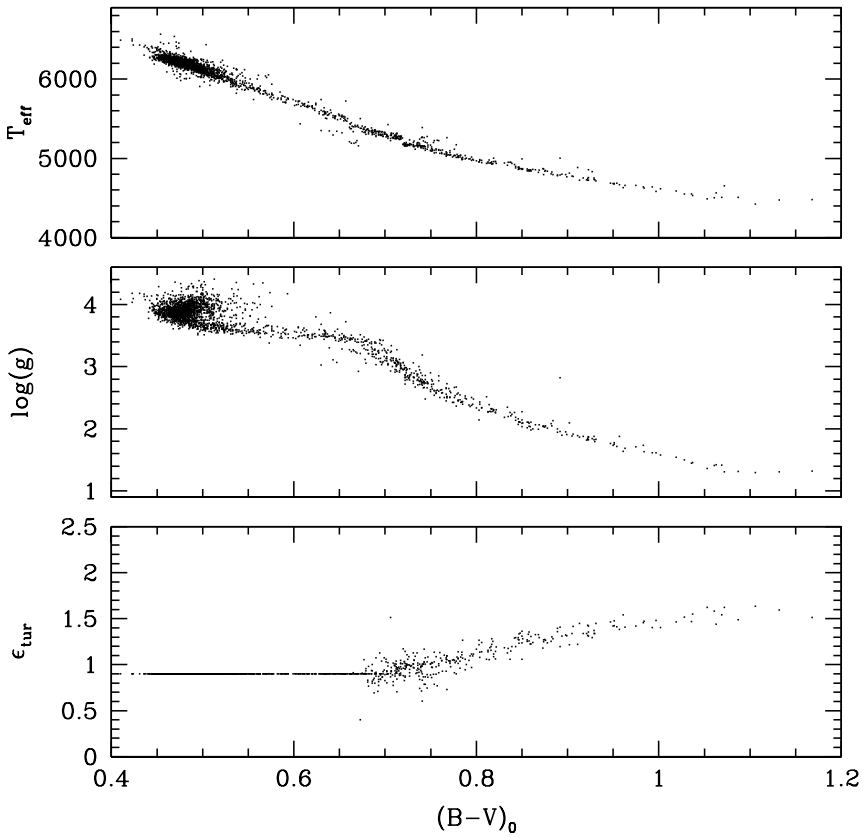

Figure 40: Atmospheric parameters after combining giants with ξ_t as a free parameter and dwarfs with fixed ξ_t . The transition between the two groups in temperature and gravity is now smoother.

We decided to fix at $\xi_t = 0.9$km/s the micro-turbulence for stars with $V_0 > 14.7$. This value has been derived according to the trend of ξ_t defined by the RGB extrapolated at the given magnitude. With this contraint, the number of star successfully analyzed grows up to 2235. The resulting atmosphere parameters are shown in Figure 40. The transition between the two groups is now smooth.

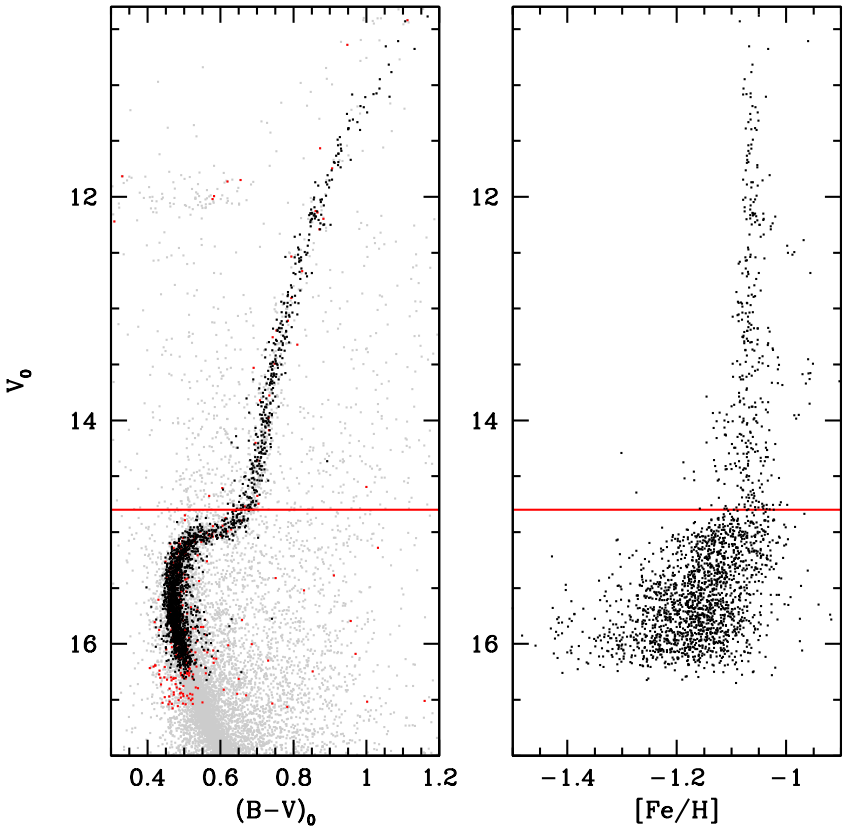

Figure 41: Metallicity as a function of V_0 magnitude is shown here. For Main Sequence and Sub Giants Branch stars a micro-turbulent velocity of $\xi_t = 0.9$km/s has been assumed (stars under the red line). The average metallicity for MS and SGB stars, $[Fe/H] = -1.162 \pm 0.002$ dex, is significantly lower than RGB stars. Assuming lower values for ξ_t does not reduce significantly this difference.

Finally the trend of metallicity with magnitude is shown in Figure 41. The mean metallicity measured from MS and SGB stars is significantly lower than the one measured from RGB stars. From 1869 stars with $V_0 > 14.7$, we have an average metallicity of $[Fe/H] = -1.162 \pm 0.002$ dex, with a dispersion of $\sigma = 0.09$ dex. Assuming lower values of ξ_t results in a small difference of a few hundredths of dex in [Fe/H], and it cannot explain such a difference. More investigations on this point are required.

4.7 CONCLUSIONS

We presented a new algorithm to automatically determine the continuum normalization level for spectra characterized by a very short and almost continuum-less wavelength range, very low SNR and strong contamination from sky background. We coupled this algorithm with a modified χ^2 minimization algorithm for the derivation of stellar parameters by comparison with spectral syntheses. Data points have been weighted according to the sensitivity of spectral features with respect to variations in the atmosphere parameters.

The robustness of our algorithm has been tested over 322 RGB stars down to $V \simeq 16$ resulting in a constant metallicity along the entire RGB and an average value of $[Fe/H] = -1.068 \pm 0.001$ dex, consistent with the one derived in Marino et al. [2008] from 105 high-resolution UVES spectra of bright stars ($V < 14$) and using Equivalent Width analysis. We stress the fact that not only the same results are recovered despite using observations with lower spectral resolution and a wavelength range that is almost ten times shorter, but we are able to extend the analysis to stars that are two magnitudes fainter than the ones in the high-resolution sample.

Atmosphere parameter determination has proved to be more difficult for the faintness of the stars. We fixed the micro-turbulent velocity at $\xi_t = 0.9$km/s to remove one degree of freedom in the χ^2 minimization. From 1869 stars in the Sub Giant Branch and Main Sequence we have an average metallicity of $[Fe/H] = -1.162 \pm 0.002$ dex, with a dispersion of $\sigma = 0.09$ dex. Assuming slightly different values of micro-turbulent velocity does not remove the observed difference in metallicity.A decrease in surface metallicity around the Turn-Off has been already observed in others globular clusters and attributed to diffusive processes (sedimentation, levitation) coupled with turbulent mixing below the outer convection, see ad example Lind et al. 2008. Due to the tricky characteristics of the spectra in this temperature range and the low SNR involved, we have decided to perform more tests before confirming this result.

5

ATMOSPHERIC STELLAR PARAMETERS FROM CROSS-CORRELATION FUNCTION

In the last chapter of my thesis we have decided to face the atmospheric parameter determination problem from an original perspective. From a qualitative point of view it is well know that the shape of Cross-Correlation Function (CCF) depends somehow by the characteristics of the spectrum under analysis. A recent work in this direction has been made by Buchhave et al. [2012], with the stellar parameters determined by cross-correlating the observed spectrum with a library of synthetic spectra and selecting the combination of parameters that results in the most peaked CCF. Torres et al. [2012] have shown however that strong correlations exist between temperatures, gravities and metallicities derived with this technique. Our approach differs substantially from the previous one. The main core of our technique is the determination of several CCFs using different line lists, where neutral iron lines are selected according to their sensitivity to changes in temperature to break the temperature-metallicity degeneracy, and a separate CCFs is determined using ionized Iron lines for gravity determination.

After the introduction of the *numerical* CCFs and an analytical description of the link between equivalent width and CCF, we introduce our method to calibrate the CCF parameters with respect to atmosphere parameters obtained with Equivalent Width analysis of high-resolution and high SNR spectra. To different calibrations are presented. The first is purely empirical and it provides temperature and metallicity from the observed CCF areas for stars with gravity in the range $\log g 4.2 - 4.6$ (solar-type stars) . The second calibration is based on the transformation of synthetic equivalent width in CCF areas and it allows the determination of gravity with temperature and metallicity, but derived parameters are slightly less precise than the first calibration. We have tested the performance of our calibration using the individual exposures available for each star.

5.1 CROSS-CORRELATION FUNCTION USING A NUMERICAL MASK

The CORAVEL-type cross-correlation function (CCF) has been introduced by Baranne et al. [1996] and it consists in the numerical transposition of the optical cross-correlation made by instrument like CORAVEL [Baranne et al., 1979]. The core of this technique is a wavelength mask where only the center of the lines that are wanted to be used have non-null values; the CCF is computed step by step for each radial

velocity (RV) point by multiplying the observed spectrum with the wavelength mask, conveniently shifted according to the radial velocity of the given point, and the result is integrated. No rebinning or merging of orders is necessary.

The extracted spectrum is not usually ready for the CCF determination and some precautions must be taken in order to ensure that different exposures of a given stars results in CCFs with the same characteristic parameters, after correcting for the observed radial velocity of the star. This is particularly true in our case, since we want to compare CCF parameters from different stars, and we must be sure that observed difference in the CCF parameters are due to different star characteristic and not to the different observing conditions.

The extracted raw spectrum is mainly affected by several instrumental and atmospheric effects: the blaze function of the Echelle grating, the change efficiency with wavelength of the combined instrument and telescope, the dispersion caused by the atmosphere, and many others of less importance. Two steps are required to reduce the impact of these effects. The first one is to divide the science spectrum with a normalized spectrum of a continuum lamp: blaze function, CCD defects and in general all the instrumental effects that affect a science spectrum are corrected with this step, since star and lamp light are subject to the instrumental effects.
The second one is more subtle because one has to deal with the change of slope in the continuum of the spectrum. When an object is observed at different elevation, the differential atmospheric refraction can alter the amount of light entering in the spectrograph. Because of differential refraction, the position of the star in the focal plane of the telescope at short wavelengths (blue part of the spectra) differs from the one at long wavelengths (red part of the spectra). Guiding systems like the one installed on HARPS to keep the fibre at the center of the stars usually work with white light (no filter in wavelength), with the consequence that less light enters in the spectrograph at those wavelengths that are farther from the one used to determine its position, i. e. the one at which the flux of the stars peaks. An *Atmospheric Dispersion Corrector* (ADC) can greatly reduce this effect, but some differences between exposures are likely to remain because these systems work on the basis of the average behavior of the weather observed at the site, and not by applying a correction on-the-fly.

5.2 PREPARING THE SPECTRA FOR THE CCF

In classical spectral analysis the spectrum is normalized either to a local continuum when determining the equivalent widths, or using a more general continuum function when comparing with spectral synthesis; differences in the slope of the continuum between obser-

vations even of the same stars are not relevant and no correction is needed for them.

In the determination of the numerical CCF, the differential refraction introduces two subtle effects. First, a given line will have different weights on different measurements, altering the shape of the CCF and eventually its center if the rest wavelength of the line is not precisely known. This condition is very likely since the rest wavelength seems to be a function of the spectral type [Allende Prieto et al., 2002]. Second, changes in the spectra introduced by the increasing loss of light towards the blue wavelengths will result in a slope in the continuum of the CCF, hence the position of the center will be affected proportionally to the amount of differential refraction.

In typical spectral reduction pipelines (such as that implemented for HARPS and HARPS-N data) these effects are usually corrected using a few standard references to calibrate the flux. The resulting spectrum does not take into account the stellar parameters of the observed stars. Using the same reference for all the observations will still ensure that internal systematics between different exposures of the same stars are removed, but the corrected flux distribution will be likely different from the outside-atmosphere one.

The ideal procedure would be to correct for blaze function $b_p(\lambda)$ and differential refraction $a(\lambda)$, but in doing this we would lose the informations about the SNR of each line, i. e. we would give the same weight to all the lines including the very noisy ones (e. g. at the edge of the lines, or in part of the spectrum strongly affected by differential refraction). The correction should be performed splitting each function as a product of a constant determined at the central wavelength of the line λ_0 and a normalized function:

$$\begin{aligned} c(\lambda, n_0) &= c_n(\lambda, n_0) * c(\lambda_0, n_0) \\ b(\lambda, n_0) &= b_n(\lambda, n_0) * b(\lambda_0, n_0) \\ a(\lambda) &= a_n(\lambda) * a(\lambda_0) \end{aligned} \tag{21}$$

In this way instrumental effects and slopes would be corrected without affecting the SNR of the spectrum. The blaze correction and the continuum function depends also on the considered order, since the same wavelength can appear on two contiguous orders; other effects that alter the continuum slope (e. g. differential refraction, sensitivity of the CCD) are dependent only on the wavelength We denote with a c the spectrum corrected using the function on the left side of Equation 21, while we use a n when the spectrum is corrected using the normalized function.

The corrected observed spectra are given by equations:

$$f_c(\lambda) = f_p(\lambda, n_o)/b(\lambda, n_o) * a(\lambda) \tag{22}$$

$$f_n(\lambda, n_o) = f_p(\lambda, n_o)/b_n(\lambda, n_o) * a_n(\lambda) \tag{23}$$

In Equation 22 f_c loses its dependence from the considered order, since the lost of efficiency caused by the blaze function is corrected. In Equation 23 this dependence is still present because we are correcting for the *shape* introduced by the instrument but not for the loss of light caused by it.

After the differential refraction correction, the corrected continuum function $c_c(\lambda)$ of the science spectrum is related to the template flux-calibrated continuum $c_t(\lambda)$ through a scaling factor m, that is simply the ratio between the template flux and the corrected science flux at any arbitrary λ_m:

$$\begin{aligned} c_c(\lambda) &= c_t(\lambda) * \frac{f_c(\lambda_m)}{f_t(\lambda_m)} \\ c_n(\lambda) &= c_t(\lambda) * \frac{f_n(\lambda_m) * a(\lambda_m)/b(\lambda_m)}{f_t(\lambda_m)} * b(\lambda, n_0)/a(\lambda) \\ m &= \frac{f_c(\lambda_m)}{f_t(\lambda_m)} = \frac{f_n(\lambda_m) * a(\lambda_m)/b(\lambda_m)}{f_t(\lambda_m)} \end{aligned} \tag{24}$$

To correctly derive this factor, the general correction (Equation 22) must be applied to the observed spectrum. When using Equation 23 the factors $a(\lambda)$ and $b(\lambda)$ appear because the continuum function must be rescaled for the actual value of $f_c(\lambda)$. In practice, it is useful to use the same normalization point used for a, so that $a(\lambda_m) = 1.0$.

Note finally that in the wavelength region immediately adjecent to a spectral line the continuum should have very small variation. This allows the use of the following approximation:

$$c_t(\lambda) = c_t(\lambda_0) \tag{25}$$

The advantage of substitutions in Equation 24 and 25 is that there is no need of determining the continuum of the observed spectrum anymore; since the continuum normalization is the most difficult task in the analysis of low SNR spectrum, the advantage is considerable. The drawback is that now the obtained values will depends on the goodness of the differential refraction correction in restoring the correct flux in function of wavelength, and this in turn depends also on how close the spectral template is to the observed spectrum. In the next sections we show in detail how to derive the stellar parameters from CCFs. Once that a first guess in temperature and metallicity has been derived, the correction for differential refraction can be performed again using a more appropriate template, if the atmosphere parameters of previous template was differing too much from the ones just derived.

5.3 THE LINK BETWEEN THE CCF AND THE EW

Given λ_i and λ_e as the extreme of the line integration, the usual formula for the EW is:

$$EW = \int_{\lambda=\lambda_i}^{\lambda_e} \frac{c(\lambda) - f(\lambda)}{c(\lambda)} \tag{26}$$

Spectral lines are usually approximated with a gaussian shape, when in the linear part of the curve of growth [Gray, 2005]. Knowing the *line contrast* (i. e. the depth of the line in the normalized spectrum) a and its full width half maximum FWHM, we can calculate its area using the well-known expression:

$$AREA = \int a e^{-\frac{(x-x_0)^2}{2\sigma^2}} = \sqrt{2\pi} a \sigma = \frac{\sqrt{2\pi}}{2.3548} a FWHM \tag{27}$$

Equation 27 applies to the line both in the wavelength and in the radial velocity space:

$$AREA_\lambda = EW = \frac{\sqrt{2\pi}}{2.3548} FWHM_\lambda \frac{c_c(\lambda_0) - f_c(\lambda_0)}{c_c(\lambda_0)} \tag{28}$$

$$AREA_{RV} = \frac{\sqrt{2\pi}}{2.3548} FWHM_{RV} \left(CCF^{cont}(0) - CCF^{obs}(0) \right) \tag{29}$$

We need to find a link between the EW and the area subtended by the CCF. The CCF is constructed using the technique introduced by Baranne et al. [1996] and improved by Pepe et al. [2002], also known as *CORAVEL-type* or *numerical CCF*. A *CCF mask* is defined as a collection of wavelengths λ_i, where each entry i has a wavelength bin $\delta\lambda_i$ (also called *holes*) and a weight w_i associated. The weight is determined using a synthetic spectrum and it takes into account the depth of the line and its width, while the hole is related to the spectral resolution of the instrument. The CCF mask is used to build a function with wavelength as parameter and that has value w_i in the interval $\lambda_i \pm \delta_i/2$ for each i, and zero otherwise. The value of the CCF at a given δRV is calculated by applying the Doppler radial-velocity shift to the wavelength of the function defined above, and integrating the product of the doppler-shifted function with the observed spectrum.

In the case of a mask with a single hole of size $\delta\lambda$,at a given δRV the value of the CCF is the integration over the size of the mask hole at the line wavelength, accordingly shifted for the corresponding δRV:

$$CCF(\delta RV) = \int_{\lambda_{\delta RV}-\delta\lambda/2}^{\lambda_{\delta RV}+\delta\lambda/2} f(\lambda) d\lambda \tag{30}$$

Radial velocities of stars are always well below the relativistic limits (i.e. $RV_{star} < 1000 km/s$), so the classical formula is a good approximation for the wavelength shift.

$$\lambda_{\delta RV} = \lambda_0 \left(\frac{\delta RV}{c} + 1 \right) \tag{31}$$

In determining the CCF, the size of the *hole* will change the amount of flux in each RV bin, changing its contrast, while the sampling is not actually influencing the shape of the CCF so it is irrelevant in our discussion (it influences the goodness of the gaussian fit, of course). Then the FWHM in the wavelength space easily converts in the radial velocity space:

$$FWHM_{RV} = \frac{c}{\lambda_0} FWHM_{\lambda} \tag{32}$$

The depth of the CCF is given by Equation 30 evaluated at $\delta RV = 0$, while the expected value of the CCF at point of minimum is given by the same equation but substituting $f(\lambda)$ with $c(\lambda)$:

$$CCF(\delta RV) = \int_{\lambda_{\delta RV} - \delta\lambda/2}^{\lambda_{\delta RV} + \delta\lambda/2} (c_c(\lambda) - f_c(\lambda))\, d\lambda \tag{33}$$

In a generic case solving Equation 33 to obtain the integrand it would be difficult. In practice, the size of the *hole* is always chosen to be very small compared to the FWHM of the instrumental profile, so we can approximate $f_c(\lambda)$ and $c_c(\lambda)$ as constant between the integration limits. Equation 33 becomes:

$$a_{RV} = (c_c(\lambda_0) - f_c(\lambda_0))\, \delta\lambda \tag{34}$$

We transpose Equations 32 and 34 into Equation 28 to derive the relationship between the EW and the CCF for a mask composed of a single line:

$$\begin{aligned} EW &= \frac{\sqrt{(2\pi)}}{2.3548} \frac{\lambda_0}{c} \frac{a_{RV} FWHM_{RV}}{c_c(\lambda_0)\ \delta\lambda} \\ &= \frac{\lambda_0}{c} \frac{AREA_{RV}}{c_c(\lambda_0)\ \delta\lambda} \\ &= \frac{\lambda_0}{c} \frac{AREA_{RV}}{m c_t(\lambda_0)\ \delta\lambda} \end{aligned} \tag{35}$$

If the spectra used to determine $AREA_{RV}$ have been corrected with Equation 23, Equation 35 take the form:

$$\begin{aligned} EW &= \frac{\lambda_0}{c} \frac{AREA_{RV}}{c_n(\lambda_0)\ \delta\lambda} \\ &= \frac{\lambda_0}{c} \frac{AREA_{RV}}{m c_t(\lambda_0) b(\lambda_0, n_0)/a(\lambda_0, n_0)\ \delta\lambda} \end{aligned} \tag{36}$$

If we could use the EW directly, there would be no advantage in switching in the radial velocity space. The advantage of the CCF is that a CCF from a mask of several lines is simply the coaddition of the *individual* CCF from each line [Pepe et al., 2002]. A synthetic CCF Area can be determined using Equation 37.

$$\begin{aligned} CCF &= \sum_i AREA_{RV} \\ &= \sum_i \frac{c}{\lambda_i} \frac{mc_t(\lambda_i) b(\lambda_i, n_i)\ \delta\lambda}{a(\lambda_i, n_i)} EW_i \end{aligned} \tag{37}$$

Until now we implicitly assumed that the template spectrum, the observed one and all the derived functions are continuous in wavelength. In practice, what we really have is the integration of the given functions over the size (in wavelength) of a pixel. The above formulation is still valid if we substitute each function in the previous section with its integrated counterpart divided by the size of the pixel of integration, for example for a continuous function $f(\lambda)$ we have:

$$f(\lambda) = \frac{f^int(\lambda_j)}{\Delta\lambda_j} \tag{38}$$

where $f^int(\lambda_j)$ is the integral of the function $f(\lambda)$ over the interval $\Delta\lambda_j$, corresponding to the size of the pixel j, and centered in λ_j.

5.4 CORRECTION OF FLUX DISTRIBUTION

Here we describe a general algorithm to bring the observed spectra to a standard reference flux distribution, in order to remove differences between exposures due to changes in observing conditions. The correction is performed using the synthetic stellar library from Coelho et al. [2005]: this library spans a wide rane of stellar atmosphere parameters, has high resolution spectra ($R > 200000$) and it covers a wide range of wavelengths (from 300nm to 1.0μm). The algorithm includes the following steps:

- Before proceeding, the spectra must be corrected for the blaze function, to remove order-dependent instrumental effects.
- For each extracted order the average flux for unit wavelength is determined. Integration is performed on the central half of the extracted spectra (from 1024 to 3072 pixel); the mean wavelength between the two integration limits is taken as reference value for the wavelength.
- The same operation is performed on a template spectrum with similar initial stellar parameters, using as integration limits the wavelengths corresponding to the pixels used in the previous step;

- Normalization is performed dividing the obtained values by the integrated flux from the 68th order (centered on 5500Å) for the observed spectra and the respective value for the template.
- The template-derived values are divided by the respective observed-derived ones to determine the correction factor for each wavelength step.
- The correction factors are interpolated as a function of wavelength with a B-Spline of the 4th order.

We noticed that some points have a correction factor systematically lower in some wavelength domains, like in the 3900-4000 Å region: this is due to the presence of strong lines (in the mentioned region, the CaII H-K doublet) where the differences between the observed spectra and the used template become important; the affected orders are excluded from the fit in the last step.

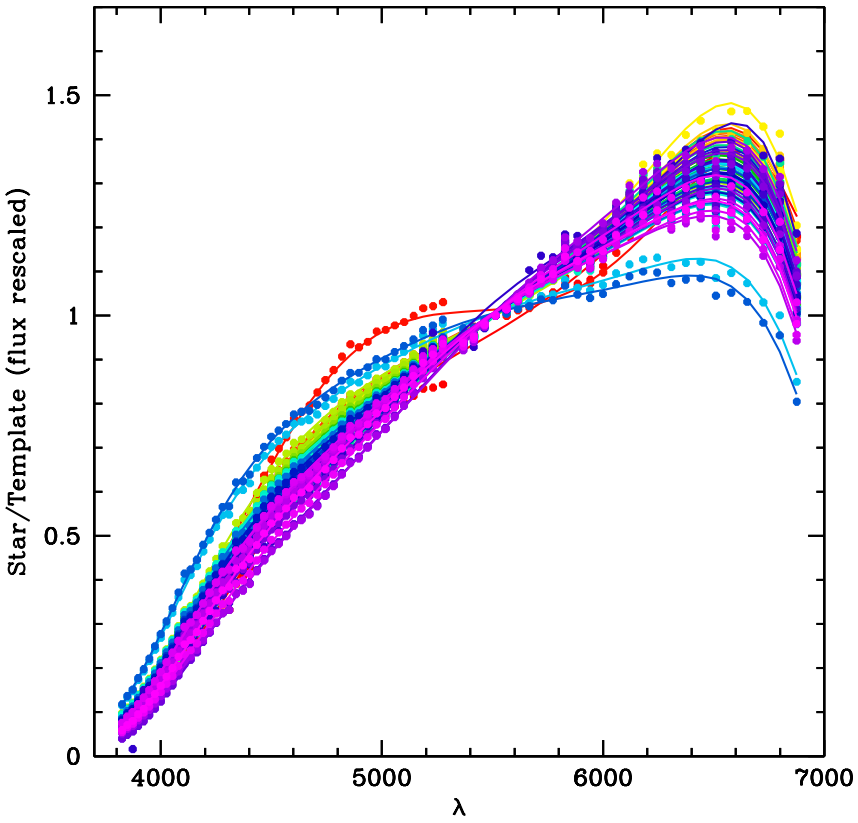

Figure 42: Flux correction for 75 exposures of HD125612.

In Figure 42 the algorithm has been applied to 75 exposures of the star HD125612. The resulting function is a combination of a general trend caused by the wavelength-dependent light loss due to telescope (e. g. mirror coating) and instrument (e. g. CCD sensitivity curve), and

small variations resulting from the difference between actual atmospheric refraction effect and the one predicted by the ADC for a given airmass.

5.5 CHARACTERISTIC OF THE SAMPLE OF STARS

A sample of stars widely spanning the range in the stellar parameters that we want to measure is essential. At our disposal we have 1111 stars from several HARPS long-term programs [Mayor et al., 2003, Lo Curto et al., 2010, Santos et al., 2011]. Since these stars are the targets of extensive surveys for exoplanet search, a large number of exposures have been collected during the ten years of operations of HARPS: this allows us to use the same stars as *calibrators* when considering the coadded, high SNR spectrum, and as *test case* when a single low SNR spectrum is under study.

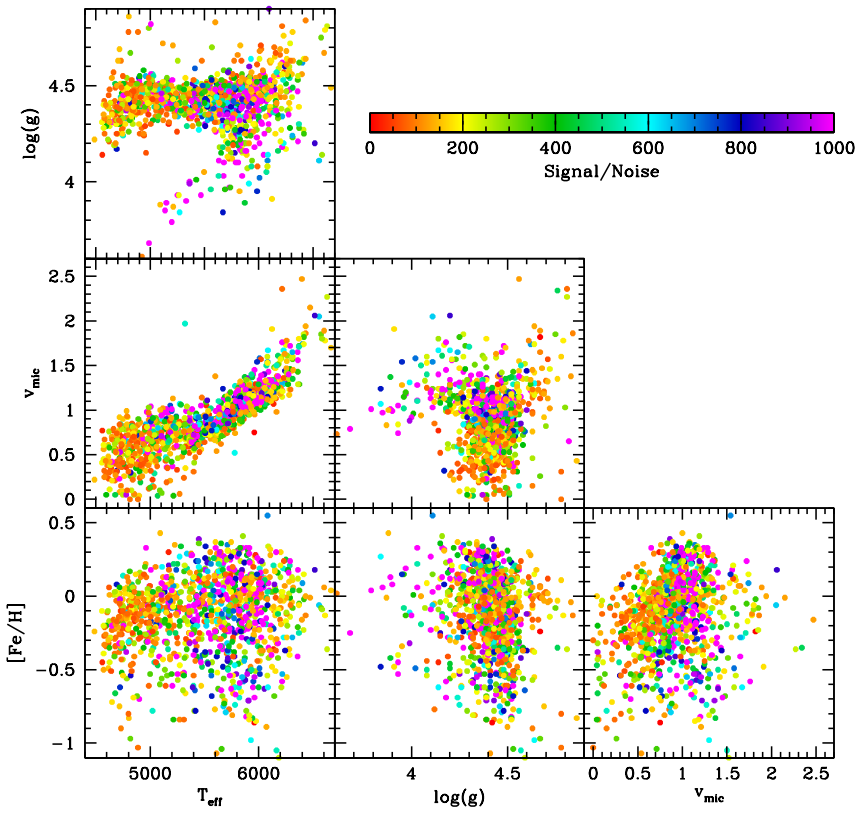

Figure 43: The stellar atmosphere parameters of the *calibration* sample, derived using high SNR spectrum. Except for ξ_t , no trend between atmosphere parameters is present.

Stellar atmosphere parameters have been derived by Adibekyan et al. [2012] from co-added spectrum using classical EW analysis; the automatic continuum determination and EW measurements per-

formed with ARES [Sousa et al., 2007], ensure a high degree of homogeneity on the derived parameters of the sample. The distribution of the atmosphere parameters are displayed in Figure 43: the sample contains only dwarf stars with $\overline{\log g} \simeq 4.4$ and a few exceptions, a clear trend of micro-turbulence ξ_t with effective temperature is clearly visible, while other parameters are almost independent and span a wide range in T_{eff} and [Fe/H]. The quoted error are $\sigma_{T_{eff}} = 30$ K, $\sigma_{\log g} = 0.06$ dex and $\sigma_{[Fe/H]} = 0.03$ dex.

5.6 LINE SELECTION FOR THE CCF MASKS

The HARPS Pipeline automatically determines the CCF and its parameters at the end of each exposure using one of three available masks (F5, G2, M2). Each mask contains several thousands of lines chosen among several chemical elements and ionization states and weighted according to the amount of Radial Velocity information they carry, e. g. RV shifts of deep and sharp lines are easier to measure [Pepe et al., 2002].

Since every line is carrying information on the radial velocity shift of the star regardless its chemical origin, and the principal goal of the instrument is to measure the differential shift of the star to a precision of less than 1m/s, all the available unsaturated lines are included.

An extended number of lines increase the SNR of the CCF so that a more precise radial velocity is measurable. But we have a more ambitious goal: to determine relationships between the CCF parameters and the atmosphere parameters of the stars. For this work, including lines with unknown atomic parameters, or from chemical elements that can have large star-to-star abundance variations at a given metallicity, can add significant noise to the derived relationships. We must make a more stringent selection of the lines included in the CCF mask in order to minimize these potential problems:

- including lines from several chemical elements will decrease the precision of the calibration, because the chemical composition is not a function of temperature, i. e. we are adding more parameters without calibrating them
- including lines with unknown atomic parameters will make difficult to disentangle the effects of Temperature and Metallicity in the CCF parameters behavior.

To select a good sample of lines for our mask, we start with the list from Sousa et al. [2010], which is in turn an extension of Iron lines list from Sousa et al. [2007] and several chemical elements from Neves et al. [2009] for a total number of 498 lines. The steps followed to derive the line list are briefly explained here. The initial line list, along with the atomic parameters (including an initial estimate of the oscillator strengths) is taken from the VALD4 online database (Piskunov

et al. 1995, Kupka et al. 2000) using the solar stellar parameters as input ($T_{eff} = 5777K$, $log(g) = 4.44$,$v_{mic} = 1.0km/s$) in the spectral region from 4500Å to 6910Å. The following criteria are followed:

- Lines must not be strongly blended in the Kurucz Solar Flux Atlas
- Lines must not be too weak ($EW < 5mÅ$) or too strong ($EW > 200mÅ$)
- Lines must not reside in the wings of strong lines (e. g. Hα, Hβ and MgI lines)
- EWs measured in the Kurucz Solar Flux Atlas and in the solar reflected light spectrum of the Ceres asteroid (obtained with HARPS) must agree within 10%.
- Lines must be reproducible by the ARES program Sousa et al. [2007]

Finally empirical oscillator strengths are obtained through an inverse analysis with the MOOG driver *ewfind* and the measured EWs from the solar spectrum, using the parameters described in Santos et al. [2004].

The provided wavelengths are used as input positions by the ARES program to perform a Gaussian fit of the lines in order to derive their EWs, but they lack the required precision to be directly used in a cross-correlation mask; for example, a difference of 0.01Å in the position of the line corresponds to a misplacement of the CCF of $0,545km/s$ at 5500Å. We have several reason to require better precision:

- our tool will be used on very low SNR spectra, so a-priori knowledge of the position of the lines is required.
- Using more precise line positions will reduce the FWHM and increase the contrast of the CCF, resulting in a better gaussian fit and a general improvement of the derived parameters.
- with precise line wavelengths, bisector spans and line shifts in function if the spectral type of the star can be determined more precisely.

To re-determine accurately the line positions, a multi-gaussian fit is performed around the wavelength in the line-list. The more precise values are already included in the tables below.

In Figure 44 the distribution of the selected FeI lines versus the basic atomic line parameters is displayed. The selected lines are homogeneously distributed in wavelength and cover a good range in Excitation Potential and Oscillator Strength. The same distributions for FeII lines are plotted in Figure 45.

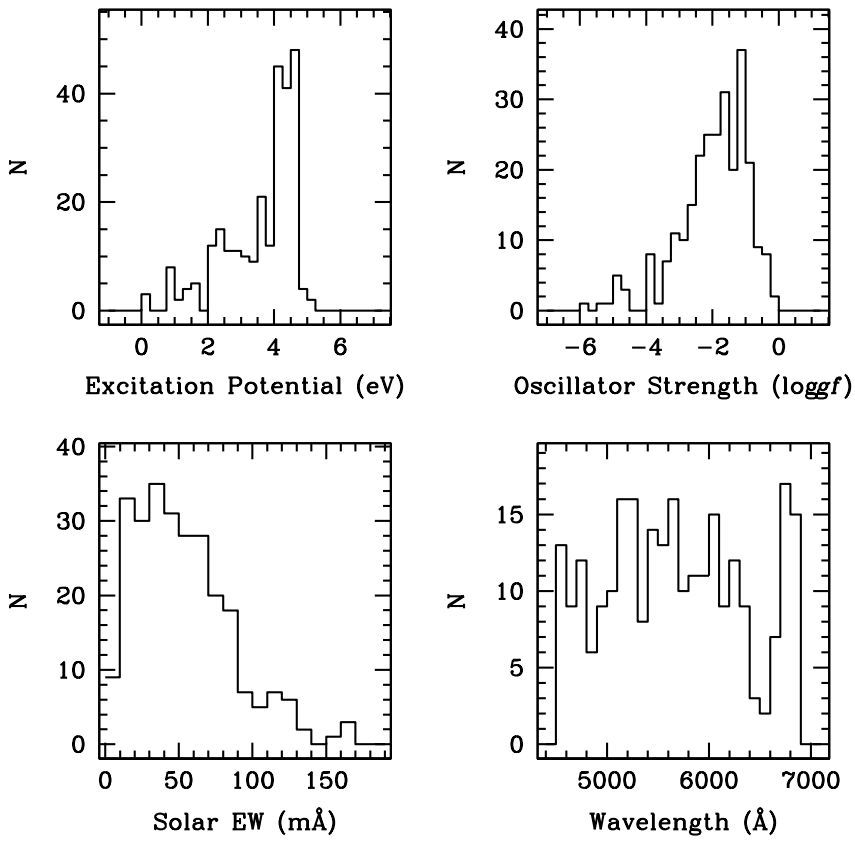

Figure 44: Distribution of selected FeI line according to their value of Excitation Potential, Oscillator Strength, Equivalent Width measured in the Solar Spectrum and Wavelength.

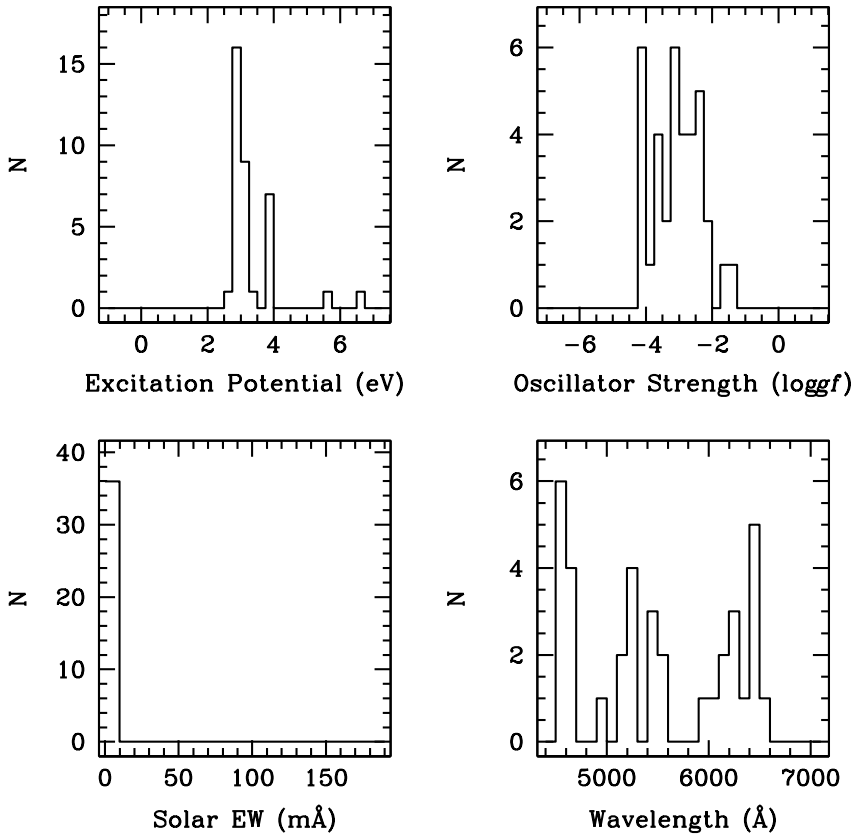

Figure 45: Same as Figure 44 but for FeII

In Figure 46 the measured continuum-normalized area under the CCFs determined by the HARPS pipeline (left panel) are compared with the CCF areas obtained when only FeI lines are used (right panel), for the same set of stars and as a function of Effective Temperature and Metallicity. The jump in color in the left panel is due to the masks used for different spectral types. Color transition is slightly smoother on the right panel, since this CCF is not influenced by star-to-star chemical variations.

Every line is both dependent on the effective temperature of the photosphere and the abundance of the chemical element that is producing that line, and this behavior is conserved when the lines are co-added into a CCF. It then appears clear why a single CCF area is not enough to constrain both Metallicity and Temperature if no external information is available, and this is true regardless of the mask that is used.

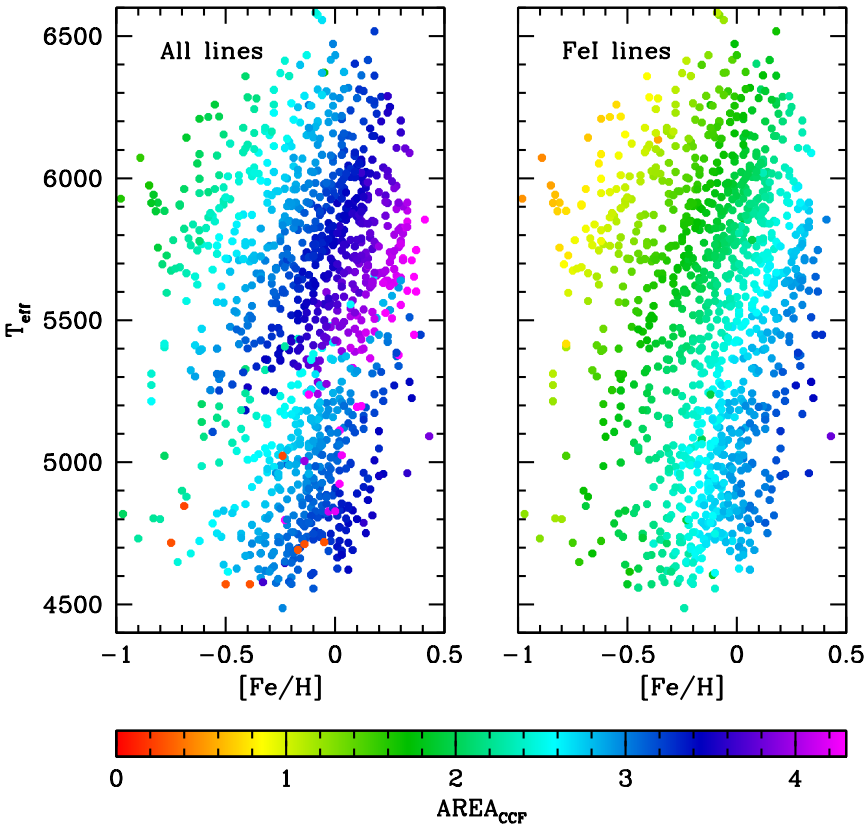

Figure 46: In the left panel, the CCF Area in function as determined by the HARPS Pipeline in the T_{eff}- [Fe/H] plane; the jump in the area values is due to the different mask used in the pipeline. In the left panel, the area of the CCF when using only FeI lines.

To break this degeneracy, we need to construct masks with lines having different behavior with temperature, in such a way that the

lines connecting the same values of CCF areas will have different slopes on the T_{eff}-[Fe/H] plane, i. e. they will cross in a point corresponding to the star parameters.

In Equivalent Width analysis a key parameter when determining the Effective Temperature of the star is the Excitation Potential (hereafter EP) The excitation potential quantifies the energy required to excite a line to a given state from the ground one. After the determination of the EW of each line, a stellar atmosphere model with given temperature, gravity and overall metallicity is assumed, and then a value of the expected abundance is derived from each EW measurement. If the assumed temperature is too far from the real value, a trend of derived abundances with EP will appear. For example, if the assumed temperature is higher than the real value, the theoretical EW will decrease and the abundance of the chemical element will be increased in order to match the observed EW;

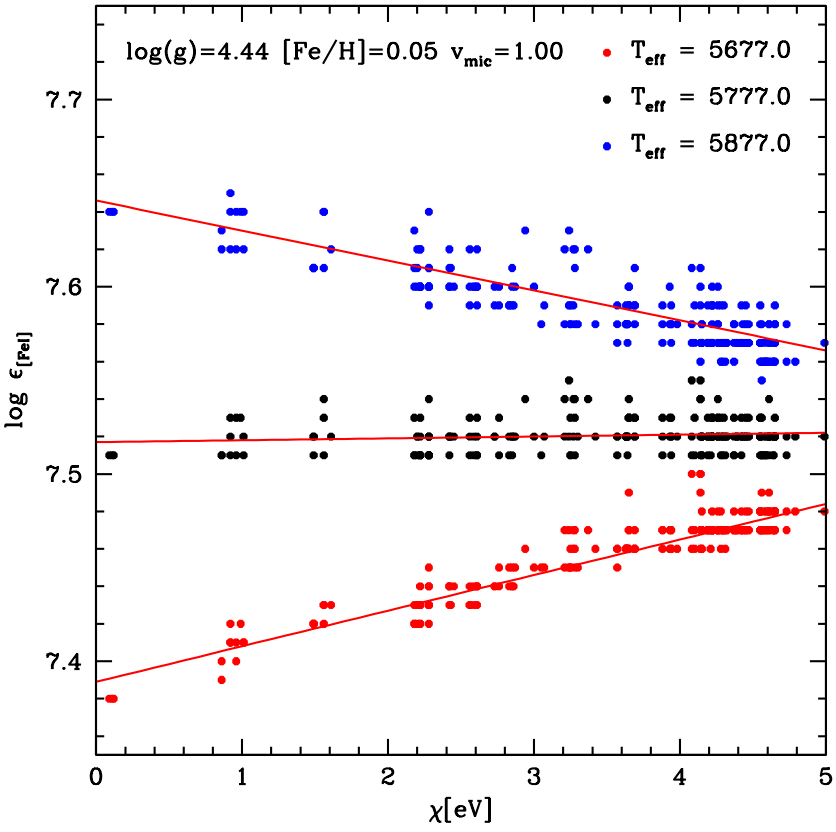

Figure 47: Temperature affects lines accordingly to their Excitation Potential, as shown in this plot. Iron abundances for single lines are determined using Solar EWs and neutral iron line atomic parameters from Sousa et al. [2007], and Kurucz atmosphere models with fixed log g, ξ_t and [Fe/H] and three different values for T_{eff}.

This situation is exemplified in Figure 47 using Solar EWs and neutral iron line atomic parameters from Sousa et al. [2007] and three Kurucz model atmosphere with $\log g = 4.44$, $\xi_t = 1.0 \mathrm{km\ s^{-1}}$ and [Fe/H] = 0.05 and three different values for the Effective Temperature. The used value for metallicity is a consequence of the different Solar Abundance used to derive the line atomic parameters ($\log \epsilon(\mathrm{Fe}) = 7.47$) and the value coded in MOOG ($\log \epsilon(\mathrm{Fe}) = 7.52$).

We then decided to split our linelist in three part to create three masks:

- one with low excitation potential lines, comprising 59 lines with EW < 3.0eV, Table 5;
- one with 47 intermediate EP lines ($3.0 <$ EP < 4.0, 6);
- the last one with 136 high excitation potential lines (EP > 4.0, Table 7)

Table 5: Atomic parameter of FeI lines selected for the *low*-EP mask (EP $<$ 3.0)

n	Name	Atm n	Wavelength [Å]	EP	log *gf*	EW_{solar}	Line contrast
208	FeI	26.00	4554.46183	2.870	-2.752	42.600	0.350
210	FeI	26.00	4561.41331	2.760	-2.879	40.000	0.368
213	FeI	26.00	4574.72204	2.280	-2.823	63.800	0.621
215	FeI	26.00	4579.32787	2.830	-3.024	30.600	0.309
219	FeI	26.00	4602.00417	1.610	-3.163	76.300	0.703
225	FeI	26.00	4630.12495	2.280	-2.488	79.400	0.688
227	FeI	26.00	4635.84972	2.850	-2.384	58.100	0.581
232	FeI	26.00	4741.53332	2.830	-2.027	73.000	0.686
235	FeI	26.00	4772.82163	1.560	-2.752	96.200	0.743
242	FeI	26.00	4794.35880	2.420	-3.892	14.500	0.129
252	FeI	26.00	4924.77429	2.280	-2.058	103.900	0.667
260	FeI	26.00	5044.21272	2.850	-2.038	82.700	0.641
273	FeI	26.00	5127.36070	0.920	-3.317	106.100	0.745
276	FeI	26.00	5141.73918	2.420	-2.125	94.400	0.689
277	FeI	26.00	5143.72111	2.200	-3.811	25.100	0.246
287	FeI	26.00	5198.71057	2.220	-2.164	103.900	0.668
290	FeI	26.00	5225.52764	0.110	-4.773	76.000	0.684
296	FeI	26.00	5247.05128	0.090	-4.941	70.200	0.657
297	FeI	26.00	5250.20998	0.120	-4.870	71.600	0.647
298	FeI	26.00	5253.02006	2.280	-3.855	20.800	0.200
325	FeI	26.00	5436.58739	2.280	-3.267	47.500	0.410
343	FeI	26.00	5567.38250	2.610	-2.594	64.500	0.462
364	FeI	26.00	5701.54631	2.560	-2.162	89.400	0.639
381	FeI	26.00	5853.14823	1.490	-5.130	8.700	0.069

387	FeI	26.00	5916.24934	2.450	-2.920	57.800	0.468
392	FeI	26.00	5956.69433	0.860	-4.526	57.300	0.466
397	FeI	26.00	6005.54069	2.590	-3.479	25.200	0.204
407	FeI	26.00	6082.70994	2.220	-3.566	38.000	0.310
414	FeI	26.00	6120.25010	0.920	-5.894	6.000	0.049
417	FeI	26.00	6151.61714	2.180	-3.298	52.800	0.428
421	FeI	26.00	6173.33582	2.220	-2.877	71.400	0.535
422	FeI	26.00	6180.20434	2.730	-2.632	59.300	0.440
424	FeI	26.00	6200.31502	2.610	-2.397	77.000	0.545
425	FeI	26.00	6213.43081	2.220	-2.580	86.400	0.602
426	FeI	26.00	6219.28102	2.200	-2.463	93.200	0.622
429	FeI	26.00	6229.22871	2.850	-2.866	42.000	0.305
433	FeI	26.00	6240.64581	2.220	-3.292	51.800	0.413
435	FeI	26.00	6265.13490	2.180	-2.559	89.500	0.607
436	FeI	26.00	6270.22502	2.860	-2.573	56.100	0.425
438	FeI	26.00	6297.79344	2.220	-2.777	76.700	0.540
441	FeI	26.00	6322.68852	2.590	-2.368	80.300	0.546
442	FeI	26.00	6335.33238	2.200	-2.339	101.300	0.644
444	FeI	26.00	6358.67818	0.860	-3.907	89.000	0.537
448	FeI	26.00	6392.53843	2.280	-3.942	20.200	0.156
455	FeI	26.00	6481.87130	2.280	-2.929	67.600	0.488
456	FeI	26.00	6498.93819	0.960	-4.627	49.500	0.380
458	FeI	26.00	6581.20918	1.490	-4.706	21.300	0.144
459	FeI	26.00	6593.87467	2.430	-2.384	89.100	0.569
460	FeI	26.00	6608.02500	2.280	-3.967	19.700	0.149
461	FeI	26.00	6609.11166	2.560	-2.632	69.600	0.476
462	FeI	26.00	6625.01990	1.010	-5.284	18.000	0.129
464	FeI	26.00	6646.93157	2.610	-3.915	11.800	0.078
467	FeI	26.00	6703.56751	2.760	-3.022	40.000	0.296
469	FeI	26.00	6710.31819	1.490	-4.810	17.900	0.131
477	FeI	26.00	6739.51992	1.560	-4.902	13.200	0.100
480	FeI	26.00	6750.15315	2.420	-2.615	78.400	0.527
484	FeI	26.00	6806.84455	2.730	-3.103	37.900	0.274
490	FeI	26.00	6839.83044	2.560	-3.377	33.000	0.238
496	FeI	26.00	6861.93642	2.420	-3.795	21.300	0.153

Table 6: Atomic parameter of FeI lines selected for the *intermediate*-EP mask ($3.0 < \mathrm{EP} < 4.0$)

n	Name	Atm n	Wavelength [Å]	EP	log *gf*	EW_{solar}	Line contrast
202	FeI	26.00	4523.39995	3.650	-1.871	49.200	0.480
203	FeI	26.00	4531.62442	3.210	-1.801	68.600	0.607
205	FeI	26.00	4537.67212	3.270	-2.870	20.100	0.207
207	FeI	26.00	4551.65019	3.940	-1.928	33.100	0.302

209	FeI	26.00	4556.92795	3.250	-2.644	29.600	0.300
211	FeI	26.00	4566.51932	3.300	-2.156	49.000	0.485
212	FeI	26.00	4574.21981	3.210	-2.353	46.400	0.433
217	FeI	26.00	4593.52751	3.940	-1.921	33.600	0.310
218	FeI	26.00	4596.40916	3.650	-2.090	38.600	0.358
221	FeI	26.00	4607.64640	3.980	-0.754	89.000	0.000
231	FeI	26.00	4690.13992	3.690	-1.550	63.800	0.555
234	FeI	26.00	4757.58093	3.270	-1.988	58.900	0.547
236	FeI	26.00	4779.44173	3.420	-2.162	44.200	0.441
237	FeI	26.00	4780.81529	3.250	-3.236	10.800	0.097
238	FeI	26.00	4787.82901	3.000	-2.532	48.500	0.386
239	FeI	26.00	4788.76101	3.240	-1.791	67.100	0.617
240	FeI	26.00	4789.65505	3.550	-1.171	85.900	0.691
241	FeI	26.00	4793.97035	3.050	-3.464	10.100	0.094
243	FeI	26.00	4799.40975	3.640	-2.073	39.100	0.387
245	FeI	26.00	4802.88406	3.690	-1.527	62.300	0.600
246	FeI	26.00	4808.15066	3.250	-2.630	30.800	0.305
247	FeI	26.00	4809.94075	3.570	-2.542	22.100	0.210
248	FeI	26.00	4811.04762	3.070	-3.182	16.700	0.151
249	FeI	26.00	4885.43556	3.880	-1.136	73.300	0.643
250	FeI	26.00	4905.13730	3.930	-1.818	40.200	0.279
255	FeI	26.00	4961.92180	3.630	-2.301	29.700	0.273
262	FeI	26.00	5054.64101	3.640	-1.992	43.400	0.374
274	FeI	26.00	5129.62469	3.940	-1.469	57.500	0.377
289	FeI	26.00	5223.18298	3.630	-2.252	32.500	0.300
294	FeI	26.00	5242.49264	3.630	-1.124	85.900	0.691
299	FeI	26.00	5253.46189	3.280	-1.591	89.300	0.600
304	FeI	26.00	5288.52737	3.690	-1.612	59.300	0.511
306	FeI	26.00	5294.54620	3.640	-2.627	17.800	0.151
311	FeI	26.00	5379.57412	3.690	-1.552	62.700	0.542
329	FeI	26.00	5466.98344	3.650	-2.141	37.300	0.312
337	FeI	26.00	5538.51395	3.630	-2.090	40.800	0.327
345	FeI	26.00	5584.76799	3.570	-2.189	38.600	0.307
352	FeI	26.00	5636.69636	3.640	-2.511	22.400	0.195
375	FeI	26.00	5809.21420	3.880	-1.614	57.100	0.388
379	FeI	26.00	5827.87613	3.280	-3.147	13.600	0.103
391	FeI	26.00	5934.65536	3.930	-1.091	87.500	0.542
411	FeI	26.00	6096.66388	3.980	-1.776	43.400	0.306
423	FeI	26.00	6187.98857	3.940	-1.620	54.100	0.368
427	FeI	26.00	6220.77974	3.880	-2.330	21.900	0.160
428	FeI	26.00	6226.73322	3.880	-2.069	33.200	0.236
430	FeI	26.00	6232.64103	3.650	-1.240	83.400	0.550
487	FeI	26.00	6828.59124	3.570	-1.871	61.800	0.385

Table 7: Atomic parameter of FeI lines selected for the *high*-EP mask (EP > 4.0)

n	Name	Atm n	Wavelength [Å]	EP	log *gf*	EW_{solar}	Line contrast
226	FeI	26.00	4631.48650	4.550	-1.890	13.700	0.146
229	FeI	26.00	4661.53666	4.560	-1.186	41.600	0.406
233	FeI	26.00	4749.94943	4.560	-1.236	38.200	0.350
244	FeI	26.00	4802.52112	4.610	-1.714	17.200	0.153
254	FeI	26.00	4952.64668	4.210	-1.184	57.400	0.473
256	FeI	26.00	4962.57581	4.180	-1.209	57.100	0.493
257	FeI	26.00	4967.90191	4.190	-0.720	82.400	0.647
258	FeI	26.00	4993.70294	4.210	-1.138	59.400	0.420
263	FeI	26.00	5067.15575	4.220	-0.860	78.200	0.043
265	FeI	26.00	5072.66965	4.220	-0.983	67.300	0.542
268	FeI	26.00	5088.14985	4.150	-1.554	43.100	0.344
269	FeI	26.00	5090.77305	4.260	-0.514	100.700	0.581
272	FeI	26.00	5109.64817	4.300	-0.736	82.600	0.605
279	FeI	26.00	5159.05835	4.280	-0.846	76.400	0.592
280	FeI	26.00	5180.06099	4.470	-1.105	52.600	0.403
281	FeI	26.00	5187.90836	4.140	-1.202	59.700	0.469
283	FeI	26.00	5195.47182	4.220	-0.441	99.300	0.696
284	FeI	26.00	5196.05774	4.260	-0.827	73.800	0.600
286	FeI	26.00	5197.93464	4.300	-1.491	39.700	0.345
291	FeI	26.00	5228.37558	4.220	-1.095	64.200	0.482
295	FeI	26.00	5243.77602	4.260	-1.022	66.700	0.540
305	FeI	26.00	5293.95889	4.140	-1.747	33.700	0.291
307	FeI	26.00	5295.31254	4.420	-1.518	33.200	0.284
308	FeI	26.00	5361.61257	4.420	-1.205	49.300	0.324
309	FeI	26.00	5373.70739	4.470	-0.841	68.200	0.511
310	FeI	26.00	5376.82941	4.290	-2.040	16.900	0.159
312	FeI	26.00	5386.33104	4.150	-1.709	36.500	0.310
313	FeI	26.00	5389.47982	4.420	-0.534	90.800	0.630
314	FeI	26.00	5395.21392	4.450	-1.724	23.100	0.201
315	FeI	26.00	5398.27883	4.450	-0.684	78.800	0.581
316	FeI	26.00	5401.26648	4.320	-1.712	28.500	0.241
317	FeI	26.00	5406.77323	4.370	-1.429	39.900	0.333
318	FeI	26.00	5409.13444	4.370	-1.051	60.200	0.459
320	FeI	26.00	5417.03633	4.420	-1.404	39.000	0.311
323	FeI	26.00	5432.94694	4.450	-0.729	75.800	0.553
324	FeI	26.00	5436.29431	4.390	-1.319	44.700	0.359
326	FeI	26.00	5441.33730	4.310	-1.558	36.200	0.286
327	FeI	26.00	5461.54806	4.450	-1.570	29.600	0.236
328	FeI	26.00	5464.27838	4.140	-1.595	41.100	0.356
330	FeI	26.00	5473.16097	4.190	-1.986	22.600	0.182

331	FeI	26.00	5481.24737	4.100	-1.203	65.800	0.484
332	FeI	26.00	5491.82954	4.190	-2.195	15.300	0.120
333	FeI	26.00	5522.44459	4.210	-1.419	47.900	0.392
335	FeI	26.00	5534.65905	4.150	-2.391	11.400	0.099
338	FeI	26.00	5543.93447	4.220	-1.070	66.500	0.517
339	FeI	26.00	5546.50457	4.370	-1.124	56.300	0.446
340	FeI	26.00	5546.99323	4.220	-1.741	31.700	0.224
341	FeI	26.00	5553.57512	4.430	-1.321	42.700	0.374
342	FeI	26.00	5560.21027	4.430	-1.064	56.700	0.451
346	FeI	26.00	5587.57247	4.140	-1.656	38.400	0.329
347	FeI	26.00	5594.65811	4.550	-0.858	59.600	0.492
348	FeI	26.00	5618.63153	4.210	-1.298	54.600	0.427
349	FeI	26.00	5619.59830	4.390	-1.435	38.700	0.298
350	FeI	26.00	5633.94543	4.990	-0.385	63.200	0.468
351	FeI	26.00	5635.82130	4.260	-1.556	39.500	0.317
353	FeI	26.00	5638.26130	4.220	-0.809	82.200	0.582
354	FeI	26.00	5641.43805	4.260	-0.969	70.500	0.502
356	FeI	26.00	5651.46766	4.470	-1.763	21.200	0.173
357	FeI	26.00	5652.31480	4.260	-1.751	29.700	0.240
358	FeI	26.00	5653.86288	4.390	-1.402	40.500	0.320
359	FeI	26.00	5661.34298	4.280	-1.828	25.600	0.202
360	FeI	26.00	5662.51611	4.180	-0.601	91.800	0.642
361	FeI	26.00	5667.51115	4.180	-1.292	54.600	0.395
362	FeI	26.00	5679.02192	4.650	-0.756	65.000	0.474
363	FeI	26.00	5680.23861	4.190	-2.330	12.000	0.100
365	FeI	26.00	5715.08280	4.280	-0.847	72.600	0.530
366	FeI	26.00	5720.89142	4.550	-1.805	17.300	0.126
367	FeI	26.00	5731.76090	4.260	-1.124	61.900	0.475
368	FeI	26.00	5738.23414	4.220	-2.164	15.900	0.113
369	FeI	26.00	5741.84618	4.260	-1.626	35.800	0.285
370	FeI	26.00	5752.03159	4.550	-0.917	56.600	0.451
371	FeI	26.00	5775.07949	4.220	-1.124	61.400	0.479
372	FeI	26.00	5778.45385	4.150	-1.967	24.700	0.210
373	FeI	26.00	5793.91327	4.220	-1.622	37.700	0.299
374	FeI	26.00	5806.72259	4.610	-0.877	59.200	0.418
376	FeI	26.00	5811.91392	4.140	-2.333	13.200	0.103
377	FeI	26.00	5814.80571	4.280	-1.820	26.200	0.204
378	FeI	26.00	5815.21533	4.150	-2.364	12.300	0.092
380	FeI	26.00	5852.21752	4.550	-1.190	44.600	0.336
382	FeI	26.00	5855.07636	4.610	-1.531	25.500	0.196
383	FeI	26.00	5856.08672	4.290	-1.572	36.100	0.296
384	FeI	26.00	5862.35804	4.550	-0.404	85.700	0.584
385	FeI	26.00	5902.47214	4.590	-1.797	16.100	0.127

386	FeI	26.00	5905.66985	4.650	-0.775	63.400	0.456
388	FeI	26.00	5927.78738	4.650	-1.057	47.200	0.353
389	FeI	26.00	5929.67531	4.550	-1.211	43.600	0.336
390	FeI	26.00	5930.18164	4.650	-0.326	94.100	0.584
393	FeI	26.00	5983.67918	4.550	-0.719	67.400	0.507
394	FeI	26.00	5984.81588	4.730	-0.335	81.300	0.478
395	FeI	26.00	5987.06295	4.790	-0.478	69.800	0.449
399	FeI	26.00	6027.05110	4.080	-1.178	65.400	0.507
400	FeI	26.00	6034.03221	4.310	-2.306	10.300	0.075
401	FeI	26.00	6035.33612	4.290	-2.476	7.600	0.056
402	FeI	26.00	6054.07144	4.370	-2.237	10.600	0.084
403	FeI	26.00	6056.00248	4.730	-0.489	77.800	0.506
405	FeI	26.00	6078.49047	4.790	-0.364	76.900	0.513
406	FeI	26.00	6079.00793	4.650	-1.008	50.000	0.367
409	FeI	26.00	6089.56801	4.580	-1.273	37.900	0.300
410	FeI	26.00	6094.37200	4.650	-1.566	22.800	0.161
412	FeI	26.00	6098.24550	4.560	-1.755	18.900	0.135
415	FeI	26.00	6127.90516	4.140	-1.417	51.000	0.388
418	FeI	26.00	6157.72729	4.080	-1.238	62.700	0.478
419	FeI	26.00	6159.37254	4.610	-1.878	13.900	0.105
420	FeI	26.00	6165.35871	4.140	-1.502	47.000	0.371
437	FeI	26.00	6290.97162	4.730	-0.576	67.700	0.420
439	FeI	26.00	6303.45605	4.320	-2.492	7.100	0.043
440	FeI	26.00	6315.81024	4.080	-1.645	43.200	0.322
446	FeI	26.00	6380.74188	4.190	-1.322	54.100	0.399
447	FeI	26.00	6385.71689	4.730	-1.828	12.500	0.089
463	FeI	26.00	6627.54280	4.550	-1.475	31.500	0.212
465	FeI	26.00	6653.85179	4.150	-2.407	12.200	0.081
466	FeI	26.00	6699.14019	4.590	-2.106	9.400	0.063
468	FeI	26.00	6705.10137	4.610	-1.057	48.300	0.330
470	FeI	26.00	6713.04325	4.610	-1.514	26.800	0.194
471	FeI	26.00	6713.19149	4.140	-2.438	11.500	0.080
472	FeI	26.00	6713.74007	4.790	-1.425	24.000	0.162
473	FeI	26.00	6725.35584	4.100	-2.187	20.300	0.135
474	FeI	26.00	6726.66633	4.610	-1.045	49.200	0.338
475	FeI	26.00	6732.06232	4.580	-2.144	8.900	0.061
476	FeI	26.00	6733.14961	4.640	-1.429	29.800	0.204
478	FeI	26.00	6745.09985	4.580	-2.094	9.900	0.069
479	FeI	26.00	6745.95787	4.080	-2.657	8.400	0.056
481	FeI	26.00	6752.70544	4.640	-1.234	39.400	0.261
482	FeI	26.00	6786.85804	4.190	-1.886	29.200	0.192
483	FeI	26.00	6793.25527	4.580	-1.901	14.500	0.102
485	FeI	26.00	6810.26021	4.610	-0.995	54.900	0.351

486	FeI	26.00	6820.36730	4.640	-1.134	45.300	0.291
488	FeI	26.00	6833.22509	4.640	-1.975	11.400	0.072
489	FeI	26.00	6837.00260	4.590	-1.732	19.300	0.136
491	FeI	26.00	6842.68496	4.640	-1.169	43.600	0.265
492	FeI	26.00	6843.65358	4.550	-0.867	65.000	0.410
493	FeI	26.00	6855.71702	4.610	-1.674	21.100	0.134
494	FeI	26.00	6857.24593	4.080	-2.075	24.800	0.173
495	FeI	26.00	6858.14771	4.610	-0.972	55.500	0.361
497	FeI	26.00	6862.49138	4.560	-1.437	33.400	0.225
498	FeI	26.00	6864.31012	4.560	-2.229	7.900	0.051

The values obtained with these three mask will be identified respectively by the symbols A_{low},A_{med},A_{high} .

In Figure 48 we can see the result of using these three masks. If in each of the three plots we draw an imaginary line connecting the points with the resulting area for the corresponding masks from a given star, we can see that these lines have different slopes and cross in only one point in the T_{eff}-[Fe/H] plane, i. e. only a single (T_{eff},[Fe/H]) pair will match the observed CCF Area, thus breaking this parameter degeneracy.

One may wonder why we are using three masks when we only have two unknowns. We could have included the lines from the *intermediate*-EP mask in the *low*-EP mask, but then the closeness of the masks would make harder to break the degeneracy. Discarding the lines of the intermediate mask would not be advisable either, since they still contain useful information. Additionally, a 3-masks calibration can ensure some robustness for very low SNR observations (which are the main target of this technique) when for some reason one of the CCF Area measurement has been affected by some problem.

5.7 MODELING THE TEMPERATURE-METALLICITY PLANE

In the previous section we have derived three CCFs mask composed of neutral Iron lines. For solar-type stars the strengths of these lines do not depend much on gravity Gray [2005], so we can use these CCFs to derive a direct empirical calibration for T_{eff} and [Fe/H] suitable for dwarf stars ($\log g \simeq 4.4 \pm 0.3$dex). Gravity determination from CCFs will be studied in depth in the next section.

The easiest way to determine Temperature and Metallicity from the available CCF Areas is to simply calibrate the two parameters as functions of the available CCF Area from high SNR spectra:

$$T_{eff} = f_1(A_{low}, A_{med}, A_{high}) \tag{39}$$

$$[Fe/H] = f_2(A_{low}, A_{med}, A_{high}) \tag{40}$$

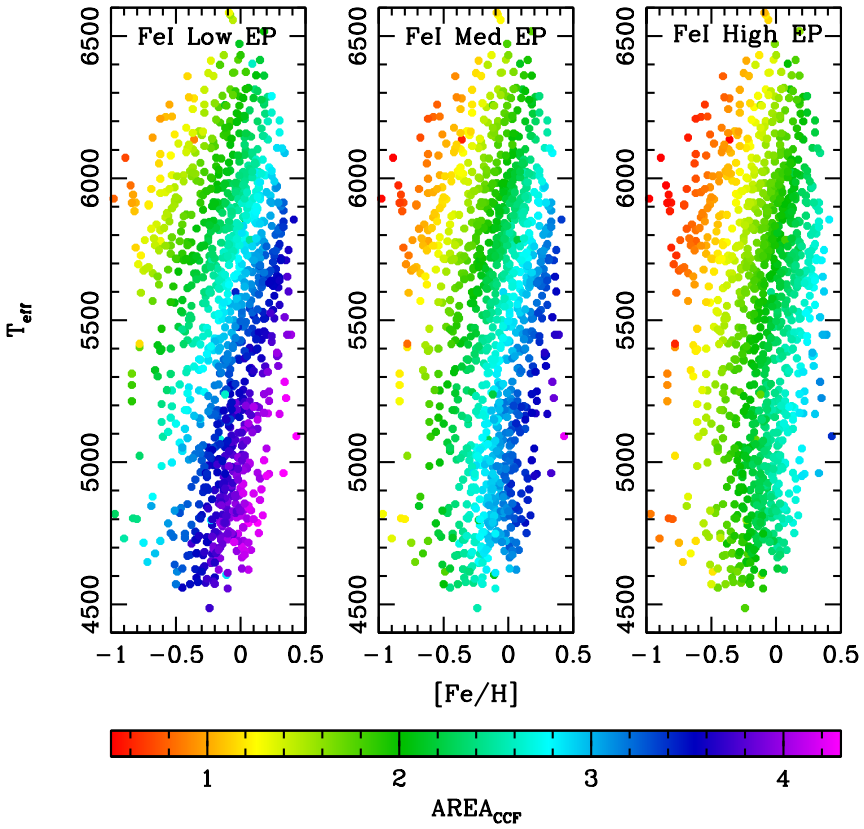

Figure 48: When using three masks with lines selected according to the above criteria, the iso-area lines have different slopes in the three plots, thus braking the T_{eff}-[Fe/H] degeneracy

The two functions f_1 and f_2 can be represented in any form, i. e. by any analytical function or a list of tabulated values to be interpolated. We found that a three-dimensional Chebyshev polynomial function of the first kind of the order 5x5x5 is good enough to fit both Equations 39 and 40. the Chebyshev polynomials are defined in the range $[-1, 1]$, so also the mid-points A^{mean} and range A^{range} for each variable must be provided. The variable trasformation is defined by

$$A' = (A^{measured} - A^{mean})/A^{range} \tag{41}$$

The coefficients of the functions are determined with the Levenberg-Marquardt algorithm (i. e. Least Squares minimization) using the values from EW analysis for T_{eff} and [Fe/H] and the measured values from high SNR spectra for the CCF Areas. The only part needed by an external program to determine the T_{eff} and [Fe/H] at this stage are:

- the three line-lists in Tables 5, 6 and 7
- the coefficients of the Functions 39 and 40
- the two parameters for each variables needed for Equation 41

The CCFs must be computed with the same technique as described in Section 5.1; since the Areas are normalized to the continuum level, the size of the *holes* in the mask is not relevant. Flux correction is playing a major role, since it's changing the weights of single lines when assembling the CCF, and it should be performed as described in Section 5.4. Note that the general trend shown in Figure 42 is due to the technical characteristics of the instrument and can be determined a priori, provided that the telescope is using an Atmospheric Dispersion Correction and that the derived values will be affected by a larger error for not correcting night-by-night deviations from the general shape.

The internal precision of the calibration is tested determining T_{eff} and [Fe/H] using the same Area values from high SNR spectra used to determine the function coefficients. Figure 49 and Figure 50 show the difference between the stellar parameters used to calibrate Equations 39 and 40 as a function of the measured CCF areas, and the value returned by these functions using the same areas. If the dispersion of the resulting values exceeds several times the precision of the calibrators, the order of the Chebyshev polynomials is too low, in the opposite situation the order is too high and we are over-fitting the data points.

We have $\sigma_{T_{eff}} = 42.6$ K and $\sigma_{[Fe/H]} = 0.027$ dex, against a reported precision of $\sigma_{T_{eff}} = 30$ K and $\sigma_{[Fe/H]} = 0.03$ dex from *EW* determination.

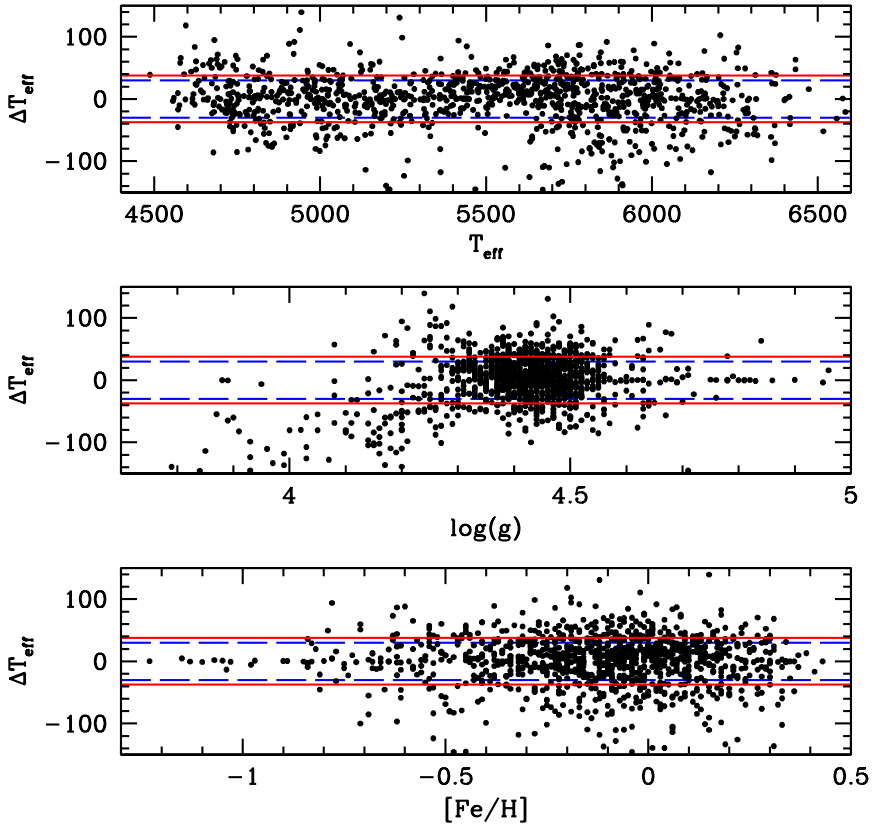

Figure 49: The internal precision of Temperature determination using our technique is shown by taking the difference between the values used to determine T_{eff}as a function of CCF areas, and the values returned by the calibrated function Equation 39 for the same CCF areas. Blue lines show the precision of EWmeasurements, i. e. the values used to calibrate the function, red lines show the $1-\sigma$ precision of our calibration.

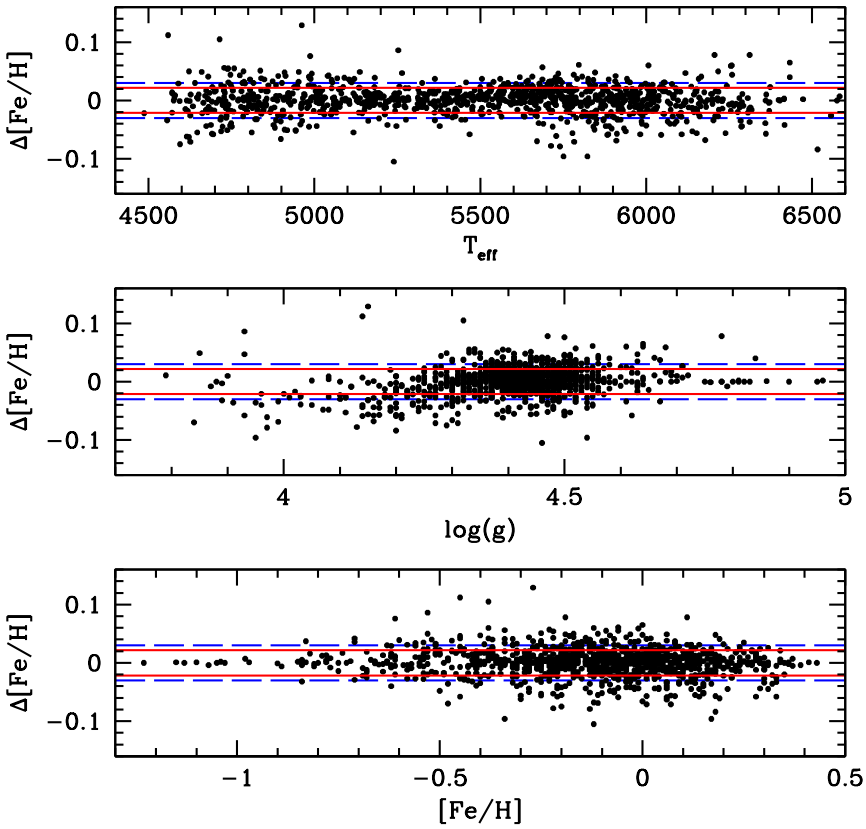

Figure 50: As in Figure 49, but for Metallicity

A small trend for lower values of gravity can be detected. In the next section we analyze the intrinsic difficulty in deriving the gravity. The precision at different SNR is analyzed in Section 5.9.

5.8 DERIVATION OF TEMPERATURE, SURFACE GRAVITY AND METALLICITY

To determine the gravity of a star we first need to identify a parameter that is sensitive to $\log g$ variations.

In principle, all the lines are sensitive to $\log g$: with increasing density, a larger number of absorbers per volume is available and the atomic lines are stronger. For cool stars however the strengths of neutral lines of most elements do not depend much on gravity (see again Figures 49 and Figure 50).

On the other hand, lines arising from the ionized species of most elements are often very dependent on gravity. The rate of ionization of atoms depends on temperature, but the recombination rate is a function of temperature and gravity (Saha equation): recombination reduces the number of ionized atoms and it is faster with increasing pressure (gravity), thus lines from most ionized states are very sensitive to this parameters.

In classical *EWs* analysis, Iron abundances are determined from individual FeII lines using T_{eff} and [Fe/H] derived from neutral Iron lines, and the value of $\log g$ is derived imposing the ionization equilibrium. Temperature and Metallicity are derived again with the updated value of Gravity, and the cycle is iterated.

Following the approach in Section 5.6 and Section 5.7, we define a CCF mask including isolated weak FeII lines taken from Sousa et al. [2010] and listed in 8.

Table 8: Atomic parameter of FeII lines

n	Name	Atm n	Wavelength [Å]	EP	log *gf*	EW_{solar}	Line contrast
200	FeI	26.10	4508.28524	2.860	-2.403	0.000	0.739
201	FeI	26.10	4520.22464	2.810	-2.563	0.000	0.670
204	FeI	26.10	4534.16496	2.860	-3.203	0.000	0.518
206	FeI	26.10	4541.51925	2.860	-2.762	0.000	0.561
214	FeI	26.10	4576.33502	2.840	-2.947	0.000	0.586
216	FeI	26.10	4582.83436	2.840	-3.076	0.000	0.537
222	FeI	26.10	4620.51653	2.830	-3.234	0.000	0.453
224	FeI	26.10	4629.34183	2.810	-2.262	0.000	0.740
228	FeI	26.10	4656.97832	2.890	-3.676	0.000	0.364
230	FeI	26.10	4670.17295	2.580	-4.011	0.000	0.276
251	FeI	26.10	4923.92946	2.890	-1.541	0.000	0.786
275	FeI	26.10	5132.66514	2.810	-4.008	0.000	0.232
285	FeI	26.10	5197.56981	3.230	-2.293	0.000	0.635

293	FeI	26.10	5234.62583	3.220	-2.235	0.000	0.644
300	FeI	26.10	5256.93532	2.890	-4.019	0.000	0.184
302	FeI	26.10	5264.80325	3.230	-3.091	0.000	0.425
303	FeI	26.10	5284.10577	2.890	-3.132	0.000	0.503
319	FeI	26.10	5414.06837	3.220	-3.568	0.000	0.252
321	FeI	26.10	5425.24758	3.200	-3.234	0.000	0.365
322	FeI	26.10	5427.80354	6.720	-1.481	0.000	0.043
334	FeI	26.10	5525.11639	3.270	-4.000	0.000	0.107
336	FeI	26.10	5534.83998	3.250	-2.792	0.000	0.493
396	FeI	26.10	5991.36987	3.150	-3.539	0.000	0.253
408	FeI	26.10	6084.10390	3.200	-3.774	0.000	0.173
413	FeI	26.10	6113.31904	3.220	-4.108	0.000	0.094
416	FeI	26.10	6149.24401	3.890	-2.719	0.000	0.286
431	FeI	26.10	6238.38089	3.890	-2.526	0.000	0.306
432	FeI	26.10	6239.93850	3.890	-3.448	0.000	0.098
434	FeI	26.10	6247.55980	3.890	-2.347	0.000	0.398
445	FeI	26.10	6369.45953	2.890	-4.131	0.000	0.153
449	FeI	26.10	6407.28259	3.890	-2.899	0.000	0.178
450	FeI	26.10	6416.92187	3.890	-2.625	0.000	0.302
452	FeI	26.10	6432.67939	2.890	-3.572	0.000	0.309
453	FeI	26.10	6442.95149	5.550	-2.399	0.000	0.031
454	FeI	26.10	6456.38211	3.900	-2.110	0.000	0.442
457	FeI	26.10	6516.08007	2.890	-3.279	0.000	0.384

Unfortunately the coverage in $\log g$ of our data sample is not sufficient to derive a direct calibration as done for T_{eff} and [Fe/H]in Section 5.7. Keeping in mind that the CCF Area from the FeII line mask is also a function of Temperature and Metallicity, we follow this strategy:

1. For each star, the expected CCF Area from EW ($A^{\star}_{EW}$) is determined using Equation 37 and the atmospheric stellar parameters from Adibekyan et al. [2012]

2. Then $A^{\log g}_{EW}$ is evaluated for a grid of $\log g$ values in the range $[3.6, 4.9]$

3. The correction factor $F_C(T_{eff}, [Fe/H]) = A^{\star}_{CCF}/A^{\star}_{EW}(T_{eff}, [Fe/H])$ is determined using stars with $\log g = 4.4 \pm 0.1$ and modeled with a 2D Chebyshev polynomial function of the first kind (order 3x3) as a function of Temperature and Metallicity (Figure 51)

4. $A^{\log g}_{EW}$ are rescaled to the measured values by using the function determined in the previous step

EWs are calculated using the LTE Spectral Synthesis program MOOGwith the ewfind driver. For the continuum determination, we decided to

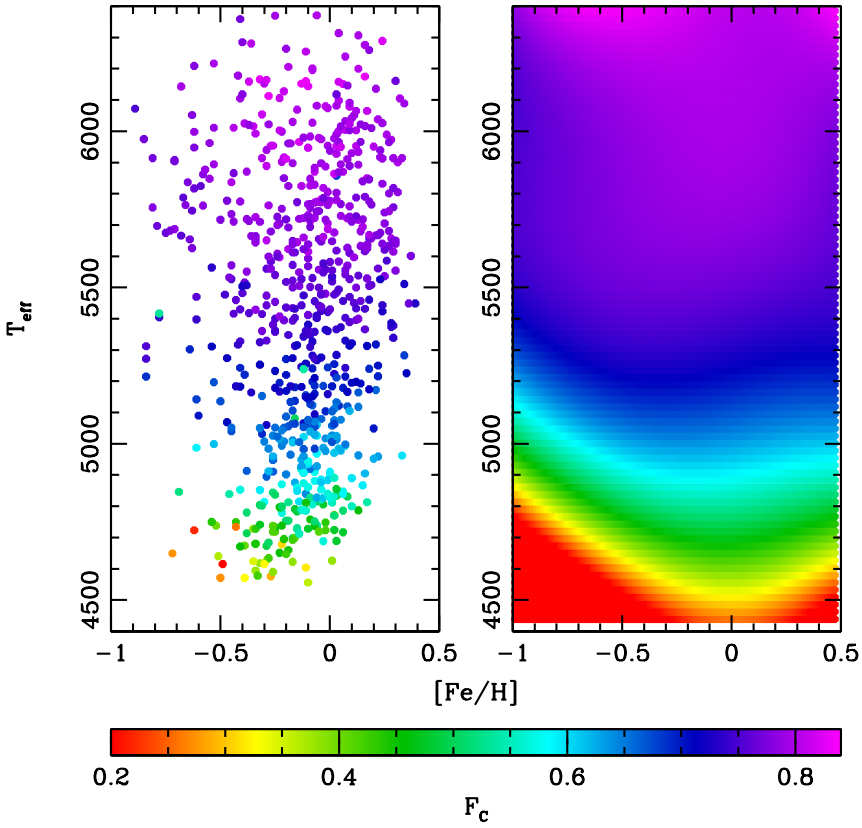

Figure 51: The correction factor $F_C(T_{eff}, [Fe/H])$ needed to rescale the synthesis-derived CCF Areas to the actual values measured in real stars. Left-hand panel, the measured values for stars with $4.3 < \log g < 4.5$; right-hand panel, the resulting 2D Chebyshev polynomial fit.

use the synthetic flux-calibrated stellar spectrum closest to the Sun from the Coelho et al. [2005] stellar library, which has $T_{eff} = 5750$ K, $\log g = 4.5$ dex, [Fe/H] $= 0.0$ dex. The systematic offset introduced by using a single synthesis instead of the whole library is taken automatically into account in the calibration of $F_C(T_{eff}, [Fe/H])$, and using the same synthesis ensures that there are no jumps in the calibration function. The function F_C also ensures that the synthesis-derived CCF Areas are properly rescaled to the observed values for the same combination of stellar atmosphere parameters.

The difference between the two values is mainly due to spectral lines close to the ones included in the mask that are lowering the continuum in the CCF of the observed spectrum, but are not taken in account in Equation 37; this explains why the ratio F_C is systematically lower than unity and get lower at cooler temperatures, where the lines grow in depth.

5.8.1 *Determination of gravity*

We firstly attempted to calibrate gravity as a function of the three CCF Areas from neutral lines (*low. medium and high* EP) plus the CCF Area using the FeII mask, as done with T_{eff} and [Fe/H], but the resulting space proved to be too hard to fit with a simple multi-variable function. As we stated before, for several fixed values of $\log g$ the Synthetic CCF Area $A_{EW}^{\log g}$ have been determined for values of T_{eff} and $\log g$ spanning our sample of stars,with the correction factor F_C applied. These values are then modeled with a Chebyshev polynomial function of the first kind of order 3x3 as a function of T_{eff} and [Fe/H]. To determine the gravity of a star then T_{eff} and [Fe/H] must be provided as input (using for example the values derived in Section 5.7), the obtained values of $A_{EW}^{\log g}$ are interpolated as a function of the $\log g$ grid and finally the gravity value is determined using the A_{CCF} obtained from the spectra using the FeII mask.

As done for T_{eff} and [Fe/H] in Section 5.7, we test the internal precision of the calibration by comparing the obtained $\log g$for the calibrators woth respect to the assumed values. In Figure 52, the difference between the value obtained from the calibration and the one from EW analysis is plotted against Temperature, Gravity itself and Metallicity (all from EW analysis).

A clear trend with the gravity of the star is shown, and this is a sign that our calibration is not good enough. The cause of this trend has been identified in the determination of the correction factor $F_C(T_{eff}, [Fe/H])$. It turns out to be a function of $\log g$. At the moment we don't have enough data to calibrate the correction factor as a function of gravity, so a different approach has to be taken. In Figure 53 the difference between the CCF Area determined using the FeII mask and the expected synthetic value the difference between the CCF Area

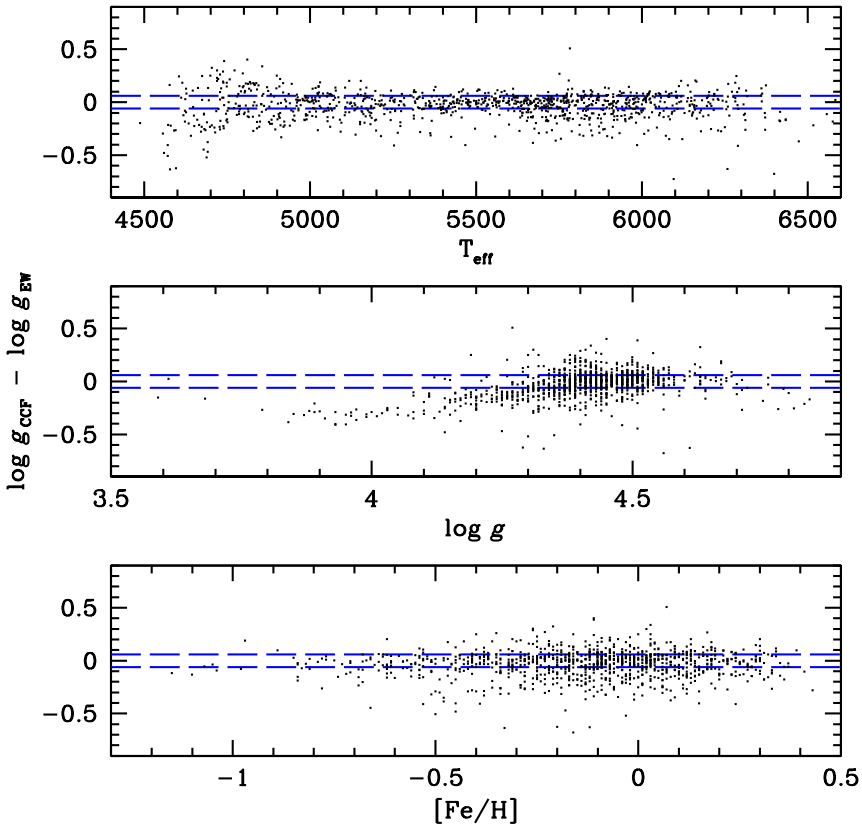

Figure 52: The difference between the value obtained from the calibration and the one from EW analysis is plotted against Temperature, Gravity itself and Metallicity (all from EW analysis). Blue line represents the error in log g from EW analysis

determined using the FeII mask and the expected synthetic value is plotted as a function of the derived gravity. The scatter around the corrected values is $\sigma\Delta Area = 0.026$, which is around $1-2\%$ of the average value of the CCF Area.

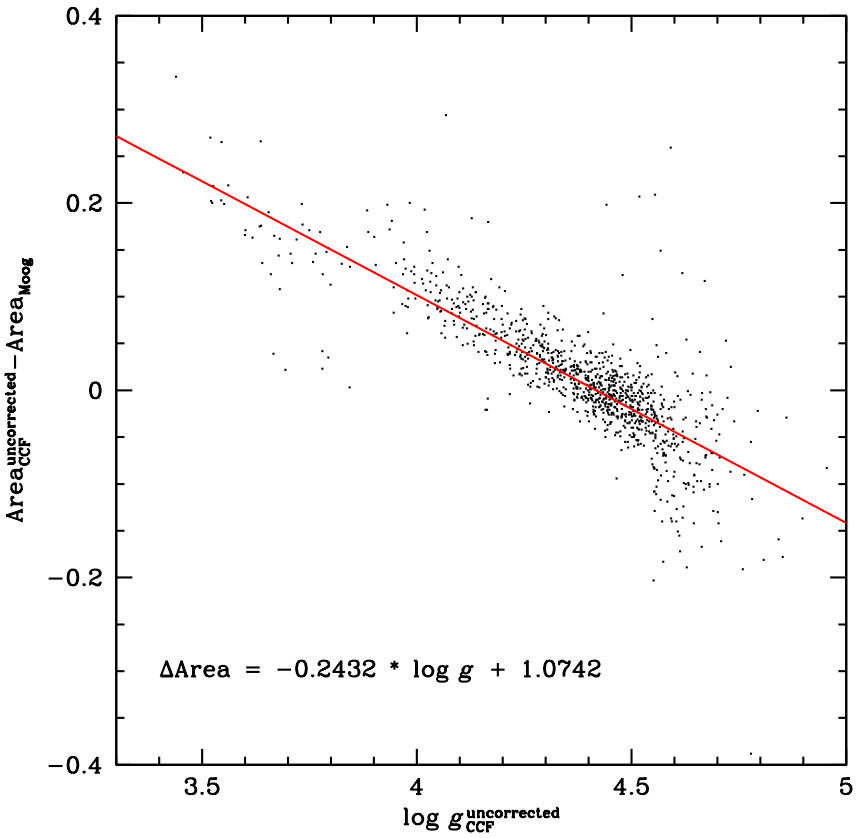

Figure 53: The difference between the CCF Area determined using the FeII mask and the expected synthetic value the difference between the CCF Area determined using the FeII mask and the expected synthetic value is plotted as a function of the derived gravity. A linear fit is performed to correct for the trend.

Using as abcissa the derived values instead of the *correct* ones let us to apply easily this correction to other stars without any assumption: gravity is firstly determined using the measured CCF Area, the Area correction value is determined using the obtained value of gravity and added to the measured CCF Area, and then this value is used to get the final estimate of gravity.

Results for calibrators stars are presented in Figure 54. The dispersion in the residual distribution is $\sigma\log g = 0.055$dex, although at higher gravities the calibration seems to be less reliable.

Single exposure spectra of calibrator stars are finally used to determine the precision of gravity determination as a function of SNR. In Figure 55 gravity has been determined using $T_{eff}$and [Fe/H]from

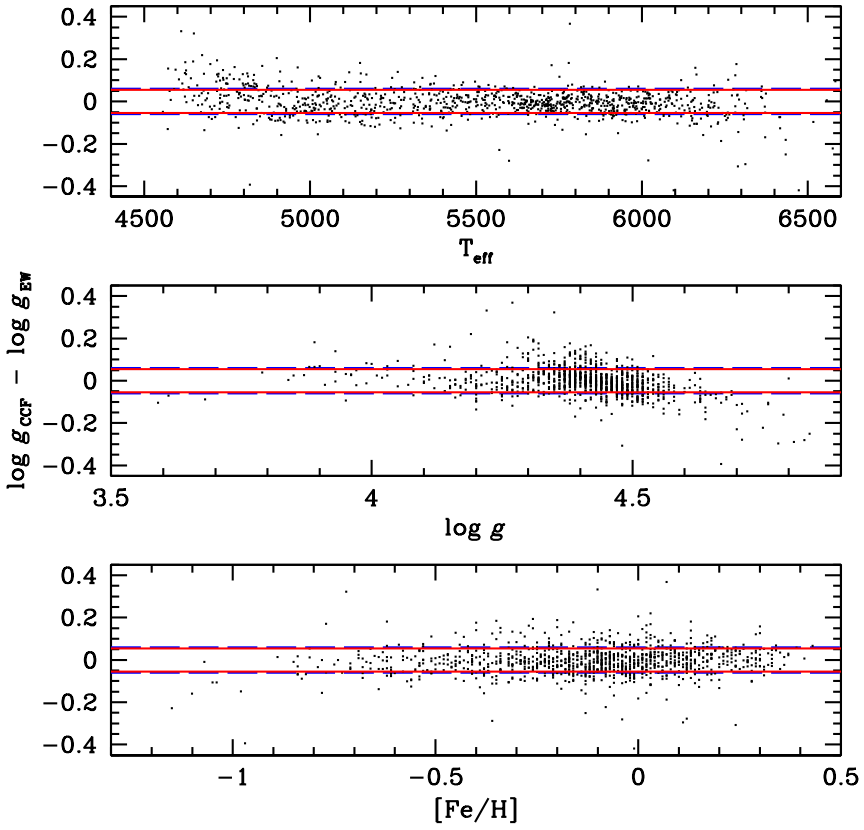

Figure 54: The difference between the value obtained from the calibration and the one from EW analysis is plotted against Temperature, Gravity itself and Metallicity (all from EW analysis). The blue lines represent the error in log g from EW analysis, red lines are the error determined for our calibrators

high SNR spectra. The combined effect of uncertainty in Temperature, Metallicity and Gravity is analyzed in the next section.

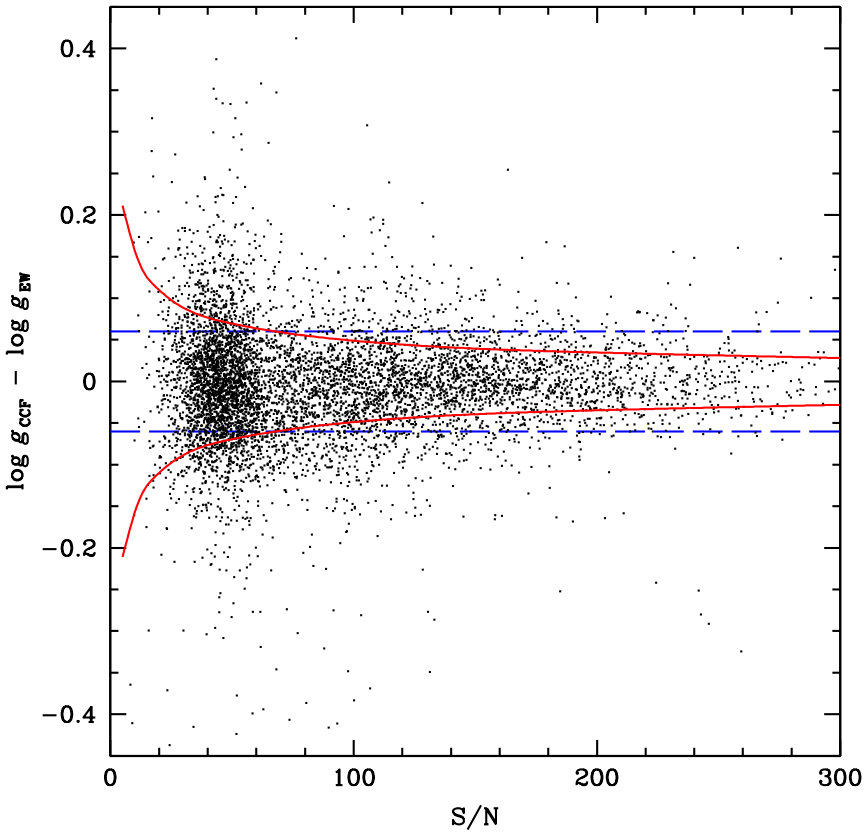

Figure 55: The difference between the gravities derived from EWs measured on high SNR spectra and the ones obtained using CCF analyses of single-exposure spectra is plotted as a function of SNR. Red line represents the dispersion of the distribution at a given SNR.

5.9 GLOBAL DERIVATION OF THE ATMOSPHERE PARAMETERS

We analyze here the derivation of the atmosphere stellar parameters from spectra of different SNR and how the estimates of the paramaters are influencing one each other during their determination. The areas have been derived using the line masks described in the previous sections.

In the first parameter estimation attempt, Temperature and Metallicity are derived using the *direct* calibration introduced in Section 5.7; Gravity is then derived using the obtained stellar parameters as described in Section 5.8.1. It is important to notice that this calibration, although very precise, is reliable only in the range of parameters used to derive the Functions 39 and 40 and when the density of data points is considerable.

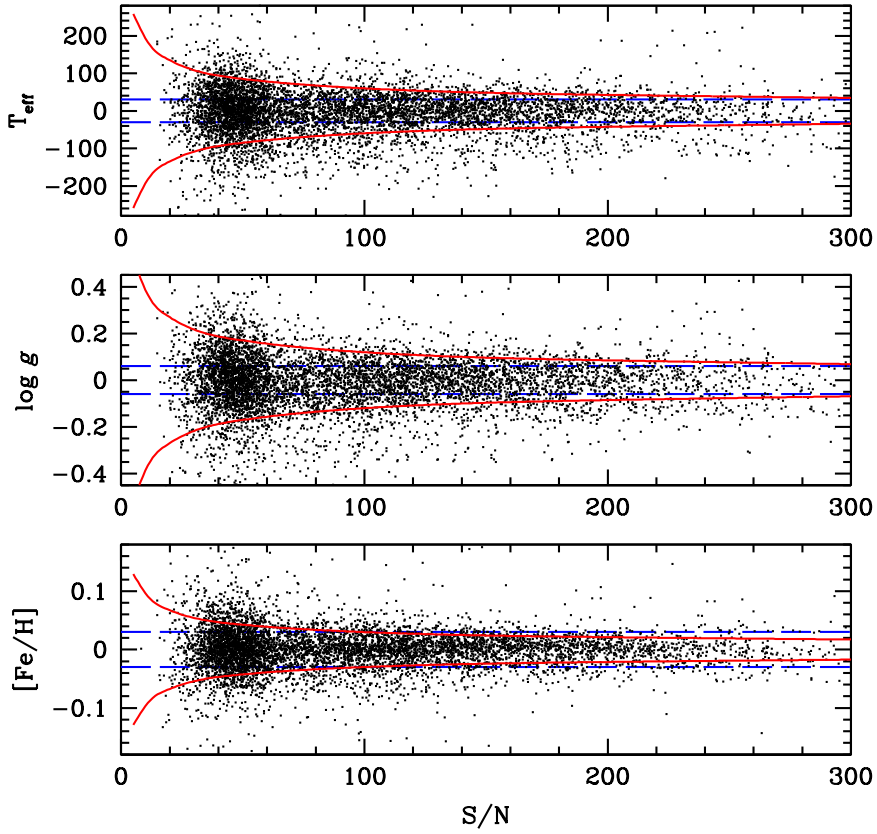

Figure 56: Difference between the derived values from CCF and the EW ones for the three stellar atmosphere parameters as a function of temperature. Temperature and Metallicity have been derived using the *direct* calibration, while Gravity with the *synthetic* one. Blue lines represent the average error of the calibrators. Red lines connecting the $1-\sigma$ of the difference distribution are a good indication of the precision of the determinations.

To determine the precision as a function of SNR, the σ of the distribution of the difference has been computed for several bins of SNR. A simple inverse square-root law of SNR provides a very good fit of the points, proving that our measurements are photon-limited even at low SNR. For SNR = 50 we have a precision of $\sigma_{T_{eff}} = 85$K, $\sigma_{\log g} = 0.17$ dex and $\sigma_{[Fe/H]} = 0.04$ dex, while the accuracy of the technique is provided by the accuracy of the atmosphere parameters of the calibrators.

Our second attempt covers a more extended range in gravity and can be extended to a broader range in temperature and metallicity with adding a reduced number of stars to the calibration sample: the F_C plane needs less data points than Functions 39 and40 to be accurately determined. A similar approach to the one described in Section 5.8 has been followed to determine a theoretical counterpart to Functions 39 and 40, with two calibration functions for each value of the $\log g$ grid; A few iterations between T_{eff}-[Fe/H] and $\log g$ determinations are needed; a value of gravity of $\log g = 4.4$dex is assumed for the first iteration. The result is shown in Figure 57, where the stars in blue are the ones with gravity lower than 4.1 dex.

In this case, the precision reached for $T_{eff}$and $\log g$are slightly higher than the ones obtained with the direct calibration, while no difference is noticeable for $\log g$. Note that there is no trend with gravity for the other parameters when using this calibration with respect to *direct*. The estimated precisions for SNR = 50 are $\sigma_{T_{eff}} = 155$K, $\sigma_{\log g} = 0.17$dex and $\sigma_{[Fe/H]} = 0.07$dex
In both cases parameters have been determined from around 6700 spectra from all the 1111 stars in the sample (we selected a maximum of ten exposure for each star), with 233 stars that have $\log g < 4.1$.

5.10 DISCUSSION

We presented here a new technique to quickly determine stellar atmosphere parameters even for low SNR spectra. With a single exposure we can determine the atmosphere stellar parameters with a precision enough to determine if the observed object have been misclassified. This tool can be extremely useful when performing time series of poorly characterized target, e. g. follow-up of faint Kepler planet-host candidates, to exclude stars that do not match a criteria selection (e. g. that are not solar-type stars) after one or two observations with a considerable optimization of telescope time. The tools are now ready to use, yet we still have space for improvement, by considering for example other elements like Titanium for the determination of gravity.

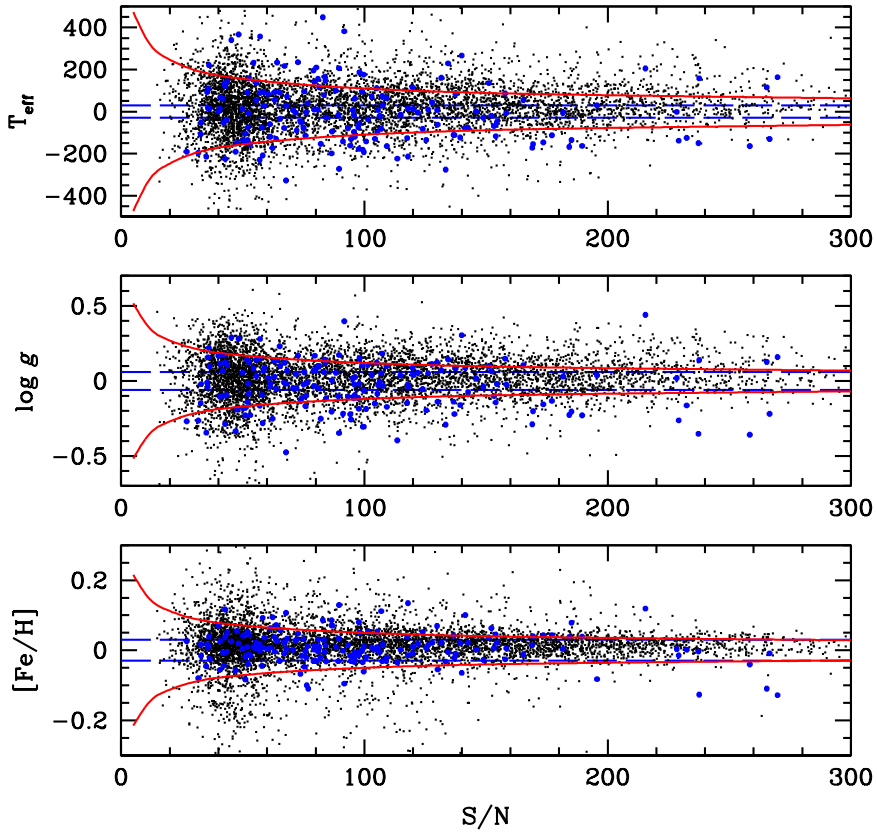

Figure 57: Difference between the derived values from CCF and the EW ones for the three stellar atmosphere parameters as a function of temperature. Parameters have been derived using the *synthetic* calibrations and iterating between T_{eff}-[Fe/H] and log g determination. The $1-\sigma$ line (in red) show the precision of the determinations. Blu points represent stars with log $g < 4.1$dex.

6

FUTURE PROJECTS

With still several years to come before 30 meters class telescopes will be available to the scientific community, we have to take advantage of every opportunity given by current facilities and do every effort to extract as much information as possible from the data at our disposal.

Standard data reduction pipelines are usually designed and tested to have good performance on average SNR data, neglecting the issues that are showing up when moving to lower SNR domains. Their black-box nature allows only a minimal fine-tuning, so the difficulties in the analysis of noisy spectra are worsened by a sub-optimal data reduction. Sometimes the astronomer just need to look at the data with a different perspective from the team who has developed the pipeline.

During the last decade our group has collected thousands of spectra using the GIRAFFE spectrograph at VLT to determine the geometrical distance and dynamical state of several Galactic Globular Clusters (GGCs) but ultimately the analysis has been hampered by systematics in data calibration and radial velocity measurements, see for example [Milone et al., 2006] and [Sommariva et al., 2009]. Well aware of these issues, we have focused our attention on the development of an alternative pipeline specifically designed to delivery state-of-the-art data reduction, wavelength calibration and radial velocity determination of very low SNR stellar spectra.

To effectively show the improvement introduced by our analysis with respect to standard procedures, we have chosen a pragmatical approach. After describing the theoretical aspects of the techniques we have introduced, we put them *in action* on data already analyzed with standard tools to highlight and resolve the unavoidable issues of practical implementation and directly show the results of our approach.

The dataset we have analyzed in this thesis comprises 7250 spectra for 2771 stars of the GCC M 4 (NGC 6121) in the wavelength region $5145 - 5360$Å obtained with GIRAFFE. Stars from the upper Red Giant Branch down to the Main Sequence have been observed in very different conditions, including nights close to full moon, and reaching SNR $\simeq 10$ for many spectra in the dataset.

Improving the data reduction and extraction has been the topic of Chapter 2. We have carefully reviewed the often-neglected details of basic data reduction techniques, like dark master determination from a reduced number of frames or correction for pixel-to-pixel gain variations, listing several ad-hoc solutions. We have applied to spec-

troscopy the well-known technique of the *empirical PSF*, which has allowed our group to obtain Color-Magnitude Diagrams with unsurpassed quality and push the study of GGCs into a new direction [Piotto et al., 2007]. The improvement in the determination of the spectral profile in the cross-dispersion direction results in a measurement of the position of the profile so precise that the $\simeq 512$th anomaly of the CCD is clearly visible as a jump in the position of the fibres. Despite the data at our disposal are not optimal for the ePSF technique since the spectra are almost perpendicular to the cross-dispersion direction, we get anyway a decrease of the noise in low SNR spectra extracted with our tools with respect to the ESO pipeline.

In Chapters 3 we have presented a new approach for the determination of accurate wavelength dispersion solutions and radial velocity shifts using simultaneous calibration fibres for not-stabilized instruments like GIRAFFE. The derived cluster radial velocity is $\langle v \rangle = 70.896 \pm 0.077$km/s, in agreement with previous measurements in literature but with an offset of $\Delta v = 0.63$km/s with the previous value derived using the same dataset by Sommariva et al. [2009]. This offset is probably due to the different technique use to determine the radial velocity. The gain in precision results in an increased number of binary candidates with respect to the one determined by the previous analysis of Sommariva et al. [2009]. We found 96 binary candidates (58 in the previous analysis), of which 30 among giants and 66 among sub-giants and dwarfs.

An average radial velocity dispersion of 4.5km/s within 3 arcsec from the center of cluster and steadingly decreasing outward has been found, in contrast with the nearly constant value of 3.5km/s from Peterson et al. [1995]. We have also attempted a preliminary study on the cluster rotation, confirming a rotation of amplitude -0.70 ± 0.2km/s and phase of $120 \pm 10^\circ$ for stars farther than $5r$; a weaker sign of rotation (amplitude -0.40 ± 0.2)km/s) seems to be present in the inner sample too, but a deeper and more detailed analysis is required.

In Chapter 5 we present a new algorithm to automatically determine the continuum normalization level for spectra characterized by a very short and almost continuum-less wavelength range, very low SNR and strong contamination from sky background. his determination is completely automatic and does not require any human interaction, thus removing one of the biggest source of uncertainties in the synthesis comparison with spectra observed in a very low SNR regime. Atmosphere stellar parameters have been derived by comparison with a fine and extensive grid of spectral syntheses. In order to avoid the introduction of any false correlations into the spectrum noise, all the steps that usually lead to sky-corrected, rest-frame, continuum-normalized stellar spectra have been instead performed in the opposite direction: deliberate degredation of the *perfect* synthetic

spectra to best match the real extracted spectra. To increase the accuracy of the parameter determination, data points have been weighted according to the sensitivity of spectral features with respect to variations in the atmosphere parameters.

For 322 RGB stars down to $V \simeq 16$ we find a constant metallicity along the entire sequence with an average value of $[Fe/H] = -1.068 \pm 0.001$ dex and a dispersion around the mean of $\sigma = 0.025$ dex. This result is consistent with the one derived in Marino et al. [2008] from 105 high-resolution UVES spectra of bright stars ($V < 14$) and using Equivalent Width analysis. We stress the fact that not only the same results are recovered despite using stars with lower spectral resolution and a wavelength range that is almost ten times shorter, but we are able to extend the analysis to stars that are two magnitudes fainter than the ones in the high-resolution sample.

For stars in the Sub-Giant Branch and Main Sequence the analysis has been made difficult by the lose of sensitivity of our spectra to micro-turbulent velocity. From 1869 stars in the SGB and MS we have an average metallicity of $[Fe/H] = -1.162 \pm 0.002$ dex with a dispersion of $\sigma = 0.09$ dex, after fixing the micro-turbulent velocity at $\xi_t = 0.9$km/s. Although a decrease in surface metallicity around the Turn-Off has been already observed in others GGC and attributed to diffusive processes (sedimentation, levitation) coupled with turbulent mixing below the outer convection, see ad example Lind et al. [2008], we prefer to perform extended analyses and tests before confirming this result.

Our plan in the next future is to characterize the radial velocity dispersion and rotation of other globular clusters that have been surveyed by FLAMES with the improved precision delivered by our alternative calibration. Several thousand of spectra are already available in the ESO Archive for the GGCs NGC 6752, NGC 6397, NGC 104 (47 Tucanae) and NGC 2808, all GGCs that have shown interesting features in their Color-Magnitude Diagrams [Piotto et al., 2007, Milone et al., 2010, Piotto et al., 2012]. The gain in precision will make possible the comparison with theoretical model like Heggie and Giersz [2008] for an extended number of cluster.

The spectral range we have used as benchwork for our project has proved to be quite poor of spectral lines from *s-* and *r-process* as well as light elements, so our immediate goal is to extend the automatic stellar atmospheric parameter determination to other spectral ranges (nominally from 6300 Å to 6800 Å , see Gonzalez et al. 2011). Once that reliable sky subtraction, continuum normalization and atmosphere parameters are obtained, the inclusion in our tools of abundance determination for chemical elements other than Iron by comparison of the observed spectra with synthesis is straightforward.

In the last chapter of my thesis we have decided to face the atmospheric parameter determination problem from an original per-

spective. From a qualitative point of view it is well know that the shape of Cross-Correlation Function (CCF) depends somehow by the characteristics of the spectrum under analysis. A recent work in this direction has been made by Buchhave et al. [2012], with the stellar parameters determined by cross-correlating the observed spectrum with a library of synthetic spectra and selecting the combination of parameters that results in the most peaked CCF. Torres et al. [2012] have shown however that strong correlations exist between temperatures, gravities and metallicities derived with this technique. Our approach differs substantially from the previous one. In Chapter 5 we used several *CORAVEL-type* CCFs obtained using different line lists, where neutral iron lines are selected according to their sensitivity to changes in temperature to break the temperature-metallicity degeneracy. We calibrated the CCF parameters with respect to atmosphere parameters obtained with Equivalent Width analysis of high-resolution and high SNR spectra. We have used the data of 1111 stars from several HARPS long-term programs during the last ten years to determine to different calibrations. The first is purely empirical and it provides temperature and metallicity from the observed CCF areas for stars with gravity in the range $\log g 4.2 - 4.6$ (solar-type stars) with a precision comparable to the one obtainable with equivalent width measurements. The second calibration is related to theoretical equivalent widths by the analytical approximations described in the chapter, and by the use of an additional mask with ionized Iron lines it provides also an estimate of gravity. This second calibration works on a more extended range in gravity, with the drawback that the derived parameters are slightly less precise than the first calibration. We have tested the performance of our calibration using the individual exposures available for each star. The tools are now ready to use, yet we still have space for improvement, by considering for example other elements like Titanium for the determination of gravity.

With the collaboration of the Exoplanet Group at Observatory of Geneva, we plan to include the atmosphere parameters determination from CCFs into the HARPS/HARPS-N pipelines within a few months. With this tool we will be able to deliver precise atmospheric parameters even for the faintest stars hosting candidate planets detected by the *Kepler* space mission that will be targeted by the follow-up program inside the Guaranteed Time Observations with HARPS-N at the Telescopio Nazionale Galileo. Additionally, since each candidate requires between ten and one hundred hours of telescope time to be confirmed (with a single exposure already taking one hour of time, due to faintness of the targets), and the Kepler mission has provided several thousands of them, excluding all the stars with mistaken photometric atmosphere parameters, i. e. giants classified as dwarfs, will provide a huge advantage in term of use of telescope time.

BIBLIOGRAPHY

V Zh Adibekyan, S G Sousa, N C Santos, E Delgado Mena, J I González Hernández, G Israelian, M Mayor, and G. Khachatryan. Chemical abundances of 1111 FGK stars from the HARPS GTO planet search program. Galactic stellar populations and planets. *A&A*, 545:32, September 2012.

C Allende Prieto, David L Lambert, Robert G Tull, and Phillip J MacQueen. Convective Wavelength Shifts in the Spectra of Late-Type Stars. *ApJ*, 566(2):L93–L96, February 2002.

J Anderson and I R King. Toward High-Precision Astrometry with WFPC2. I. Deriving an Accurate Point-Spread Function. *PASP*, 112: 1360, October 2000.

J Anderson, R R Bedin, G Piotto, R S Yadav, and A Bellini. Ground-based CCD astrometry with wide field imagers. I. Observations just a few years apart allow decontamination of field objects from members in two globular clusters. *A&A*, 454(3):1029–1045, August 2006.

J Anderson, G Piotto, I R King, R R Bedin, and P Guhathakurta. Mixed Populations in Globular Clusters: Et Tu, 47 Tuc? *ApJL*, 697: L58, May 2009.

A Baranne, M Mayor, and J L Poncet. Coravel— A new tool for radial velocity measurements. *Vistas in Astronomy*, 23:279–316, January 1979.

A Baranne, D Queloz, M Mayor, G Adrianzyk, G Knispel, D Kohler, D Lacroix, J-P Meunier, G Rimbaud, and A Vin. ELODIE: A spectrograph for accurate radial velocity measurements. *Astron Astrophys Sup*, 119:373, October 1996.

Andre Blecha, Veronique Cayatte, Pierre North, Frederic Royer, and Gilles Simond. Data-reduction software for GIRAFFE, the VLT medium-resolution multi-object fiber-fed spectrograph. *Proc. SPIE Vol. 4008*, 4008:467, August 2000.

A Bragaglia, E Carretta, R Gratton, S Lucatello, A P Milone, G Piotto, V D'Orazi, S Cassisi, C Sneden, and R R Bedin. X-shooter Observations of Main-sequence Stars in the Globular Cluster NGC 2808: First Chemical Tagging of a He-normal and a He-rich Dwarf. *ApJL*, 720(1):L41–L45, September 2010.

Jeffery A Brown and George Wallerstein. High-resolution CCD spectra of stars in globular clusters. VII - Abundances of 16 elements in 47 Tuc, M4, and M22. *AJ*, 104:1818–1830, November 1992.

Lars A Buchhave, D W Latham, Anders Johansen, Martin Bizzarro, G Torres, Jason F Rowe, N M Batalha, William J Borucki, Erik Brugamyer, Caroline Caldwell, Stephen T Bryson, David R Ciardi, William D Cochran, Michael Endl, Gilbert A Esquerdo, E B Ford, John C Geary, Ronald L Gilliland, Terese Hansen, Howard Isaacson, John B Laird, Philip W Lucas, G W Marcy, Jon A Morse, Paul Robertson, Avi Shporer, Robert P Stefanik, Martin Still, and S N Quinn. An abundance of small exoplanets around stars with a wide range of metallicities. *Nature*, 486(7):375–377, June 2012.

E Carretta, A Bragaglia, R Gratton, A Recio-Blanco, S Lucatello, V D'Orazi, and S Cassisi. Properties of stellar generations in globular clusters and relations with global parameters. *A&A*, 516:55, June 2010.

L Casagrande, I Ramírez, J Meléndez, M Bessell, and M Asplund. An absolutely calibrated Teff scale from the infrared flux method. Dwarfs and subgiants. *A&A*, 512:54, March 2010.

F Castelli and R L Kurucz. New Grids of ATLAS9 Model Atmospheres. *arXiv*, astro-ph, May 2004.

P Coelho, B Barbuy, J Meléndez, R P Schiavon, and B V Castilho. A library of high resolution synthetic stellar spectra from 300 nm to 1.8 μm with solar and α−enhanced composition. *A&A*, 443(2):735–746, November 2005.

Patrick Côté and Philippe Fischer. Spectroscopic Binaries in Globular Clusters. I. A Search for Ultra-Hard Binaries on the Main Sequence in M4. *Astronomical Journal v.112*, 112:565, August 1996.

K M Cudworth and R Rees. Astrometry and photometry in the globular cluster M4. *AJ*, 99:1491–1500, May 1990.

F D'Antona, P B Stetson, Paolo Ventura, A P Milone, G Piotto, and V Caloi. The influence of the C+N+O abundances on the determination of the relative ages of globular clusters: the case of NGC1851 and NGC6121 (M4). *Monthly Notices of the Royal Astronomical Society: Letters*, page L319, September 2009.

R. I. Dixon and A. J. Longmore. Fundamental Parameters for m4 the Nearest Globular Cluster. *Monthly Notice of the Royal Astronomical Society*, 265:395, November 1993.

V D'Orazi and A F Marino. Lithium Abundances in Red Giants of M4: Evidence for Asymptotic Giant Branch Star Pollution in Globular Clusters? *ApJL*, 716(2):L166–L169, June 2010.

J J Drake, Verne V Smith, and Nicholas B Suntzeff. Sodium, aluminum, and oxygen abundance variations in giants in the globular cluster M4. *ApJ*, 395:L95–L98, August 1992.

J J Drake, V V Smith, and N B Suntzeff. The metallicity of M4: Accurate spectroscopic fundamental parameters for four giants. *ApJ*, 430:610–623, August 1994.

D Enard. European Southern Observatory (ESO) coudeechelle spectrometer. *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 331:232–242, October 1982.

K C Freeman and J E Norris. The chemical composition, structure, and dynamics of globular clusters. *Annual Reviews in Astronomy and Astrophysics*, 19:319–356, January 1981.

D Geisler. Magnesium, silicon, and iron abundances and the cluster metallicity scale. *ApJ*, 287:L85–L89, December 1984.

O A Gonzalez, M Rejkuba, M Zoccali, V Hill, G Battaglia, C Babusiaux, D Minniti, B Barbuy, A Alves-Brito, A Renzini, A Gómez, and S Ortolani. Alpha element abundances and gradients in the Milky Way bulge from FLAMES-GIRAFFE spectra of 650 K giants. *A&A*, 530:54, June 2011.

R Gratton, M. L. Quarta, and S Ortolani. The metal abundance of metal-rich globular clusters. II. *Astronomy and Astrophysics (ISSN 0004-6361)*, 169:208–218, November 1986.

R Gratton, E Carretta, and F Castelli. Abundances of light elements in metal-poor stars. I. Atmospheric parameters and a new T_eff_ scale. *A&A*, 314:191–203, October 1996.

R Gratton, C Sneden, and E Carretta. Abundance Variations Within Globular Clusters. *Annual Review of Astronomy &Astrophysics*, 42: 385, September 2004.

D. F. Gray. *The Observation and Analysis of Stellar Photospheres*. September 2005.

W E Harris. A Catalog of Parameters for Globular Clusters in the Milky Way. *AJ*, 112:1487, October 1996.

Douglas C Heggie and Mirek Giersz. Monte Carlo simulations of star clusters - V. The globular cluster M4. *Monthly Notice of the Royal Astronomical Society*, 389(4):1858–1870, October 2008.

Benjamin Hendricks, Peter B Stetson, Don A Vandenberg, and M Dall'Ora. A New Reddening Law for M4. *AJ*, 144(1):25, July 2012.

K. Hinkle, L. Wallace, J. Valenti, and D. Harmer. *Visible and Near Infrared Atlas of the Arcturus Spectrum 3727-9300 A*. Not Available, 2000.

K Horne. An optimal extraction algorithm for CCD spectroscopy. *PASP*, 98:609, June 1986.

M L Houdashelt, R A Bell, and A V Sweigart. Improved Color-Temperature Relations and Bolometric Corrections for Cool Stars. *AJ*, 119:1448, March 2000.

I I Ivans, C Sneden, R P Kraft, Nicholas B Suntzeff, Verne V Smith, G Edward Langer, and Jon P Fulbright. Star-to-Star Abundance Variations among Bright Giants in the Mildly Metal-poor Globular Cluster M4. *AJ*, 118(3):1273–1300, September 1999.

I I Ivans, Jennifer Simmerer, C Sneden, James E Lawler, John J Cowan, Roberto Gallino, and Sara Bisterzo. Near-Ultraviolet Observations of HD 221170: New Insights into the Nature of r-Process–rich Stars. *ApJ*, 645(1):613–633, July 2006.

H L Johnson and W W Morgan. Fundamental stellar photometry for standards of spectral type on the revised system of the Yerkes spectral atlas. *Astrophysical Journal*, 117:313, May 1953.

R P Kraft. On the nonhomogeneity of metal abundances in stars of globular clusters and satellite subsystems of the Galaxy. *Annual Reviews in Astronomy and Astrophysics*, 17:309–343, January 1979.

R P Kraft. Abundance differences among globular-cluster giants: Primordial versus evolutionary scenarios. *PASP*, 106:553, June 1994.

R P Kraft and I I Ivans. A Globular Cluster Metallicity Scale Based on the Abundance of Fe II. *PASP*, 115(8):143–169, February 2003.

F. G. Kupka, T. A. Ryabchikova, N E Piskunov, H C Stempels, and W. W. Weiss. VALD-2 – The New Vienna Atomic Line Database. *Baltic Astronomy*, 9:590–594, January 2000.

Michael J Kurtz and Douglas J Mink. RVSAO 2.0: Digital Redshifts and Radial Velocities. *PASP*, 110(7):934–977, August 1998.

R L Kurucz. Model Atmospheres for Population Synthesis. In *The Stellar Populations of Galaxies: Proceedings of the 149th Symposium of the International Astronomical Union*, page 225, January 1992.

R. L. Kurucz, I. Furenlid, J. Brault, and L. Testerman. *Solar flux atlas from 296 to 1300 nm*. National Solar Observatory, 1984.

K Lind, A J Korn, P S Barklem, and F Grundahl. Atomic diffusion and mixing in old stars. II. Observations of stars in the globular cluster NGC 6397 with VLT/FLAMES-GIRAFFE. *A&A*, 490(2):777–786, November 2008.

G Lo Curto, M Mayor, W Benz, F Bouchy, C Lovis, C Moutou, D Naef, F Pepe, D Queloz, N C Santos, D Ségransan, and S Udry. The HARPS search for southern extra-solar planets . XXII. Multiple planet systems from the HARPS volume limited sample. *A&A*, 512: 48, March 2010.

C Lovis and F Pepe. A new list of thorium and argon spectral lines in the visible. *A&A*, 468(3):1115–1121, June 2007.

A F Marino, S Villanova, G Piotto, A P Milone, Y Momany, R R Bedin, and A M Medling. Spectroscopic and photometric evidence of two stellar populations in the Galactic globular cluster NGC 6121 (M4). *A&A*, 490(2):625–640, November 2008.

A F Marino, S Villanova, A P Milone, G Piotto, K Lind, D Geisler, and P B Stetson. Sodium-Oxygen Anticorrelation Among Horizontal Branch Stars in the Globular Cluster M4. *ApJL*, 730(2):L16, April 2011.

Alessandro Massarotti, D W Latham, Robert P Stefanik, and Jeffrey Fogel. Rotational and Radial Velocities for a Sample of 761 HIPPARCOS Giants and the Role of Binarity. *AJ*, 135(1):209–231, January 2008.

M Mayor, F Pepe, D Queloz, F Bouchy, G Rupprecht, G Lo Curto, G Avila, W Benz, J-L Bertaux, X Bonfils, T H Dall, H Dekker, B Delabre, W Eckert, M Fleury, A Gilliotte, D Gojak, J C Guzman, D Kohler, J-L Lizon, A Longinotti, C Lovis, D Megevand, L Pasquini, J Reyes, J-P Sivan, D Sosnowska, R Soto, S Udry, A Van Kesteren, L Weber, and U Weilenmann. Setting New Standards with HARPS. *The Messenger*, 114:20, December 2003.

A P Milone, Sandro Villanova, R R Bedin, G Piotto, Giovanni Carraro, J Anderson, I R King, and S Zaggia. Absolute motions of globular clusters. II. HST astrometry and VLT radial velocities in NGC 6397. *A&A*, 456:517, September 2006.

A P Milone, P B Stetson, G Piotto, R R Bedin, J Anderson, S Cassisi, and Maurizio Salaris. The radial distribution of the two stellar populations in NGC 1851. *A&A*, 503:755, September 2009.

A P Milone, G Piotto, I R King, R R Bedin, J Anderson, A F Marino, Y Momany, Luca Malavolta, and S Villanova. Multiple Stellar Populations in the Galactic Globular Cluster NGC 6752. *ApJ*, 709(2): 1183–1194, February 2010.

A P Milone, G Piotto, R R Bedin, A Aparicio, J Anderson, A Sarajedini, A F Marino, A Moretti, M B Davies, B. Chaboyer, A. Dotter, M Hempel, A Marín-Franch, S. Majewski, N. E. Q. Paust, I N Reid, A. Rosenberg, and M H Siegel. The ACS survey of Galactic globular

clusters. XII. Photometric binaries along the main sequence. *A&A*, 540:16, April 2012.

M T Murphy, P Tzanavaris, J K Webb, and C Lovis. Selection of ThAr lines for wavelength calibration of echelle spectra and implications for variations in the fine-structure constant. *Monthly Notice of the Royal Astronomical Society*, 378:221, June 2007.

Michael T. Murphy, Clayton R. Locke, Philip S. Light, Andre N. Luiten, and Jon S. Lawrence. Laser frequency comb techniques for precise astronomical spectroscopy. *Monthly Notice of the Royal Astronomical Society*, 422(1):761–771, May 2012.

V Neves, N C Santos, S G Sousa, A C M Correia, and G Israelian. Chemical abundances of 451 stars from the HARPS GTO planet search program. *A&A*, 497(2):563–581, April 2009.

J E Norris. The cyanogen distribution of M4 and the possible connection between horizontal branch morphology and chemical inhomogeneity. *ApJ*, 248:177–188, August 1981.

L Pasquini, Gerardo Avila, Eric Allaert, Pascal Ballester, Peter Biereichel, Bernard Buzzoni, Cyril Cavadore, Hans Dekker, Bernard Delabre, F Ferraro, Vanessa Hill, A Kaufer, Heinz Kotzlowski, Jean-Louis Lizon, Antonio Longinotti, S. Moureau, R. Palsa, and S Zaggia. FLAMES: a multi-object fiber facility for the VLT. *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 4008: 129–140, August 2000.

F Pepe, M Mayor, F Galland, D Naef, D Queloz, N C Santos, S Udry, and M Burnet. The CORALIE survey for southern extra-solar planets VII. Two short-period Saturnian companions to HD 108147 and HD 168746. *A&A*, 388:632–638, June 2002.

F Pepe, M Mayor, D Queloz, W Benz, X Bonfils, F Bouchy, G Lo Curto, C Lovis, D Megevand, C Moutou, D Naef, G Rupprecht, N C Santos, J P Sivan, D Sosnowska, and S Udry. The HARPS search for southern extra-solar planets. I. HD 330075 b: A new "hot Jupiter". *A&A*, 423:385–389, August 2004.

R C Peterson, Richard F Rees, and Kyle M Cudworth. Radial velocities of stars in the globular cluster M4 and the cluster distance. *ApJ*, 443:124, April 1995.

G Piotto, Sandro Villanova, R R Bedin, R Gratton, S Cassisi, Yazan Momany, Alejandra Recio-Blanco, S Lucatello, J Anderson, I R King, A Pietrinferni, and Giovanni Carraro. Metallicities on the Double Main Sequence of ω Centauri Imply Large Helium Enhancement. *ApJ*, 621:777, March 2005.

G Piotto, R R Bedin, J Anderson, I R King, S Cassisi, A P Milone, Sandro Villanova, A Pietrinferni, and Alvio Renzini. A Triple Main Sequence in the Globular Cluster NGC 2808. *ApJ*, 661:L53, May 2007.

G Piotto, A P Milone, J Anderson, R R Bedin, A Bellini, S Cassisi, A F Marino, A Aparicio, and V Nascimbeni. Hubble Space Telescope Reveals Multiple Sub-giant Branch in Eight Globular Clusters. *ApJ*, 760(1):39, November 2012.

N E Piskunov, F Kupka, T. A. Ryabchikova, W. W. Weiss, and C S Jeffery. VALD: The Vienna Atomic Line Data Base. *Astron Astrophys Sup*, 112:525, September 1995.

William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. Numerical recipes in FORTRAN. The art of scientific computing. *Nature*, -1, 1992.

I Ramírez and Jorge Meléndez. The Effective Temperature Scale of FGK Stars. II. Teff:Color:[Fe/H] Calibrations. *ApJ*, 626:465, June 2005.

Frederic Royer, Andre Blecha, Pierre North, Gilles Simond, S Baratchart, Veronique Cayatte, Laurent Chemin, and Ralf Palsa. Toward accurate radial velocities with the fiber-fed GIRAFFE multi-object VLT spectrograph. *Astronomical Data Analysis II. Edited by Starck*, 4847:184, December 2002.

N C Santos, G Israelian, and M Mayor. Spectroscopic [Fe/H] for 98 extra-solar planet-host stars. Exploring the probability of planet formation. *A&A*, 415:1153, March 2004.

N C Santos, M Mayor, X Bonfils, X Dumusque, F Bouchy, P Figueira, C Lovis, C Melo, F Pepe, D Queloz, D Ségransan, S G Sousa, and S Udry. The HARPS search for southern extrasolar planets. XXV. Results from the metal-poor sample. *A&A*, 526:112, February 2011.

G H Smith and M M Briley. CN Abundance Inhomogeneities in the Globular Cluster Messier 4 (NGC 6121): Results Based on Merged Data Sets from the Literature. *PASP*, 117(8):895–910, September 2005.

C Sneden. The nitrogen abundance of the very metal-poor star HD 122563. *Astrophysical Journal*, 184:839–849, September 1973.

V Sommariva, G Piotto, M Rejkuba, R R Bedin, D C Heggie, R D Mathieu, and S Villanova. A search for spectroscopic binaries in the Galactic globular cluster M4. *A&A*, 493(3):947–958, January 2009.

S G Sousa, N C Santos, G Israelian, M Mayor, and M J P F G Monteiro. A new code for automatic determination of equivalent widths:

Automatic Routine for line Equivalent widths in stellar Spectra (ARES). *A&A*, 469:783, July 2007.

S G Sousa, A Alapini, G Israelian, and N C Santos. An effective temperature calibration for solar type stars using equivalent width ratios. *A&A*, 512:A13, March 2010.

S G Sousa, N C Santos, G Israelian, M Mayor, and S Udry. Spectroscopic stellar parameters for 582 FGK stars in the HARPS volume-limited sample. *A&A*, 533:A141, September 2011.

Peter B Stetson, J E Hesser, and T A Smecker Hane. Homogeneous Photometry for Star Clusters and Resolved Galaxies. I. A Survey of Bright Stars in the Fornax Dwarf Spheroidal Galaxy. *PASP*, 110(7): 533–552, May 1998.

P Stumpff. Two Self-Consistent FORTRAN Subroutines for the Computation of the Earth's Motion. *Astron Astrophys Sup*, 41:1, June 1980.

Y Takeda, H Kaneko, N Matsumoto, S Oshino, H Ito, and T Shibuya. Potassium Abundances in Red Giants of Mildly to Very Metal-Poor Globular Clusters. *Publications of the Astronomical Society of Japan*, 61: 563–, June 2009.

J Tonry and M Davis. A survey of galaxy redshifts. I - Data reduction techniques. *Astronomical Journal*, 84:1511–1525, October 1979.

G Torres, Debra A Fischer, Alessandro Sozzetti, Lars A Buchhave, Joshua N Winn, M J Holman, and Joshua A Carter. Improved Spectroscopic Parameters for Transiting Planet Hosts. *ApJ*, 757(2):161, October 2012.

T Wilken, C Lovis, A Manescau, T Steinmetz, L Pasquini, G Lo Curto, T W Hänsch, R Holzwarth, and Th Udem. High-precision calibration of spectrographs. *Monthly Notices of the Royal Astronomical Society: Letters*, 405:L16, June 2010.

David Yong, Amanda I Karakas, David L Lambert, Alessandro Chieffi, and Marco Limongi. Heavy Element Abundances in Giant Stars of the Globular Clusters M4 and M5. *ApJ*, 689(2):1031–1043, December 2008a.

David Yong, David L Lambert, Diane B Paulson, and Bruce W Carney. Rubidium and Lead Abundances in Giant Stars of the Globular Clusters M4 and M5. *ApJ*, 673(2):854–863, February 2008b.