Flat-relative optimal extraction*
A quick and efficient algorithm for stabilised spectrographs

M. Zechmeister¹, G. Anglada-Escudé¹,², and A. Reiners¹

¹ Institut für Astrophysik, Georg-August-Universität, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany
² Astronomy Unit, Queen Mary University of London, Mile End Road, London E1 4NS, UK

Received / Accepted

ABSTRACT

Context. Optimal extraction is a key step in processing the raw images of spectra as registered by two-dimensional detector arrays to a one-dimensional format. Previously reported algorithms reconstruct models for a mean one-dimensional spatial profile to assist a properly weighted extraction.

Aims. We outline a simple optimal extraction algorithm including error propagation, which is very suitable for stabilised, fibre-fed spectrographs and does not model the spatial profile shape.

Methods. A high signal-to-noise, master-flat image serves as reference image and is directly used as an extraction profile mask. Each extracted spectral value is the scaling factor relative to the cross-section of the unnormalised master-flat which contains all information about the spatial profile as well as pixel-to-pixel variations, fringing, and blaze. The extracted spectrum is measured relative to the flat spectrum.

Results. Using echelle spectra of the HARPS spectrograph we demonstrate a competitive extraction performance in terms of signal-to-noise and show that extracted spectra can be used for high precision radial velocity measurement.

Conclusions. Pre- or post-flat-fielding of the data is not necessary, since all spectrograph inefficiencies inherent to the extraction mask are automatically accounted for. Also the reconstruction of the mean spatial profile by models is not needed, thereby reducing the number of operations to extract spectra. Flat-relative optimal extraction is a simple, efficient, and robust method that can be applied easily to stabilised, fibre-fed spectrographs.

Key words. instrumentation: spectrographs – methods: data analysis – techniques: image processing – techniques: radial velocities

1. Introduction

Spectra of astronomical objects provide a wealth of information, and the increasing need for higher precision has led to the development of very stable and fibre-fed spectrographs. A prospering example is the search for exoplanets with Doppler spectroscopy. Cross-dispersed, high-resolution echelle spectrographs are usually employed to measure radial velocities at a precision of 1 m/s, which corresponds to about 1/1000 of a pixel. Therefore careful calibration, as well as high mechanical, thermal, and pressure stability, is essential.

If aiming for high precision, not only is the hardware important, but also the software algorithms to process the images. Typically, a reduction of echelle spectra consists of several steps: bias subtraction, dark subtraction, scattered light subtraction, flat-fielding, extraction to 1D, deblazing, wavelength calibration, order merging, flux normalisation (e.g. Baranne et al. 1996). In this work we focus on extraction, which is a crucial step in image processing. A widely used method is the so-called optimal extraction (Horne 1986) and its variants (e.g. Marsh 1989; Piskunov & Valenti 2002, Table 1), which basically scales 1D cross-sectional profiles to the imaged spectrum, and the scaling factor is the best flux estimate. Additionally, most algorithms try to model and reconstruct the spatial profile/slit function with polynomial, Gaussian, or other smooth functions. However, for stabilised spectrographs, the order profiles and positions are object- and time-independent, which simplifies the extraction; in particular, there is not necessarily any need to model the spatial profile with empirical functions. We exploit this circumstance, and we derive and test our concept of flat-relative optimal extraction.

2. Principle of flat-relative optimal extraction (FOX)

In the following, we assume that the image processing steps (e.g., bias, dark, and background subtraction) preceding extraction are properly done and that, in the image, the main dispersion is oriented in a more or less horizontal direction (x) and the cross-dispersion (echelle orders) in a vertical direction (y).

A spectrograph consists of dispersive elements and a camera that images the slit or the fibre exit to wavelength-dependent positions and shapes. The observed light distribution $S(x,y)$ is a convolution of the input spectrum $s(\lambda)$ with a wavelength-dependent instrumental point spread function (iPSF) of the spectrograph $\Psi(x,y,\lambda)$. Therefore a model $\hat{S}(x,y)$ for the observed spectrum can be formulated as

$$\hat{S}(x,y) = \Psi(x,y,\lambda) \otimes s(\lambda)$$

where wavelength $\lambda$ and the positions $x$ and $y$ in the detector plane are continuous variables (the hat indicates the model or

* Based on data obtained from the ESO Science Archive Facility under request number ZECHMEISTER73978. Based on observations made with the HARPS instrument on the ESO 3.6 m telescope under programme ID 074.D-0380.
best estimate, while without the hat it indicates observations with noise: \( S(x, y) = \hat{S}(x, y) + \sigma(x, y) \). Furthermore, the spectrum is recorded and binned by detector pixels. Hence it is convenient to grid (serves as model for “perfect” extraction).

Optimal extraction uses the column numbers as extractionective point spread function (ePSF, Anderson & King 2010) penalised functions in spatial direction employs spatial subpixel grid

\[ \hat{S}_{x,y} = \sum_i \Psi_{x,y,i} \delta_{i,x} \]  

where \( x \) and \( y \) now correspond to pixel indices, and a finite integration over wavelength \( \lambda \) still has to be done. The calibration matrix \( \Psi_{x,y,i} \) tabulates the effective response function of the spectrograph and detector. In Bolton & Schlegel (2010), Eq. \( 2 \) serves as “perfect” extraction.

Optimal extraction uses the column numbers as extraction grid \((\lambda \rightarrow x)\) and basically assumes 1D slit functions; i.e., any input wavelength \( \lambda \) that corresponds to the pixel \( x \) is imaged only to that column, meaning only pixels in the spatial direction \( y \) are affected, but no neighbouring columns. Under these circumstances the calibration matrix can be separated as

\[ \Psi_{x,y,i} = \delta_{i,x} \psi_{i,y} \]  

where \( \delta_{i,x} \) is the Kronecker delta\(^1\) and \( \psi_{i,y} \) the wavelength-dependent cross-section. Then the model image in Eq. \( 2 \) simplifies to

\[ \hat{S}_{x,y} = \psi_{i,y} s_{x,y} \]  

The response \( \psi_{i,y} \) could be measured directly with a uniform input spectrum \( s_{x,y} = 1 \), but it is much more challenging to determine \( \Psi_{x,y,i} \). In practice, exposures of flat lamps \( F_{x,y} \) are usually taken as part of regular calibration sets. Those flat exposures have high signal-to-noise ratios (S/Ns), and it can be further increased in master-flats by coadding many flat exposures. This means that the errors are negligible compared to science exposures and that we can set \( F_{x,y} = F_{x,y} \). The spectrum of a flat lamp \( f_{x} \) is generally not uniform overall, but continuous and featureless, and it varies slowly varying with wavelength. If \( f_{x} \) is known, we measure the response as \( \psi_{i,y} = \frac{F_{x,y}}{f_{x}} \) directly. However, \( f_{x} \) is usually not known in advance and actually it cannot be measured from the flat exposure alone. It should be characterised externally, either in advance by another flux-calibrated instrument or afterwards by observations of standard stars (with known spectra) with the same instrument as part of the flux calibration step. Since both might not be available, we extract the science spectrum \( s_{x} \) relative to the flat spectrum \( f_{x} \) and write for the model

\[ \hat{S}_{x,y} = F_{x,y} s_{x,y} \]  

\(^1\) \( \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases} \)

Table 1. Optimal extraction algorithms.

| Reference                  | Cross-section model                      | Comment                                  |
|----------------------------|------------------------------------------|------------------------------------------|
| Hewett et al. (1985)       | average along dispersion                 | assumes no order tilt                    |
| Horne (1986); Robertson (1986) | polynomials along dispersion             | assumes small order tilt                 |
| Urry & Reichert (1988)     | Gaussian function                        |                                          |
| Marsh (1989)               | coupled polynomials along dispersion     | employs spatial subpixel grid            |
| Piskunov & Valenti (2002)  | penalised functions in spatial direction | employs spatial subpixel grid            |
| this work                  | (master flat)                            | requires stabilised spectrograph         |

We obtain the spectrum \( \hat{S}_{x,y} \) by minimising the residuals between observations \( S_{x,y} \) and model \( \hat{S}_{x,y} \), i.e., solving the linear least-square problem

\[ \chi^2 \approx \sum_{x,y} w_{x,y} \left[ S_{x,y} - \frac{s_{x,y}}{f_{x}} \right]^2 = \text{minimum} \]  

where \( w_{x,y} \) are the weights for each pixel using a noise model \( \sigma_{x,y} \) (e.g. photon and readout noise; see Sect. 3). These weights may also include a map \( M_{x,y} \), masking bad pixels and pixels outside the extraction aperture \((w_{x,y} = \frac{M_{x,y}}{\sigma_{x,y}^2})\). Setting the derivative of \( \chi^2 \) with respect to \( \frac{s_{x,y}}{f_{x}} \) equal to zero, we get a set of decoupled equations with the solution at each position

\[ r_x \equiv \frac{s_{x,y}}{f_{x}} = \frac{\sum_y w_{x,y} F_{x,y} \hat{S}_{x,y}}{\sum_y w_{x,y} F_{x,y}^2} \]  

where \( r_x \) is the best fitting amplitude at each spectral bin \( x \). Figure illustrates the principle of flat-relative optimal extraction (FOX).

Equation \( 7 \) is quite similar to the well known optimal extraction equation (e.g. Horne 1986). The difference is that we extract the spectrum \( s_{x} \) relative to the flat spectrum \( f_{x} \), and \( F_{x,y} \) is not normalised and includes the natural spatial profile and all flat-field effects.

3. Noise model and error estimation

The pixels on the detector (and therefore the extracted values \( s_{x}/f_{x} \)) are affected by noise \( \sigma_{x,y} \). This consists mainly of photon noise \( \sigma_{ph}^2 = \hat{S}_{x,y} \) (if both \( \sigma_{ph} \) and \( \hat{S}_{x,y} \) are in units of photon counts\(^2\) and detector noise (e.g. the readout noise \( \sigma_{rn} \) of a CCD). Therefore, a simple noise model for each pixel is, for instance (now all units in ADU),

\[ \sigma_{x,y}^2 = \sigma_{m}^2 + g \hat{S}_{x,y} \]  

where constant gain \( g \) (which converts the number of photoelectrons to ADU\(^3\)) and constant readout noise \( \sigma_{m}^2 \) are assumed.

\( ^2 \) If a sky- or stray light background was subtracted from the image before, its contribution to photon noise should be also taken into account.

\( ^3 \) There are different definitions of the gain in the literature. Here we use \( \hat{S}[\text{ADU}] = g[\text{ADU}/e^-] \cdot \hat{S}[e^-] \). The photon noise in units of ADU is \( \sigma_{ph}[\text{ADU}] = g[\text{ADU}/e^-] \cdot \sigma_{ph}[e^-] = g[\text{ADU}/e^-] \cdot \sqrt{\hat{S}[e^-]} = \sqrt{g[\text{ADU}/e^-] \cdot \hat{S}[\text{ADU}]} \).

Article number, page 2 of [8]
\[ S_{x,y} - \hat{S}_{x,y} > \kappa \sqrt{\text{var}(S_{x,y} - \hat{S}_{x,y})} = \kappa \cdot \left( \sigma_{x,y}^2 - F_{x,y}^2 \sigma_{x,y}^2 / f_{x,y} \right), \]

(see Appendix A for a derivation of this equation), where the clipping value is typically \( \kappa \approx 3 - 5 \) to reject cosmics. A lower threshold could be set as well, e.g. to detected unmasked cold pixels. Moreover, other or additional criteria are also used (Baranne et al. 1996). The outliers are masked and the extraction process is repeated.

We note that the variance in the residual image \( S_{x,y} - \hat{S}_{x,y} \) is generally not just the pixel variance \( \sigma_{x,y}^2 \). In contrast to Horne (1986) the correction term \( F_{x,y}^2 \sigma_{x,y}^2 / f_{x,y} \) should be applied in Eq. (12). Figuratively, the observed dispersion in the residuals will be noticeably smaller than \( \sigma_{x,y}^2 \) because the number of pixels within the aperture is, as already mentioned, small and the profile is usually not uniform, but concentrates most flux in even fewer pixels (~3-4 pixels in fibre-fed spectographs). Imagine a concentrated profile, where one pixel has a very high weight. This pixel dominates the fit, thus forcing its own residual to zero. As an example, in Fig. 1 the pixel in the profile centre contains 400pix×13pix section of the reddest HARPS order (fibre A). Pixel-to-pixel variations, fringing, order tilt/curvature, and the spatial profile are present in both the master-flat \( F_{x,y} \) (top) and the stellar raw spectrum \( S_{x,y} \) which has additional (telluric) absorption lines (middle). The extracted spectrum \( s_{x,y} / f_{x,y} \) (bottom) is the scaling factor between both spectra for each column within the extraction mask (red lines).

Fig. 1. Principle of flat-relative optimal extraction (FOX). Each of the two upper panels shows a 400pix×13pix section of the reddest HARPS order (fibre A). Pixel-to-pixel variations, fringing, order tilt/curvature, and the spatial profile are present in both the master-flat \( F_{x,y} \) (top) and the stellar raw spectrum \( S_{x,y} \) which has additional (telluric) absorption lines (middle). The extracted spectrum \( s_{x,y} / f_{x,y} \) (bottom) is the scaling factor between both spectra for each column within the extraction mask (red lines).

With a given noise model, one can predict the extraction uncertainties through error propagation of Eq. (7)

\[ \text{var}(r_{x}) = \epsilon_{r_{x}} = \sum_{y} \left( \frac{\partial r_{x}}{\partial S_{x,y}} \right)^2 \sigma_{x,y}^2 = \sum_{y} \left( \frac{F_{x,y}/\sigma_{x,y}^2}{\sum_{y'} w_{x,y'} F_{x,y'}^2} \right)^2 \sigma_{x,y}^2 / \left( \sum_{y} F_{x,y}^2 / \sigma_{x,y}^2 \right)^2 \]  

(9)

\[ \chi^2 \text{ modelling are often rescaled by } \chi^2_{\text{red}} \text{ (the reduced } \chi^2) \text{ to provide a posteriori estimated errors} \]

\[ \sigma_{x,y} / f_{x,y} = \chi_{\text{red}} \cdot \epsilon_{x,y} / f_{x,y} \]  

(11)

where \( \chi_{\text{red}} = \sqrt{\frac{\chi^2}{N-\nu}} \) with \( \chi^2 \) from the global fit in Eq. (6). The number of degrees of freedom \( N - \nu \) in the denominator is given by the number of unmasked pixels \( N = \sum_{x,y} M_{x,y} \) within the extraction aperture and the number of fitted parameters \( \nu \) (extracted spectral values).

One can also consider another, second choice for the scaling factor, namely \( \chi_{\text{red},x} \) to be taken from the individual cross-section fits, i.e. \( \chi_{\text{red},x} = \chi_{\text{red}} / \nu \) where only one free parameter is left (\( \nu = 1 \)), \( \chi_{\text{red},x}^2 \) is the weighted sum of the residuals only in column \( x \) and \( N_x = \sum_y M_{x,y} \) is the number of unmasked pixels in that column. However, for fibre-fed spectrographs, the extraction aperture is only a few pixels (\( N_x \sim 10 \) wide). This gives low number statistics making \( \chi_{\text{red},x}^2 \) itself very uncertain (in contrast to \( \chi_{\text{red}}^2 \)). In particular, a considerable number of \( \chi_{\text{red},x}^2 \) will occur with values much less than one, when fitting the profiles along the dispersion axis. However, the predicted errors should not be smaller than the fundamental limit (e.g. photon noise).

As discussed in Horne (1986), cosmic ray hits can be efficiently detected and removed with optimal extraction. These cosmics distort the profile and can be identified as significant outliers by means of the noise model, e.g. by setting an upper threshold

\[ S_{x,y} - \hat{S}_{x,y} > \kappa \sqrt{\text{var}(S_{x,y} - \hat{S}_{x,y})} = \kappa \cdot (\sigma_{x,y}^2 - F_{x,y}^2 \sigma_{x,y}^2 / f_{x,y}) \]  

(12)

(see Appendix A for a derivation of this equation), where the clipping value is typically \( \kappa \approx 3 - 5 \) to reject cosmics. A lower threshold could be set as well, e.g. to detected unmasked cold pixels. Moreover, other or additional criteria are also used (Baranne et al. 1996). The outliers are masked and the extraction process is repeated.

4. Conceptual comparison with other optimal extraction implementations

Numerous optimal extraction algorithms (hereafter OXT) exists that are listed in Table 1. Usually, they have to assume a slowly varying spatial profile and differ in the reconstruction method for

\[ \chi^2_{\text{red},x} = \frac{1}{N_x} \sum_{x} \frac{1}{\nu} \chi^2_{x,y} \approx \frac{1}{\nu N_x} \sum_{x} \chi^2_{x,y} = \frac{1}{\nu N_x} \chi^2_{x,y} = \chi_{\text{red},x}^2 \]  

The approximation becomes an equality if the aperture width is the same for all columns, and no pixels are rejected (\( N_x = \text{const}, N_x \sim \nu (N_x) \)).
Table 2. Efficiency components of typical echelle spectrographs visible in flat-fields.

| Amplitude scale | Size scale | Source |
|-----------------|------------|--------|
| \( \varepsilon_{\text{pixel}} \) | 1 – 4% | pixel efficiency, pixel size, and quantum efficiency (CCD) |
| \( \varepsilon_{\text{fringe}} \) | 0 – 20% | fringing, interference pattern |
| \( \varepsilon_{\text{blaze}} \) | 0 – 100% | blaze function (echelle grating) |
| \( \varepsilon_{\text{d}} \) | 0 – 100% orders | wavelength-dependent efficiency of spectrograph and detector |

5 Therefore the blaze function and wavelength-dependent efficiencies are not corrected at this stage, and usually \( x_{\lambda} \cdot \varepsilon_{\text{blaze}} \cdot \varepsilon_{\lambda, x} \) is extracted (instead of \( x_{\lambda} \)).

6 http://www.eso.org/sci/facilities/lasilla/instruments/harps/
extract optimal extraction. Both the raw and reduced spectra are publicly available.\footnote{http://archive.eso.org/eso/eso_archive_main.html and http://archive.eso.org/wdb/wdb/eso/repro/form}

We extracted spectra for a standard star (HD 60753, B3IV) and a solar-like star (τ Cet, G8V). We used the REDUCE package of Piskunov & Valenti (2002) for the preprocessing: The bias frames were averaged to a master bias; five flat exposures were bias-subtracted, averaged to a master-flat and the scattered light was subtracted; one order-location frame (a flat lamp illuminating only fibre A) was used to define the order traces of fibre A; from the science images, the master bias and the scattered light were subtracted (no pre-flat-fielding was performed). Then we extracted the science spectra with our FOX algorithm with a spatial extraction width of ten pixels (five whole pixels below and five above the order trace), $\sigma_{\text{rms}} = 5$ ADU, and $g = 0.70$ ADU/$e^{-}$ ph. The wavelength solutions were taken from the DRS pipeline.

Figure 2 shows the result of the FOX extraction for two observations. As can be seen, the extracted orders match in the overlapping regions and seem to be ready to be merged directly. Order merging, however, is not the topic of this work and is not necessary for our RV computations below. Merging also needs some care, since the different resolution and sampling will lead to mismatches in the absorption and emission features, and in this way some information is also lost.

5.1 S/N measurement for the standard star

To measure the extraction quality, we use two observations of the standard star HD 60753, taken three hours apart in the night 2007-03-26; exposures times 9.5 min, and various methods of estimating S/Ns. Figure 3 displays the extracted aperture number 68 (6613–6687 Å) of the standard star (the prominent absorption line is HeI).

\[ 	ext{Fig. 2. Spectra of HD 60753 and } \tau \text{ Cet extracted with FOX. The orders (alternating colours) are not merged. The spectrum is not flux-corrected, but is relative to the spectrum of the flat lamp. The inset shows the overlap between two orders.} \]

\[ 	ext{Fig. 3. Comparison of the extraction quality for two HARPS spectra of the B3IV star HD 60753 taken in the same night (S/N~170). Top: Extraction with FOX. Second panel: Spectrum extracted with the HARPS DRS pipeline (not blaze corrected). Third panel: Ratio of both exposures for FOX (red dots) and DRS (blue dots). Bottom: Histogram of the ratio values. The black line is a Gaussian fit to the FOX histogram having a centre at 0.914 and a width of 0.00905.} \]
Using the uncertainties derived from the noise model (Eq. (9)), we estimate for FOX a quadratic mean signal-to-noise of $S/N = \sqrt{\sum \frac{\epsilon_i^2}{n}} = 169.6$ (exposure 2: 176.9) per extracted pixel around the central $n = 100$ pixels. Since no uncertainties are provided for DRS spectra, we assume pure photon noise ($\epsilon_i = \sqrt{\sigma_i}$). Then the count level in the DRS spectrum implies a mean photon $S/N$ of $\sqrt{\langle s_i \rangle} = 170.4$ (exposure 2: 178.6). These $S/N$ values are similar and provide fundamental limits. In this respect, we note that we have $\chi^2 = 1.08$ for this order.

Another way to measure the $S/N$ independently of a priori error estimates is to analyse the scatter in a continuum region. (This is the reason for choosing a standard star for the comparison.) For the same central region, the $S/N$ per extracted pixel derived as the ratio of the intensity mean and the standard deviation ($S/N = \langle \epsilon_i \rangle / \langle (\langle s_i \rangle - \langle s_i \rangle^2) \rangle^{1/2}$) is 185 (exposure 2: 163) for FOX and 177 (exposure 2: 141) for DRS. These numbers already indicate that the extraction qualities are similar.

In a further comparison (also independent of a priori error estimates), we take the ratio of the two spectra of the standard star and analyse the scatter. Taking the ratio of both spectra cancels out the remaining flat spectrum in the case of FOX and the blaze function in the case of DRS, therefore allowing for a more direct comparison over a wider range and even pixel-wise. We see in the third panel of Fig. 5 that the ratio values of FOX ($q,FOX = s_{\exp1}/s_{\exp2}$) and DRS ($q,DRS = s_{\exp1}/s_{\exp2}$) have a similar mean ($\mu_0 = 0.914$, constant over the full order) and are correlated. For both extractions, the scatter appears similar and barely distinguishable by eye. In both cases the scatter increases towards the order edges (since the flux decreases due to the blaze, see second panel of Fig. 3). As before, we measured a $S/N$ from the mean and the standard deviation of the ratio values for the central 100 pixels and find a slightly higher $S/N$ for FOX ($S/N_{q,FOX} = 133$ and $S/N_{q,DRS} = 129$).

In contrast to the previous method, we can extend this ratio method over the full order and also use regions with stellar lines (assuming that the stellar lines are static, i.e. do not vary in shape and position). A slight, relative shift between both spectra is present owing to the difference in their barycentric radial velocities (212 m/s, $\sim 0.25$ pix). The slightly increased scatter noticeable at the position of the strong stellar line is due to this shift as well as to the lower flux level (lower $S/N$).

The last panel of Fig. 3 shows a histogram for the 4096 ratio values. We see that the ratio values have nearly a Gaussian distribution, and we now measure the mean and dispersion more robustly from a Gaussian fit to the histograms. Using $S/N_q = 4$, we find values of 100.77 (FOX) and 100.87 (DRS) for order 68. The same procedure was applied to the other orders and the results plotted in Fig. 4. The overall course of the $S/N_q$ values mostly reflects the instrument efficiency (times the stellar energy distribution) along the order. Both FOX and DRS extraction deliver similar $S/N_q$ with quotients close to unity and deviation of $\sim 1\%$, whereas in this example FOX provided slightly higher $S/N$ in the blue orders.

5.2. RV performance for a Sun-like star

As a further, less direct, but probably more relevant proxy for the extraction quality we use radial velocity (RV) measurements. The RV precision depends on the $S/N$ in the observation and the RV information content of the stellar spectral lines (the number

8 Accounting for readout noise, say, 5 spatial pixels decreases the $S/N$ values by about $1 - \sqrt{\langle s_i + 5\sigma_i^2 \rangle / \langle s_i \rangle \approx 0.3\%}$.

Figure 4 shows the sharp features of a spectrum from a laser frequency comb (LFC; Murphy et al. 2007, Wilken et al. 2010) observed with HARPS (order 44). The residuals of the extraction ($S_{\exp} - S_{\exp}$) are shown in the lower panel, where $S_{\exp}$ is computed with the extracted spectrum $s_i$ and Eq. (5)). Of course, the residuals are larger in regions with larger flux (more photon noise). However, a systematic (not random) pattern remains: the resid-
Fig. 5. Radial velocities of τ Cet in the night of 2004-10-02 computed with the least square method for FOX and DRS extraction (top) and for comparison the radial velocities from cross-correlation for DRS extraction (bottom).

Fig. 6. Section of a HARPS laser frequency spectrum $S_{\lambda}$ (top, logarithmic intensity scale) and residuals $S_{\lambda} - \hat{S}_{\lambda}$ after the FOX extraction (bottom, linear scale). Coloured ticks indicate positions of cross-section cuts (see Fig. 4).

Fig. 7. Cross-sections at different positions in the laser frequency comb spectrum and flat-field image. The cut positions are at the LFC peak maximum (black asterisks), two pixels before (blue crosses), and two pixels after (red plus), as indicated by the corresponding colored ticks in Fig. 6; the coloured lines connect the corresponding profile means. The profiles are normalised to unit area.

The shortcoming of “optimal” extraction might be solved with “perfect” extraction that involves a 2D PSF as outlined in [Bolton & Schlegel (2010)]. It can theoretically also deal with stray light and ghost features. However, besides the increased computational effort, the main problem in practice is to obtain the calibration matrix $\Psi_{x,y}$. 

7. Conclusion

We have introduced a method extracting 1D spectra from 2D raw data using flat-field exposures as a measure of the instrumental profile. This method is similar to standard optimum extraction. It does not make any assumptions about the instrumental profile but requires its temporal stability between flat-field and science exposures. The method is well suited to stabilised fibre-fed spectrographs optimised for high-precision radial velocity work.

One of the main advantages of FOX is that the reconstruction of the instrumental PSF becomes unnecessary. Modelling of the PSF is a time-consuming step, and in regions of low signal, the PSF is often ill-defined in science exposures. It is therefore a strong advantage to take the PSF from well-defined flat-field exposures, which is the main idea of FOX. Following this scheme, extraction, masking, flat-fielding, and blaze correction are all carried out in one step without any need for data fitting.
When the calibration matrix is constructed. The decrease in required CPU time is significant, which is particularly relevant for large surveys like HARPS and CARMENES (Quirrenbach et al. 2012). Furthermore, FOX has no requirements concerning the spectral format (such as slowly varying spatial profiles), relaxes the need for accurate localisation of spectral orders, and does not involve any numerical unstable operations like division by the flat-field.

We compared the performance of FOX to standard optimal extraction using HARPS data and the HARPS data reduction system. The results are very similar with insignificant differences in the S/N. The spectra of individual orders extracted with FOX match well in the overlap regions showing that the inherent blaze correction works. We computed the RV series from HARPS spectra extracted with FOX and DRS and find them indistinguishable in terms of their rms scatter. We conclude that FOX is a highly efficient and very robust method for extracting astronomical spectroscopic data observed with stabilised fibre-fed spectrographs. FOX cannot overcome the limitations caused by tilted PSFs, stray light, or ghost features, but it can significantly improve the robustness and time efficiency of existing and future data reduction procedures.

Acknowledgements. MZ acknowledges support by the European Research Council under the FP7 Starting Grant agreement number 279347. We thank the referee, N. Piskunov, for many useful comments. We are also thankful to C. Lo Curto for helpful discussions about the HARPS instrument and F. Bauer for reading the manuscript.

References
Anderson, J. & King, I. R. 2000, PASP, 112, 1360
Anglada-Escudé, G. & Butler, R. P. 2012, ApJS, 200, 15
Barrane, A., Queloz, D., Mayor, M., et al. 1996, A&A, 119, 373
Bolton, A. S. & Burles, S. 2007, New Journal of Physics, 9, 443
Bolton, A. S. & Schlegel, D. J. 2010, PASP, 122, 248
Bouchy, F., Pepe, F., & Queloz, D. 2001, A&A, 374, 733
Hearnshaw, J. 2009, Astronomical Spectrographs and their History (Cambridge University Press)
Hewitt, P. C., Irwin, M. J., Bunclark, P., et al. 1985, MNRAS, 213, 971
Horne, K. 1986, PASP, 98, 669
Marsh, T. R. 1989, PASP, 101, 1032
Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20
Murphy, M. T., Uden, T., HolwARTH, R., et al. 2007, MNRAS, 380, 839
Pepe, F., LoNIV, C., Ségransan, D., et al. 2011, A&A, 534, A58
Piskunov, N. E. & Valenti, J. A. 2002, A&A, 385, 1095
Quirrenbach, A., Amado, P. J., Seifert, W., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
Robertson, J. G. 1986, PASP, 98, 1220
Urry, M. & Reichert, G. 1988, NASA IUE Newsletter, 34, 96
Wilken, T., LoNIV, C., Manescau, A., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Appendix A: Appendix
To estimate for each pixel the residual variance,
\[
\text{var}(S_{x,y} - \hat{S}_{x,y}) = \text{var}(S_{x,y}) + \text{var}(\hat{S}_{x,y}) - 2\text{cov}(S_{x,y}, \hat{S}_{x,y}).
\]
(A.1)
we need the pixel variance, which is \(\text{var}(S_{x,y}) = \sigma_{x,y}^2 = \frac{1}{w_{x,y}}\), and an expression for the model \(\hat{S}_{x,y}\). Inserting Eqs. (7) and (9) into Eq. (5) gives
\[
\hat{S}_{x,y} = F_{x,y}r_x + F_{x,y} \sum \omega_{x,y} w_{x,y} F_{x,y} S_{x,y}. 
\]
(A.2)

We see that \(\sigma_{x,y}\) itself contributes to \(\hat{S}_{x,y}\) and a covariance term will persist. Assuming that the pixels are otherwise independent, the variance of the residuals is given by
\[
\text{var}(S_{x,y} - \hat{S}_{x,y}) = \sigma_{x,y}^2 + F_{x,y} \sum w_{x,y} \sigma_{x,y}^2 - 2F_{x,y} \sigma_{x,y}^2, 
\]
(A.3)
This equation can be simplified for three special cases.

Case 1. For pixels with relatively low profile weights \((F_{x,y} \sigma_{x,y}^2 \ll \sigma_{x,y}^2)\), e.g. in the wings of spatial profiles, the residual variance just becomes \(\sigma_{x,y}^2\).

Case 2. When the pixel variance is the same for all pixels \(\left(\sigma_{x,y} = \sigma_0\right)\), e.g. readout noise dominates, we have \((\sigma_{x,y}^2 = \sigma_0^2 \sum w_{x,y})\).

Case 3. Assuming \(\sigma_{x,y}^2 = g \hat{S}_{x,y} = g r_x F_{x,y}\), i.e. photon noise dominates, we find \((\sigma_{x,y}^2 = \frac{\sigma_0^2}{w_{x,y}} F_{x,y})\).

Moreover, for a uniform profile \((F_{x,y} = 1)\) the factor in the bracket becomes \(\frac{\sigma_0^2}{F_{x,y}}\) (a well known correction factor for the unbiased variance).

Again for a uniform profile the prefactor becomes \(\frac{\sigma_0^2}{F_{x,y}}\).