A new method for instrumental profile reconstruction of high-resolution spectrographs

D. Milaković1,2 and P. Jethwa3

1 Institute for Fundamental Physics of the Universe, Via Beirut, 2, 34151 Trieste, Italy
2 INAF, Osservatorio Astronomico di Trieste, via Tiepolo 11, 34131, Trieste, Italy
e-mail: dinko@milakovic.net
3 University of Vienna, Department of Astrophysics, Türkenschanzstraße 17, A-1180 Vienna, Austria

Received ; accepted

ABSTRACT

Context. Knowledge of the spectrograph’s instrumental profile (IP) provides important information needed for wavelength calibration and for use in scientific analyses.

Aims. This work develops new methods for IP reconstruction in high-resolution spectrographs equipped with astronomical laser frequency comb (astrocomb) calibration systems and assesses the impact that assumptions on the IP shape have on achieving accurate spectroscopic measurements.

Methods. Astrocombs produce \( \approx 10000 \) bright, unresolved emission lines with known wavelengths, making them excellent probes of the IP. New methods based on Gaussian process regression were developed to extract detailed information on the IP shape from these data. Applying them to HARPS, an extremely stable spectrograph installed on the ESO 3.6m telescope, we reconstructed its IP at 512 locations of the detector, covering 60% of the total detector area.

Results. We found that the HARPS IP is asymmetric and that it varies smoothly across the detector. Empirical IP models provide a wavelength accuracy better than 10 m/s\(^{-1}\) (5 m/s\(^{-1}\)) with a 92% (64%) probability. In comparison, reaching the same accuracy has a probability of only 29% (8%) when a Gaussian IP shape is assumed. Furthermore, the Gaussian assumption is associated with intra-order and inter-order distortions in the HARPS wavelength scale as large as 60 m/s\(^{-1}\). The spatial distribution of these distortions suggests they may be related to spectrograph optics and therefore may generally appear in cross-dispersed echelle spectrographs when Gaussian IPs are used. Empirical IP models are provided as supplementary material in machine readable format. We also provide a method to correct the distortions in astrocomb calibrations made under the Gaussian IP assumption.

Conclusions. Methods presented here can be applied to other instruments equipped with astrocombs, such as ESPRESSO, but also ANDES and G-CLEF in the future. The empirical IPs are crucial for obtaining objective and unbiased measurements of fundamental constants from high-resolution spectra, as well as measurements of the redshift drift, isotopic abundances, and other science cases.

Key words. Instrumentation: spectrographs – methods: data analysis – techniques: spectroscopic

1. Introduction

Properties of the instrumental profile (IP) are important in several aspects when making accurate spectroscopic measurements.

Accurately measuring centres of spectral features is known to crucially depend on assumptions made on the IP shape, as inaccurate assumptions unavoidably bias centre measurements. When used to measure the centres of features used for instrument calibration, this inevitably leads to distortions in the wavelength calibration of the instrument. Avoiding distortions is important for fine structure constant measurements (Rahmani et al. 2013; Molaro et al. 2013; Evans et al. 2014; Whitmore & Murphy 2015; Dumont & Webb 2017) and other science cases. Furthermore, accurate IP shape models are also needed whenever theoretical models must be compared against the data during spectral modelling.

Most spectographs have an approximately Gaussian IP, with strong departures from this shape generally caused by optical design choices or (rarely) defects in the spectrograph construction. The IP shape has received comparatively little attention so far, for several reasons. Firstly, before the advent of extremely precise spectrographs, spectrograph performance was limited by systematic effects other than the IP shape. Secondly, in radial velocity studies, the exact assumptions of IP shape should be of little consequence as long as the shape is applied consistently, especially when comparing multi-epoch observations in which spectral features appear at approximately the same locations on the detector. Lastly, the ideal way for IP shape reconstruction is by observing a set of unresolved, uniformly distributed, and unblended spectral features covering a large fraction of the instrument’s spectral range and such sources just recently became available.

The IP for several instruments was previously reconstructed from spectra in which an absorption cell containing molecular iodine (I\(_2\)) was inserted in the light path while observing astronomical objects (Marcy & Butler 1992; Valenti et al. 1995; Kambe et al. 2002). The I\(_2\) cell methods have the benefit of using the light following the same optical path as the scientific observations so the reconstructed IP is exactly the one that appears in the data. However, the strong absorption from I\(_2\) makes it unfeasible to use when observing faint targets. Another drawback
comes from the fact that the lines in the I$_2$ spectrum are blended even when observed at a spectral resolution of \( \sim 10^5 \), complicating IP reconstruction. A comprehensive description on the use of the I$_2$ cell for IP reconstruction can be found in Butler et al. (1996).

A different approach is offered by the astronomical laser frequency comb (LFC) systems or ‘astrocombs’. Astrocomb lines have almost all necessary properties to make them suitable tools for IP reconstruction, as we explain in Sec. 2.1. Several other groups have recently used an astrocomb for this purpose Zhao et al. (2019), Hirano et al. (2020).

The paper is organised as follows. Sec. 2 presents the theoretical considerations of our work. New methods presented in Sec. 3 are based on Gaussian process (GP) regression and iterative procedures, allow for the reconstruction of the most likely non-parametric function describing the IP. These methods were applied to reconstruct the IP of HARPS using its astrocomb system in 2010. Improvements from knowing the empirical IP are presented in Sec. 4. The main results are summarised in Sec. 5 followed by a discussion in Sec. 6. We conclude with a discussion of future prospects in Sec. 8.

## 2. The instrumental profile of a spectrograph

### 2.1. Astrocomb as a tool for IP reconstruction

The LFC (Udem et al. 2002, Hänsch 2006) produces a train of emission lines (also known as modes) with frequencies determined by the equation

\[
f_n = f_0 + nf_r,
\]

where \( f_0 \) is the frequency of the 0th mode and \( f_0 \) and \( f_r \) are the offset and repetition frequencies. Both \( f_0 \) and \( f_r \) are actively controlled frequencies in the MHz range, stabilised against a radio-frequency standard such as an atomic clock (see, e.g., Probst 2015, for more details). As such, the LFC line frequencies are known with a relative accuracy of \( \sim 10^{-12} \) as long as a frequency lock with the reference is maintained. The accuracy and stability of the frequency standards can be transferred to the spectrograph, making LFCs the ideal wavelength calibration for astronomical spectrographs.

LFCs specifically adapted for astronomical spectrograph calibration are named astrocombs and have \( f_0 \) and \( f_r \) in the GHz range to match the spectrograph resolving power. First astrocomb prototypes were tested on the German Vacuum Tower Telescope (Steinmetz et al. 2008) and the HARPS (Wilken et al. 2010) spectrographs. The left panel of Fig. 1 shows a small part of the astrocomb spectrum observed by HARPS.

In the context of IP reconstruction considered here, astrocombs show another useful property: the comb lines are intrinsically extremely narrow, 300 kHz to 500 kHz (Tilo Steinmetz, private communication). This corresponds to a relative uncertainty of \( 5 \times 10^{-10} \) to \( 1 \times 10^{-9} \) (or \( \leq 1 \) m s$^{-1}$, when scaled by the speed of light). For comparison, the resolution element width of a spectrograph with spectral resolving power \( R = \lambda/\delta \lambda = f/\delta f = 120,000 \) (where \( \delta \lambda \) and \( \delta f \) are the full width at half maximum, FWHM, of the resolution element in wavelength and frequency space, respectively) at 550 nm is \( \delta f \approx 4.54 \) GHz. This is 10,000 times broader than the intrinsic astrocomb line width. The astrocomb lines are therefore completely unresolved by even the highest resolution astronomical spectrographs and, as such, individual modes can be considered monochromatic impulse inputs to the optical system. Their image on the detector is therefore a direct digitised representation of the 2D IP. Correspondingly, the extracted astrocomb spectrum then represents the 1D IP. For example, the middle panel of Fig. 1 shows a stack of flux normalised astrocomb lines observed by HARPS on top of their best-fit Gaussian centres.

Lastly, the astrocomb lines are high signal-to-noise (S/N), as seen in the right panel of Fig. 1. They also have very uniform flux levels: the dynamic range of the HARPS astrocomb peak amplitudes is \( \sim 4 \) dB. We are therefore able to use a single astrocomb calibration frame to reconstruct the IP without needing to worry about its variation with recorded light intensity within a single frame. In fact, exposures taken with different flux levels can be used to determine the relation between light intensity and IP shape in the future.

## 2.2. The digitised IP

This work was heavily influenced by the paper by Anderson & King (2000), henceforth referred to as A&K. While they were...
interested in performing accurate astrometric measurements and hence the 2-dimensional IP and we are interested in accurate wavelength measurements and hence the 1-dimensional IP, most of the considerations made by the authors apply here unchanged.

Similarly to Valenti et al. (1995) previously, A&K considered the fact that a pixelated detector does not directly record information on the ‘true’ instrumental profile (their ‘iPSF’). Instead, what is recorded is the integrated product of the iPSF with the 2-dimensional pixel sensitivity profile (R), where the integration limits are determined by the pixel’s boundaries (see their section 3 and formulae (1)-(5), specifically). This led them to define the ‘effective’ IP (their ‘ePSF’) as the convolution of the iPSF with R. The digitised image of a point source on the detector then provides samples of ePSF, with the values in individual pixels determined precisely by the distance of the centre of the pixel from the point source, multiplied by some flux factor intrinsic to the source.

Importantly, A&K showed that one does not need to determine the iPSF to measure the positions of features with sub-pixel accuracy. Instead, determining ePSF is sufficient because the recorded pixel values – the only observable – are already integrated over the pixel area with appropriate weights determined by the pixel response function. Going further, A&K demonstrated how repeated observations of the same point sources, each offset from the other by ≤ 1 pix, can be used to reconstruct ePSF on a subpixel scale and its variation across the detector.

Their approach has three major advantages, as noted by the authors themselves. Firstly, it avoids integrating the true IP over the pixel area when fitting for the position of the feature (the centre of a star image in their case, and the centre of an astrocomb line for us), as the integration over the area of the pixel is already included into the definition of the effective IP. Secondly, fewer computational resources are needed to solve for the effective IP than the true IP itself. Finally, because of the definition of the ePSF as the convolution of iPSF and R, the latter is automatically incorporated into the modelling process and needs not to be untangled from the recorded point source image.

We agreed with the reasoning of A&K above, and have applied it to reconstruct the 1-dimensional IP. All references to the IP in the rest of the paper therefore refer to the effective IP, unless stated otherwise.

2.3. The 1-dimensional effective IP

The flux in an extracted spectral pixel is given by:

\[ F_i = f_i \int_{-\infty}^{+\infty} \psi_E(x-x_i) Q(x-i) \, dx + B_i, \]  

where the index \( i \) identifies the pixel centred at \( x = i \) (in a single echelle order). The values \( f_i \) and \( x_i \) are the total integrated flux (brightness) and the position of a monochromatic light source, respectively, and \( B_i \) is the background level. \( \psi_E(\Delta x) \) is the 1-dimensional true IP, that is it quantifies the fraction of light (per 1d pixel unit length) that falls onto the point offset by \( \Delta x \) from \( x_i \). Equation (2) is analogous to A&K’s equation (1) and our \( \psi_i \) is a 1d analogue of their iPSF.

Q replaces A&K’s R in our equations. In addition to the 2d pixel response function contained in the latter, Q carries the information on all other operations applied to the data by the data acquisition and spectral extraction processes. This includes bias subtraction, flat-fielding, weighted summation of pixels in the spatial direction, etc. Untangling contributions of each of these operations in 1d spectra is impossible in practice, and doing so would require working on 2d raw data. However, like with R, it is not necessary to know Q’s value.

Applying the same transformations as in A&K, appropriately modified for 1d, we obtain:

\[ F_i = f_i \psi_E(i-x_i) + B_i, \]  

where:

\[ \psi_E(\Delta x) = \int_{-\infty}^{+\infty} Q'(x) \psi_E(\Delta x - x) \, dx, \]  

and \( Q'(x) = Q(-x) \). Equation (3) is the 1d analogue of A&K’s ePSF.

In above Equations, \( x \) stands for detector pixels in uncalibrated data but the same reasoning and Equations would be valid if \( x \) were replaced by \( \lambda \) in wavelength calibrated data. In fact, we do this later, deriving two sets of IP shapes. The first set was derived from non-calibrated data and used to determine the centres of astrocomb lines in pixels. These centres were used to wavelength calibrate the instrument, allowing for a second set of IP shapes to be produced from the same data using \( \lambda \). The latter were used to investigate the accuracy of the wavelength calibration in Sec. 5.

2.4. GP: non-parametric functional form for the IP

Previous studies modelled the non-Gaussian shapes using different modifications to the Gaussian, including modifying the wings (Cardelli et al. 1990) or adding satellite Gaussian profiles (Valenti et al. 1995; Zhao et al. 2019). More recently, a Gaussian with a modified exponent in combination with cubic splines to modify that shape when necessary (Hao et al. 2020), A&K used splines to describe the full IP shape. While we tried out some of these approaches, we were unable to obtain satisfying results. Ultimately, we decided to use GP regression to obtain a smooth function that is most likely representation of the HARPS instrumental profile. This is similar to what Hirano et al. (2020) have done for the InfraRed Doppler spectrograph on the Subaru telescope.

GP regression is a technique for probabilistic, non-parametric regression (Rasmussen & Williams 2006). Here we provide a basic overview; see Aigrain & Foreman-Mackey (2017) for a detailed review in an astronomical context. Given pairs of data-points \((x_i, y_i)\) with noise \( \epsilon \), the goal of regression is to find a function \( f \) such that \( y_i = f(x_i) + \epsilon_i \). A GP is a prior over this space of functions. The defining property of a GP is that for any finite sample of points \( x \), the sampled function values \( f = \{ f(x) \} \) have a multivariate normal probability distribution

\[ p(f) = \mathcal{N}(m, K), \]  

which is parameterised by a mean vector \( m \) and covariance matrix \( K \). These are, in turn, specified by a mean function \( m \) and covariance function \( K \), where

\[ m_i = m(x_i; \theta), \]  

\[ K_{ij} = K(x_i, x_j; \phi), \]  

which depend on hyperparameters \( \theta \) and \( \phi \) respectively. Assuming that the noise \( \epsilon \) is normally distributed with zero mean and scale \( \sigma_r \), i.e.

\[ p(\epsilon) = \mathcal{N}(0, \sigma_r^2), \]  

Article number, page 3 of 28
then the likelihood of the data $y$ as a function of the hyperparameters is given by
\begin{equation}
\mathcal{L}(\theta, \psi) = N(y; \mathbf{m}, \mathbf{K} + \text{diag}(\sigma^2)).
\end{equation}

Given observed data $(x_i, y_i)$, training the GP means to find the best combination of hyperparameters $(\theta, \psi)$ which maximise this likelihood. Once trained, the GP encodes a probability distribution over suitable regression functions $f$. Ensembles of functions $f$ can be sampled from the trained GP (see Agrinier & Foreman-Mackey 2023 for details), and the variance amongst these functions encodes the uncertainty in the regression. The specific choices of mean functions and covariance functions that we adopt in this work are introduced in later sections: in section 3.2 we describe a GP model for the line profile, and in section 4.6 a GP to describe the observational errors.

### 3. Methods

The instrumental profile can be estimated from the astrocomb spectra by inverting Eq. (1):
\begin{equation}
\bar{\psi}(\Delta x) = \frac{F_i - B_i}{f_i},
\end{equation}
where we have now dropped the subscript “$E$” on $\psi_E$ for the ease of notation. The hat above $\psi$ indicates estimation from observations. In the equation above, $\Delta x = i - x$, is the distance between the pixel and astrocomb line centres and $f_i$ is the integrated astrocomb line flux. Knowing $x_i$ and $f_i$, allows for determining at which offsets $\Delta x$ has the pixel centered at $x = i$ sampled $\psi$ and how the sampling has been scaled. The summation on top of each other provides subpixel sampling of the pixel and astrocomb line centres and $f_i$ allows for determining at which offset $\Delta x$ has the pixel centered at $x = i$ sampled $\psi$ and how the sampling has been scaled. The observables on top of each other provide subpixel sampling of the pixel and astrocomb line centres and $f_i$ allows for determining at which offset $\Delta x$ has the pixel centered at $x = i$ sampled $\psi$ and how the sampling has been scaled.

When performing the GP regression, we considered also the variance on $\psi$. It is, however, not simple to calculate this quantity analytically because it means calculating the variance of the ratio of two random Poisson variables (the first one being $f_i - B_i$, and the second one being $\psi_i$, one of which is conditional on the other, that is $f_i = \sum (F_i - B_i)$. Instead, we used the following simple approximation that preserves S/N in the data:
\begin{equation}
\sigma^2_{\psi} = \frac{1}{f_c} \left[ \sigma^2_{F_c} + \sigma^2_{B_c} \right].
\end{equation}

Simple MCMC calculations of the variance on Eq. (10) confirm that this approximation works very well, with between values determined through Eq. (11) and the MCMC variance differing by 5% at most, always in lowest flux pixels. We later describe how we used the data themselves to modify $\sigma^2_{\psi}$ where necessary.

#### 3.1. Obtaining line position and brightness

Accurate determination of $x$, and $f$, that go into calculating $\bar{\psi}$ requires that $\psi$ is already known and an accurate estimation of $\psi$, in turn, requires $x$, and $f$, to also be known. This motivated the use of an iterative approach.

Initial guesses for $x_i$ and $f_i$ were determined by approximating $\psi$ with a Gaussian profile and fitting it to the data with three free parameters: the amplitude $A$, the mean $\mu$, and the standard deviation $\sigma$ of the Gaussian, that is
\begin{equation}
I(x; A, \mu, \sigma) = \frac{A}{\sqrt{2\pi}\sigma} \exp\left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right].
\end{equation}

Since we are interested in the effective IP, integration under pixel area when optimising for the Gaussian parameters was not performed. Instead, we simply evaluated the Gaussian profile at pixel centres. The impact this has on estimating $x_i$ is minimal (the average change in line centres is approximately $1 \times 10^{-3}$ pix). We identify $x_i$ and $f_i$ with the mean and the integrated area under the Gaussian profile, that is $x_i = \mu$ and $f_i = A \sqrt{2\pi}$. In subsequent iterations, once $\psi$ has been modelled using our GP regression methods, the Gaussian approximation is replaced by the formula:
\begin{equation}
I(x; A_i, \mu_i, \sigma_i) = A_i \psi((\omega(x - x_i))\). \tag{13}
\end{equation}

Here, $\psi$ is the IP normalised to unit area, and $A_i$, $\mu_i$, and $\sigma_i$ are free parameters. $\psi(\Delta x)$ models were saved into a look-up table containing 800 values spanning $\Delta x = \pm 8$ pix, corresponding to 50 values per pixel, and were interpolated using cubic splines during fitting. $A_i$ and $x_i$ are the line amplitude and central position, respectively, whereas $\sigma_i$ is a parameter whose value is $\approx 1$, allowing for small stretching or compressing of the pixel scale when performing the fit. Because $\psi$ already describes integrated quantities, line brightness was calculated as $f_i = \sum N_i f_i$, where the sum goes over the $N$ pixels of the line.

#### 3.1.1. Fitting an IP model to the data

Let $I(\xi)$ be the predicted model flux, where $\xi$ is the set of all free model parameters. Parameter optimisation finds $\xi$ which minimises the normalised residuals of the model to the data:
\begin{equation}
\xi = \arg\min f(\xi), \tag{14}
\end{equation}
where
\begin{equation}
f(\xi) = \sum_{j=0}^{r_N} \left( \frac{I(\xi) - F_j}{\sigma_j} \right)^2. \tag{15}
\end{equation}

The summation in Eq. (15) goes over the $N$ pixels of the astrocomb line, where the $j = 0$ and $N$ correspond to the locations of the minima either side of the line and $w$ is the weight applied to each pixel.

The weighting scheme was implemented to ensure the best possible fit at astrocomb line centres, following the discussion and recommendations made in A&A. The weight of a pixel centred at $x = j$ and at a distance $\Delta x = j - x$, from the astrocomb line, was given by
\begin{equation}
w_j = \begin{cases} 
1 & \text{for } |\Delta x| \leq x_1 \\
1 - |\Delta x| & \text{for } x_1 < |\Delta x| \leq x_2 \\
0 & \text{for } |\Delta x| > x_2,
\end{cases} \tag{16}
\end{equation}
where $x_1 = 2.5$ pix and $x_2 = 5$ pix. The weighting scheme was implemented when fitting both our empirical IP shapes and the Gaussian IP shapes.

The reported best-fit parameters correspond to the values of $\xi$ that satisfy Eq. (14). Parameter errors were calculated from the diagonal of the covariance matrix at the best-fit solution.
3.2. A GP for the instrumental profile

The empirical $\psi$ was modelled using GP regression, conditioning the GP on the observed $\hat{\psi}(\Delta x)$ samples. Because the shapes of astrocomb lines are approximately Gaussian, the mean function was chosen to be

$$m(x; \theta) = N(x; \mu, \tau) + y_0$$

(17)

where $N$ represents the standard normal probability density function. This function takes parameters $\theta = (A, \mu, \tau, y_0)$ corresponding to the normalisation, the mean, the standard deviation, and the zero-offset, in that order.

The choice of covariance function affects the smoothness of the regression curve. The squared exponential covariance function was chosen, i.e.

$$k_{SE}(x_i, x_j; \phi) = \sigma_0^2 \exp\left(-\frac{|x_i - x_j|^2}{2l^2}\right),$$

(18)

which depends on the hyperparameters $\phi = (a, l)$ i.e. an amplitude and a length-scale. The amplitude controls how far the regression curve can stray from the mean function, while the length-scale determines typical length of wiggles in the curve.

An additional diagonal term was added to the covariance matrix in order to account for the uncertainties on the data:

$$k(x_i, x_j; \phi) = k_{SE}(x_i, x_j; \phi) + \sigma_\epsilon \delta_{ij},$$

(19)

where $\sigma_\epsilon = \sqrt{\sigma_0^2 + \sigma_\delta^2}$ and $\sigma_\delta^2$ is an additional error term that is common to all $\psi$ samples and is also a free hyperparameter.

Finally, $\delta_{ij}$ is Kronecker delta.

3.3. Defining the IP centre

Defining the location of the IP centre, that is deciding where to set $\Delta x = 0$, is crucial for line centre $x$, measurements and, by extension, the construction of the IP models themselves. For an unimodal and symmetric IP profile, any of the mean, mode, or median can be meaningfully used as the centre. While these quantities are also well defined for asymmetric profiles, they can be very different from each other and also be sensitive to small changes in the shape of the profile.

We wanted our centre estimator to be robust against small changes in hyperparameter values. We thus set to identify the optimal centre estimator for use in this work by comparing the properties the mean, the mode, the median, and the centre definition from [Anderson & King (2000)]. In this last case, the IP centre is posited to be at the location which results in equal fluxes for the two brightest pixels: $\psi(\Delta x = 0.5) = \psi(\Delta x = -0.5)$.

To find the most robust estimator, a single IP model was perturbed 1000 times and the four centre estimators were calculated for each perturbed model. The details of the comparison can be found in Appendix A. The results show that A&K’s centre estimate is the most robust of the four examined, shifting by $\lesssim 5 \times 10^{-8}$ pix across the 1000 perturbed profiles. For comparison, the mode was the least stable estimator with shifts $\lesssim 1 \times 10^{-3}$ pix.

We also followed A&K’s suggestion to apply a rigid shift to the $\psi$ samples in $\Delta x$ to ensure that the IP models are properly centred before modelling. The transformation applied was:

$$\Delta x \rightarrow \Delta x + \frac{\psi(0.5) - \psi(-0.5)}{\Delta x_{0.5}(-0.5) + \Delta x(0.5)}.$$  

(20)

Subsequently to the data shifting, GP hyperparameters were optimised once more and the GP was retrained on the shifted data.

The data shifting and hyperparameter optimisation procedures were repeated until the difference between consecutive shifts was smaller than $1 \times 10^{-3}$ pix or for a maximum of twenty times (the stopping criterion was typically satisfied after three to four repetitions). This $\psi$ at the end of this procedure was the final model of a single iteration.

4. Reconstructing the HARPS IP

4.1. Data

The methods described in the previous Sections were applied to the High Accuracy Radial-velocity Planet Searcher (HARPS) spectrograph, and its IP was reconstructed from a single astrocomb calibration frame taken as a part of an observing programme to measure the fine structure constant ($\alpha$) at $z = 1.15$ towards the quasar HE0515–4414 executed in December 2018 (ESO observing programme 0102.A-0697, PI Milaković). The primary reason for choosing HARPS is that the IP models could be applied to the science spectra in order to explore their impact on measurements of $\alpha$. The HARPS IP was previously investigated by Zhao et al. (2021), although no IP shape models were produced.

While ESPRESSO observations of the same quasar are also available and astrocomb calibrated, those observations required significant manual interventions in the data (Murphy et al., 2022) for reasons not yet fully understood (ESPRESSO consortium, private communication). Additionally, the wavelength calibration of ESPRESSO contains systematic effects which are also not yet understood. The most obvious example is the high frequency correlation of its wavelength calibration residuals with wavelength. The residuals also are not normally distributed nor do they have properties expected from photon noise statistics (Schmidt et al., 2021). These problems have not been observed in HARPS.

4.1.1. HARPS

HARPS is a fibre-fed echelle spectrograph designed for extremely precise spectroscopic measurements (Mayor et al., 2003), installed at the ESO 3.6m telescope at the La Silla Observatory. HARPS’s primary scattering element is an R4 grism which disperses the incoming light into 72 echelle orders, 89 through 161. Optical fibres transport the light from the telescope to the instrument in two channels. One channel (the primary, fibre A) generally carries the astronomical target light. The other channel (the second, fibre B) can be used to simultaneously record the sky light or the light from a wavelength calibration lamp. HARPS is equipped with three calibration lamps: a Thorium-Argon (ThAr) arc lamp, a Fabry-Pérot etalon, and an astrocomb, which can all be observed in either fibre. The nominal spectral resolution is $R = 115,000$. The entire instrument is enclosed in a temperature and pressure controlled vacuum vessel in order to minimise environmental disturbances and their impact on the scientific observations.

4.1.2. Astrocomb data description

We first set out to use the astrocomb spectrum recorded on 2018-12-05 at 08:12:52.040 UT. That spectrum was used for wavelength calibration of quasar observations in Milaković et al. (2021) and was thus the obvious choice. Upon closer inspection of that data, we discovered that the peaks of some astrocomb
Flux ($\times 10^3$)

![Flux graph]

**Fig. 2.** Astrocomb spectrum (black) and the inferred spectral background (red) for 512 pix wide sections of four echelle orders.

The relevant blaze function is saved in the observations taken on ground (red) for 512 pix wide sections of four echelle orders. The spectral background contains contributions from scattered light (~ 3%, Rodier & Lo Curto 2019) and from background associated with the astrocomb system itself. The astrocomb background is seeded by the high-power amplifier of the system required to broaden the astrocomb spectrum from the infrared to the optical wavelengths inside the photonic crystal fibre (Probst et al. 2015, 2020), and contributes as much as 30% of the total flux below 500 nm (Milaković et al. 2020).

The full spectral background was determined in each echelle order independently. We started by identifying the minima as the locations at which the first derivative of the spectral flux array is closest to zero and, simultaneously, second derivative of the flux array is positive. Before calculating the derivatives, the flux array was upsampled by a factor of 10 and smoothed using a Nuttall window function (Nuttall 1981) to more precisely determine the location of the minima and to avoid falsely detecting minima in the noise. The most appropriate width for the window function was determined by calculating the power spectrum of the flux array and identifying the frequency carrying the most power. In practice, the window widths were between 11 and 15 pixels, depending on the wavelength. The background over the entire order was then determined by connecting the minima between the lines using straight lines and smoothing the resulting array with a Nuttall window function with a window size of 51 pix. Small sections of the background is shown in Fig. 2 for four echelle orders. We assumed that the background follows Poisson statistics, that is $\sigma_i^2 = B_i$.

**4.3. Line detection and line properties**

Boundaries between individual astrocomb lines were identified by applying the minima detection algorithm to the background-subtracted spectra. 10576 astrocomb lines were detected, with mode numbers between 24087 and 33384 (1472 modes falling in the spectral region covered by two adjacent orders are seen twice). The separation between lines ranges between 10 pix at 500 nm and 18 pix at 690 nm, whereas the FWHM of the best fitting Gaussian for each line was <3.7 pix everywhere. Lines were therefore sufficiently separated such that line blending is expected to be negligible. This provided > 100 000 samples of $\phi(x)$ across a 60% of the detector area.

Visual examination of lines revealed asymmetry in their shapes, with most lines being skewed to the right, and this asymmetry also appeared to vary across the detector. We confirmed this by examining differences between two line centre estimates: the centroid ($x$, a flux-weighted average position of the line) and the centre of the best fitting Gaussian, $x_G$. If the lines are approximately symmetric, there should be zero difference between these two centre estimates whereas positive or negative values imply that the line is skewed to the right or to the left, respectively.

The main panel of Fig. 3 shows values of this quantity for all detected lines as coloured squares. Lines appearing in a single order are aligned along the $x$-axis and echelle orders are separated along the $y$-axis, so that the location of the squares in the panel approximately corresponds to their location on the detector. Wavelength within each order increases with increasing pixel number (left to right), whereas the central wavelength of the order increases (order number decreases) from bottom to top.

Correlated departures from a symmetrical shape across the detector are immediately noticeable. The most obvious is the division of the detector area into two halves: left, with positive $x - x_G$ values (coloured red) and right, with values around zero (white) or negative (light blue). Lines appearing in the middle of each order have values closest to zero. Maximal positive departures from the symmetrical shape are seen in the top left corner of the panel, with $\overline{x} - x_G$ as large as 0.14 pix, and maximal negative departures are seen in the top right corner of the panel, with values of approximately −0.03 pix. In between the two regions the values vary quickly. Negative values seem to appear around 3000 pix in most orders, but in some orders $\overline{x} - x_G$ recovers to zero at the right edge. The width of regions with negative values seem to increase as one moves vertically from the middle of the detector towards the top and the bottom. Contours in the same Figure illustrate this more clearly. We conclude from this that the HARPS IP is not symmetric and that its shape varies strongly across the detector in a correlated way.

The top panel of Fig. 3 shows the histogram of the values from the main panel. The mode of the distribution is around zero pixels, and its median is 0.020 pix. The central 68% of the dis...
Fig. 3. Astrocomb line shape asymmetry as a function of position on the HARPS detector. Main panel: Each coloured square shows one of the 10576 detected astrocomb lines in the HARPS astrocomb spectrum described in the text. The position of the square in the (x,y) plane approximately matches its position on the HARPS detector frame. The gap seen at ≈ 525 nm is due one echelle order falling between the two HARPS detector CCD chips so is not recorded. Wavelength within a single echelle order increases towards the right. The colour of the square is proportional to the difference (in units pixels) between the measured line centroid and the line position determined by fitting a Gaussian IP to the data. The zero point of the colour bar to the right is set to zero, such that the blue and the red colours denote negative and positive differences, respectively. Smooth variations across the detector most likely correspond to IP shape variations. Contours over smoothed data are plotted to visualise trend more clearly and emphasise where certain values cluster. Contour levels correspond to the 16th, 50th, and 68th percentile of the values, i.e., −0.002 pix, 0.020 pix, and 0.083 pix. Full lines are used for positive levels and dashed lines for negative. The cyan rectangle indicates the location of the 1st segment in optical order 90, for which we show an empirical IP model in Fig. 4. Top panel: Histogram of the values shown in the main panel. The secondary x-axis on the top shows the line centre differences in units m s$^{-1}$ (1 pix = 820 m s$^{-1}$). The vertical dashed red lines show the positions of the distribution median and the central 68% distribution limits. The same quantities are shown as thick black lines in the colour bar.

### 4.4. Determining the IP locally

To reduce the impact from positional IP variability on our re-constructed models as much as possible, the detector was artificially segmented in both the main and the cross dispersion directions and the IP models was then determined for each segment independently. Echelle orders provide a natural division in the cross dispersion direction and we further divided individual orders into 16 segments, each spanning 256 pix or ≈ 0.5 nm, in the main dispersion direction. The 256 pix width was chosen to avoid a known “stitching” issue of the HARPS detector which appears as small changes in the detector pixel size every 512 pix in the main dispersion direction and arises from imperfections during CCD construction (Wilken et al. 2010; Bauer et al. 2015; Coffinet et al. 2019; Milaković et al. 2020). The area covered by one such segment is shown as a cyan rectangle in the top left corner in the main panel of Fig. 3.

While we explicitly assumed that IP does not change within the segment when performing the modelling, we allowed for smooth variation of $\psi$ when fitting IP models to the data, as follows. Let the sets $\{\psi_1, \ldots, \psi_{16}\}$ and $\{X_1, \ldots, X_{16}\}$ denote the 16 IP models calculated within single echelle order and the central segment positions, respectively. We accepted the models as correct at those locations, $\psi_{i+X_1} \equiv \psi_i$, where $i$ is the ordinal number of the segment and $x$ is now a coordinate within the echelle order and not $\Delta x$. The IP at an arbitrary $x$ was then interpolated from the two closest $\psi$ models by applying weights inversely proportional to the distance between $x$ and the two nearest segment centres:

$$\psi(\Delta x)|_{X_{i-1}\leq x < X_{i+1}} = \frac{d_2^2}{D} \psi(\Delta x)|_{X_{i-1}} + \frac{d_1^2}{D} \psi(\Delta x)|_{X_{i+1}}. \quad (21)$$

Here, $d_1 = |x - X_i|$, $d_2 = |x - X_{i+1}|$, and $D = |X_{i+1} - X_i|$. If $x < X_1$ or $x > X_{16}$, no interpolation was done and only the single closest IP was used ($\psi_1$ or $\psi_{16}$).

### 4.5. Hyperparameter determination

To model the IP for one segment, we optimise $\mathcal{L}$ from Eq. (9) using $\theta$ samples falling within the segment boundaries to find the maximum likelihood values of the hyperparameters $\theta$ and $\phi$. The mathematical framework for hyperparameter optimisation was implemented using the GP library tinygp (Foreman-Mackey 2023), which was performed using the limited memory Broyden & Fletcher - Goldfarb - Shanno algorithm with box constraints (also known as L-BFGS-B, Byrd et al. 1995) as implemented in the Python optimisation libraryjaxopt (Blondel et al. 2021). After this step, known as training, we can draw sample functions $f$ from the GP evaluated at arbitrary locations, and the variation of these samples give us an estimate of the uncertainty in the regression function $f$. The mean $f$ is the IP model, $\psi$.

During training, parameters $(A, \mu, \tau, \nu_0)$ were initialised at the values determined from fitting Eq. (17) to $\psi(\Delta x)$ using non-linear least-squares. These parameters were constrained not to move further than ±5 standard deviations, which were calculated from the covariance matrix diagonal at the best fit solution. We required hyperparameters $\alpha, l$, and $\sigma_0$ to always be positive so they were expressed as logarithms. Their initial values were $\log \alpha = 1$, $\log l = 0$, and $\log \sigma_0 = -5$. These parameters were constrained to have the following values: $-4 \leq \log \alpha \leq 4$, $-1 \leq \log l \leq 2$, and $-15 \leq \log \sigma_0 \leq 1.5$. Only $\sigma_0$ ever reaches its imposed boundary (on the lower side).

We also explored a fully Bayesian treatment, whereby we introduce hyperprior on $\theta$ and $\phi$ and perform Markov Chain Monte Carlo (MCMC) sampling of the posterior distribution, thus propagating additional uncertainty to the regression function. We found MCMC sampling prohibitively slow for general use (several hours per segment using MCMC compared to one
minute per segment using L-BFGS-B). However, our tests suggest that our maximum likelihood uncertainty estimates are very similar to the fully Bayesian approach and the IP models derived through L-BFGS-B are indistinguishable from the MCMC one. More details are provided in Appendix B.

Fig. 4. GP fits to the IP. Top panel: The large black points are the 251 \( \psi \) samples with corresponding error bars in the first segment of order 90 (average wavelength \( \lambda = 676.5 \) nm), also indicated as a rectangle in Fig. 3. The thick red line is our reconstruction of \( \psi \) using GP regression (Sec. 2.4) after re-centring, in the first iteration. The secondary GP for variance estimation was not used, resulting in a \( \chi^2 \) of 5.01 for this fit. The shaded red bands show the 1σ ranges of the model. The dashed blue line is the mean function of the GP, \( m(x; \theta) \), Eq. (17). The small black points above the model show the residuals normalised with respect to the error on the data, with the grey shaded region indicating ±1σ. Parameters \( \theta \) (controlling \( m \)) and \( \phi \) (controlling the correlation matrix \( K_{ij} \), Eq. (18)) were fitted simultaneously. The magenta dotted line is the difference between the GP model and the mean function, that is the function controlled by \( K_{ij} \). Departures from the Gaussian shape are observed almost everywhere. The thin dotted black line indicates zero flux. Rectangles indicate regions plotted in the bottom panels. Panels with numbers (1)-(4); Zoom-ins of the top panel. The number in the top left corner identifies one rectangle in the top panel.

More details are provided in Appendix B.

Fig. 5. GP for empirical variance estimation. \( g(\Delta x) \) multiplies \( \sigma_S \) to capture the observational variances more correctly. Black squares with error bars are the logarithms of \( S^2 \), where \( S^2 \) is calculated as explained in the text. The solid black line is the mean prediction of \( g(\Delta x) \) from GP regression and the grey shaded are the 1σ ranges. The y-axis label on the right shows \( S^2 \) and \( g(\Delta x) \) on a linear scale.

4.6. Empirical variance estimation

We already commented that our \( \sigma_\psi \) estimates from Eq (11) are formally incorrect. Fig. 4 shows the model derived from \( \psi \) samples in the 1st segment of order 90 (wavelength range 676.1 nm \( \leq \lambda \leq 676.7 \) nm) in detail. While the fit is generally good, the residuals scatter beyond what is expected from errors on individual data points. This is most obviously seen in the peak region, where the residuals show correlated scatter outside of the expected ±1σ range indicated by the grey shaded area in the top of the Figure’s main panel. A detailed view, shown in panels (1)-(4) of the same Figure, reveals that the scatter between nearby \( \psi \) samples is inconsistent with the scatter expected from their estimated errors and that this inconsistency changes with \( \Delta x \). For example, panel (3) shows that the errors in the region close to the profile peak are underestimated.

The increased scatter also affects the goodness of fit statistic, reduced \( \chi^2 : \chi^2 = \chi^2 / \nu \), where \( \nu \) is the number of degrees of freedom in the fit. \( \chi^2 \) is the weighted sum of squared deviations from the expected value

\[
\chi^2 = \sum_{i=1}^{N} \left( \frac{O_i - E_i}{\sigma_i} \right)^2 ,
\]

where \( O_i \) and \( E_i \) are the observed and the expected values of the \( i \)th observation, and \( \sigma_i \) is its standard error. In this case, \( O_i = \psi_i \), \( E_i = \psi(\Delta x_i) \), \( \sigma_i = \sigma_S \), and the summation goes over the \( N \) samples of a single segment. For the model shown in Fig. 4, \( \chi^2 = 5.01 \), significantly higher than the desirable value of \( \approx 1 \).

Models for other segments had similar \( \chi^2 \) values, most likely due to incorrect errors on \( \psi \). Even though we also obtained models with \( \chi^2 \approx 1 \), strong departures of the normalised residuals from the ±1σ range, such as those seen in Fig. 4, were observed also in those models.

More details are provided in Appendix B.
4.6.1. A secondary GP for variances

Since these models underestimated the observational variance, the resulting uncertainty estimates on line positions were also likely to be underestimated. We therefore implemented additional methods to deal with this issue by extending our modelling procedures to more correctly capture the observed variance. A second GP was introduced to modify $\sigma_i$ from Eq. (19), of the form

$$g(\Delta x) \sim N(C_g, G_g).$$

(23)

where $C_g$ is a constant hyperparameter and $G$ is a squared exponential kernel with the amplitude $a_g$ and the length-scale $l_g$ as hyperparameters. The mean prediction of $g(\Delta x)$ acts as a multiplicative factor modifying the observed variance of $\hat{\psi}_i$ such that

$$\sigma_i^2 = \frac{1}{\hat{\psi}_i^2} \left( g(\sigma_i^2) + \sigma_0^2 \right).$$

(24)

where $\sigma_0^2$ is the unmodified error on $\hat{\psi}_i$ and $g_i = g(\Delta x_i)$.

The training data for $g$ was derived from the normalised model residuals, $(\hat{\psi}_i - \psi_i) / \sigma_i$, as follows. In the first step, the $N$ residuals were divided into equally spaced bins in $\Delta x$ such that there were at least 15 points in every bin. Sample variance, $S^2$, was then calculated together with the variance on $S^2$ for each bin:

$$\text{Var}(S^2) = \frac{1}{n} \left( \mu_i - n - 3 \right) \left( n - 1 \right)^{-1}.$$

(25)

Here, $n$ is the number of residuals in the bin and $\mu_i$ is their fourth central moment.

Hyperparameters $\phi_g = \{ C_g, a_g, l_g \}$ were then optimised using equations in Sec. 2.4 on pairs of points $(x_b, \log(S^2))$, where $x_b$ are bin centres. Limits on $\phi_g$ were: $-10 \leq \log C_g \leq 5$, $-3 \leq \log a_g \leq 3$, and $-1 \leq \log l_g \leq 3$. Logarithms were used to ensure $g(x)$ and hyperparameters $\phi_g$ are always positive. The value $\log(\text{Var}(S^2))$ was added to the diagonal of $G$ to account for the uncertainty on the variances analogously to Eq. (19). An example of $g(\Delta x)$ determined in this way is shown in Fig. 5 using the data and the model previously shown in Fig. 4. This determination of $g(\Delta x)$ is also in excellent agreement with MCMC determinations (Appendix B).

Finally, hyperparameters of the GP describing the instrumental profile were recalculated using the empirical variances. Appendix C shows a detailed plot for the same data as in Fig. 4 obtained in this way. The multiplication function applied to the variances on $\hat{\psi}$ is the same as shown in Fig. 5. The improvement from applying this method is obvious: the residuals have improved significantly and show no dependence on $\Delta x$ and the value of $\chi^2$ is much closer to unity. Similar improvements are seen for all segments.

4.7. The final IP models

We produced ten iterations of IP models for each of the 16 segments of the 32 echelle orders illuminated by the astrocomb (5120 models in total). In each iteration, all astrocomb lines were refitted using the $\psi$ models from the preceding iteration, where the IP shape was also interpolated to the location of the astrocomb line, as explained in Sec. 4.4. Surprisingly, we found that the likelihood of the model does not necessarily increase with each subsequent iteration. Furthermore, models with lower likelihoods were also associated with worse fits to the data and larger uncertainties on model parameters. We thus also separately saved a set of 512 “most likely” $\psi$ models from the set of 5120 by choosing the model with the largest likelihood from the set of 10 models produced for one segment. This set of models also results in the average $\chi^2$ (calculated over all astrocomb lines) being the closest to unity.

IP models in pixel space for all 32 orders are shown in Fig. 6 where one order is shown per panel. Looking at any of the panels, the strong variation of the IP shape within the order is immediately noticeable, with the bluest line (corresponding to the segment covering the bluest wavelength range) generally being the most sharply peaked and the reddest line (corresponding to the reddest wavelength covered by that order) being more broad. The IPs in the reddest echelle orders exhibit a tail on their right side that is the strongest for blue lines and that almost completely disappears for the reddest line, changing smoothly with wavelength. Inter-order variations are best appreciated by comparing the bluest lines in each panel, revealing that the aforementioned tail grows with increasing wavelength.

The most likely values for all ten hyperparameters are shown as a function of detector position in Fig. 7 together with their histograms (above each panel). Most notable is that most (if not all) hyperparameters change smoothly across the detector area, albeit the degree of smoothness varies between them. We also noted similarities between spatial distributions of some hyperparameters, such as between $\mu$ and $\nu_0$, and $a$ and $l$. The distribution of $\mu$ is consistent with what is seen in Fig. 3 where the lines in top left corner of the detector showed the strongest signs of asymmetry in their shapes and lines in the right detector edge were more symmetric, albeit not fully. Similar behaviour is seen in all echelle orders, once more consistent with the general picture on line shapes provided by Fig. 3.

5. Improving wavelength calibration accuracy

5.1. Fit quality of astrocomb lines and uncertainties on line position measurements

Fig. 8 shows the models for one randomly chosen astrocomb line fitted with three different IPs: (1) the empirical IP at the line location (i.e. $\psi$ is linearly interpolated from the nearest two segment centres), left panel; (2) the empirical IP from the nearest segment centre (no $\psi$ interpolation was done), middle panel; and (3) a Gaussian IP, right panel. The best fit quality, in terms of $\chi^2$ is provided by the interpolated empirical IP; followed by the empirical IP without interpolation, and finally by the Gaussian IP. The non-interpolated empirical IP is shown here for comparison and is not used elsewhere. We note the large difference in $\chi^2$ between the Gaussian IP model fit and either one of the two empirical IP model fits, demonstrating again that the Gaussian IP is inadequate. Similar results are obtained for all other examined lines.

All 10576 astrocomb lines were independently fitted using our empirical IP shapes and a Gaussian IP shape, yielding two sets of model parameters, including line centres (see Eq. (12) and (13)). In the following text, we use subscripts to indicate which model parameters are referred to (“E” for our empirical IP models and “G” for Gaussian IP models). The $\chi^2_E$ for all lines fitted using the final empirical IP models are shown in Fig. 9 as a function of line position on the detector. No data weighting was applied in calculating $\chi^2_E$.

Looking at the spatial distribution of values (main panel), we see that the overall fit quality is excellent although lines closer
Fig. 6. HARPS IP models in pixel space for orders 89 through 122. Each panel shows the IP models for a single echelle order, whose central wavelength is printed above it. The 16 coloured lines inside each panel show how the IP changes with wavelength (line colour changes from blue to red with increasing wavelength). The same plots with background reference grids (which make it easier to see the profile changes) can be found on the arXiv version of this manuscript [https://doi.org/10.48550/arXiv.2311.05240].
Fig. 7. Values of hyperparameters describing $\psi$ as a function of detector position, also plotted as histograms above the main panels. White colour in the main panels corresponds to the median hyperparameter value. To better visualise trends, the colour bar and histogram limits only show the central 95% of the plotted values. The units on the colour bar are indicated above the panels. The vertical dashed red lines showing the median and the central 68% limits (calculated over all values) which are also shown as thick black lines in the colour bar.

to order edges have systematically higher $\chi^2_v$ values than lines in the middle of the detector. This indicates issues in those regions relating either to our models or to the observed fluxes and their uncertainties. Inaccuracies in the blaze correction function would adversely impact both on the construction of $\psi$ models (by increasing scatter between neighbouring $\psi$ samples) and on the fitting procedure (by providing incorrect spectral fluxes and errors). Upon inspection, we found significant correlation between the blaze correction function and larger $\chi^2_v$ values.

As seen from the histogram in the Figure’s top panel, the most commonly observed $\chi^2_v$ values is in the range $0.9 < \chi^2_v < 1.1$ and the median $\chi^2_v = 1.41$. The central 68% of the distribution is within $0.73 < \chi^2_v < 2.67$, with 70% (95%) of the values $\chi^2_v < 1.96$ (4.19). For comparison, the median $\chi^2_v$ for the Gaussian fits is 155.91, and 70% (95%) of the lines have $\chi^2_v < 186.22$ (270.30). Overall, the fit quality is significantly better when using our $\psi$ models than when using a Gaussian IP shape approximation. The improvement is also appreciated by comparing our Fig. 8 and 9 to figures 4 and 5 in (2020).

Fig. 10 shows the uncertainties on line positions, $\sigma_{x,E}$, determined from the covariance matrix at the best-fit solution. As expected, larger uncertainties are associated with lines with higher $\chi^2_v$ values (e.g. at detector edges) and in regions with low S/N in the data. The median uncertainty is 2.2 m pix and the central 68% distribution limits are 1.4 m pix and 3.4 m pix. 95% of all lines have $\sigma_{x,E} < 4.9$ m pix. Approximating $\sigma_{x,E}$ in units m s$^{-1}$ (1 pix = 820 m s$^{-1}$), distribution median corresponds to 1.79 m s$^{-1}$ and the 16th, 68th, and 95th percentiles correspond to 1.17 m s$^{-1}$, 2.76 m s$^{-1}$, and 4.01 m s$^{-1}$, respectively.

5.2. Differences between Gaussian and empirical IP centres and the impact on wavelength calibration

First we compared $x_0$ measured by fitting empirical IP shapes to $x_0$ determined by fitting Gaussian IP shapes. Fig. 11 shows $\Delta x_0 = x_{0,G} - x_{0,E}$ for all 10576 lines, revealing striking results. Firstly, zero differences are seldom observed. The maximum observed $\Delta x_0 = 0.193$ pix (corresponding to a velocity shift of 158 m s$^{-1}$), appears at the blue wavelength edges of the reddest order (red area in the top left corner). Negative values are only ever observed in the very low S/N part of the spectrum (light blue area at the bottom centre). The 16th, 50th, and 84th percentiles are 0.020 pix (16.6 m s$^{-1}$), 0.039 pix (31.8 m s$^{-1}$), and 0.083 pix (67.7 m s$^{-1}$), in that order. 50% of the lines have $\Delta x_0 < 0.039$ pix (31.8 m s$^{-1}$) and 95% of them have $\Delta x_0 < 0.116$ pix (94.9 m s$^{-1}$).

Secondly, the differences vary smoothly both within and across orders. For all orders, the maximum difference occurs at their blue wavelength edges (left in the main panel). Moving towards the red wavelength edge, $\Delta x_0$ approaches zero approximately in the middle of the order (coloured white) before increasing again at its red edge (turning more brightly red). The contours show this more clearly and also reveals that this pattern closely resembles the pattern for $\mu$ in Fig. 7.

Next, the two sets of $x_0$ measurements were used to wavelength calibrate HARPS. In producing the calibration, a single echelle order was divided into eight 512 pixel wide segments and a seventh order polynomial was fitted in each segment individually. No continuity conditions were imposed, meaning that the wavelength calibration is allowed to jump at each segment boundary. This approach was shown to be preferable over using a single continuous polynomial covering the entire order, even when very high order polynomials are used and detector stitch-
Fig. 8. Results of fitting the same astrocob line using three different IPs. They are, in order from left to right, (1) the most likely empirical IP, allowed to vary locally within the echelle order (interpolated), (2) when the empirical IP from the nearest segment is used instead (not interpolated), and (3) the Gaussian IP. In each top panel, the solid black histogram with error bars shows the astrocob line flux and the solid red line with crosses shows the best fit model of that data. The line model is described by Eq. (13) when fitting the empirical IP and by Eq. (12) when fitting a Gaussian IP. The location of the best-fit line centre, $x_\text{fit}$, is indicated by the vertical dotted black line in each panel, with the ticks and labels on the top indicating $\Delta x$. The two vertical grey shaded bands illustrate the limits of the weighting scheme, Eq. (10), $\chi^2$ for each fit is printed in the top right corner of the panel. In calculating $\chi^2$, no data weighting was applied. Finally, the continuous dashed grey line shows the IP model, $\psi$, shifted and rescaled to match the data (for visual comparison purposes only). Dots in each bottom panel are the normalised fit residuals and the grey shaded band indicates the ±1σ region. Dot opacity is proportional to the pixel weight in the fit (Eq. (16)), which is also printed above each dot. Dots falling within the darker grey band in the top panel all had a weight of unity, whereas the dots falling outside of both bands had zero weight (so are not plotted). Dots falling in between, that is inside the lighter grey band of the top panel, had weights that change linearly between zero and unity. We note that the y-axis scale differs between the bottom panels.

The two wavelength calibrations were next compared to each other on a pixel-by-pixel basis. Expressing the differences in pixel wavelengths as velocity shifts, we found the same general pattern already seen in Fig. 11 so this is not shown. The only difference with respect to the pattern in Fig. 11 is that the pattern seen in wavelength calibration comparison is of the opposite sign. This can be understood by considering that, generally, $x_{\text{E},E} < x_{\text{E},G}$ so that the same wavelength (determined from Eq. (11)) is associated with a smaller pixel number when $x_{\text{E},E}$ is used. Once these centres are used for wavelength calibration, the same pixel is assigned a lower wavelength in the empirical IP case than in the Gaussian IP case, resulting in the velocity shift between the two calibrations having an opposite sign with respect to the sign of $\Delta x$. The median velocity shift between the two calibrations is $-27.74\text{ m s}^{-1}$ (Gaussian minus empirical IP calibration), with the central 68% of the values falling between $-76.04\text{ m s}^{-1}$ and $-16.06\text{ m s}^{-1}$. The maximum negative difference is approximately $-150\text{ m s}^{-1}$, appearing at the same locations where the maximum positive $\Delta x$, is observed in Fig. 11. This is consistent in amplitude with the simple conversion based on the velocity content of HARPS pixels.

5.3. Differences between wavelength measurements made using Gaussian and empirical IP models

To quantify the difference of our empirical IP models make on wavelength measurements from the data, the astrocob spectrum was next treated as if it were a science spectrum, so it was calibrated using the wavelength calibration produced in Sec. 5.2 and line wavelengths, $\lambda_0$, were measured from this spectrum without any prior knowledge from Eq. (1).

For this purpose, we additionally created IP models in velocity space (shown in Appendix B). The data and the procedures used for creating these models were exactly the same as described in Sections 5 and 6 (including the secondary GP for variance estimation), except that $\Delta x$ was expressed in terms of velocity instead of pixel. The velocity of the $i^{th}$ pixel is:

$$v_i = \frac{\lambda_i - \lambda_0}{\lambda_e},$$

(26)

where $c$ is the speed of light, and $\lambda_0$ is the wavelength at pixel centre. Wavelengths for all pixels were obtained from the wavelength calibration, which was done independently in each iteration using the appropriate $x_0$ estimates. $\lambda_e$ is the wavelength
at the astrocomb line centre determined by fitting our velocity space IP to the data in exactly the same way as was previously done in pixel space (3.1), except for the weights in Eq. (16) changing to be $x_1 = 2 \, \text{km s}^{-1}$ and $x_2 = 4 \, \text{km s}^{-1}$. After a total of ten iterations, a set of most likely IP models in velocity space was saved separately and used here.

Finally, the wavelength of each line ($\lambda_*$) was measured from the wavelength calibrated data. Astrocomb lines were approximated by a Dirac $\delta$ function, for which the convolution with an IP model returns the IP model itself, so $\lambda_*$ was determined by fitting the appropriate IP shapes to the observations directly. Two sets of $\lambda_*$ measurements were produced: one by using our empirical IP shapes and the second by using Gaussians. To allow a more straightforward interpretation of the results, each kind of IP shape was used consistently to measure $x_\ast$ and $\lambda_\ast$. In both cases, the fitting followed the methods explained in Sec. 3.1.1 with $\lambda_\ast$ taking the place of $x_\ast$. Empirical IPs were fitted in velocity space because of the way they were constructed and because it was easier. Gaussian IPs were fitted in wavelength space instead. Doing so did not impact $\lambda_\ast$ measurements: the transformation from wavelength to velocity space results in the multiplication of the Gaussian function (Eq. 12) by $\exp(-(c/\lambda_\ast)^2)$, where $\lambda_\ast$ is the Gaussian centre and also defines zero velocity. This transformation can be written as a simple rescaling of Gaussian amplitude, width, and the centre remains unchanged. However, the transformation introduces correlations between the amplitude, the width and $\lambda_\ast$, so it was avoided.

Differences between the two sets of $\lambda_\ast$ measurements were expressed in velocity units:

$$\frac{\Delta \nu}{c} = \frac{\lambda_\ast - \lambda_{\ast,E}}{\lambda_{\ast,F}}.$$  

where subscripts G and E on $\lambda_\ast$, again indicate the IP used and $\lambda_{\ast,F}$ is the expected wavelength of the line determined through Eq. (1). Fig. 12 shows the results, with $\Delta \nu_*$ varying smoothly across the detector in a way similar to Fig. 11. The largest positive (but also absolute) departures from zero can be found in the blue wavelength edges of the reddest echelle orders, reaching up to 35 m s$^{-1}$. Distribution median is 2.74 m s$^{-1}$ and the central 68% of the values are within $-0.02 \, \text{m s}^{-1} \leq \Delta \nu_* \leq 13.32 \, \text{m s}^{-1}$, 16% of the lines have $\Delta \nu_*$ below zero. In absolute terms, the 50%, 70%, and 95% of the lines have $|\Delta \nu_*| \leq 2.56 \, \text{m s}^{-1}$, 8.76 m s$^{-1}$, and 19.37 m s$^{-1}$, respectively.

While Fig. 11 shows the differences between $x_\ast$ when two different IP shapes are assumed, Fig. 12 shows the difference between the end results of a chain of processes performed under two different assumptions for IP shape. These processes are, in order: measuring $x_\ast$, wavelength calibrating HARPS using $x_\ast$, constructing IP models in velocity space from wavelength calibrated data, and measuring $\lambda_\ast$ by fitting these models to the wavelength calibrated data. The fact that differences observed in Fig. 12 are smaller than those in Fig. 11 demonstrate that wavelength calibration diminishes the initial differences between $x_\ast$ seen in Fig. 11.

The uncertainties on $\lambda_\ast$, associated with empirical IPs, $\sigma_{\lambda_\ast,G}$, resemble those on $x_{\ast,E}$ both in the distribution on the detector and in their histogram (so are not shown). The median value of $\sigma_{\lambda_\ast,G}$ is 1.74 m s$^{-1}$, with the 16th, 68th, and 95th percentiles 1.13 m s$^{-1}$, 2.75 m s$^{-1}$, and 4.09 m s$^{-1}$ (respectively). These numbers are in excellent agreement with the uncertainties on $x_{\ast,E}$ expressed

---

**Fig. 9.** Reduced $\chi^2$ for 10576 fits of astrocomb line shapes using our empirical IP models. Main panel: $\chi^2$ as a function of position on the detector. The colour bar’s zero point corresponds to $\chi^2 = 1$, such that blue and red colours show lines whose $\chi^2$ is smaller and larger than 1 (respectively). Top panel: Histogram of $\chi^2$ values plotted in the main panel. The vertical dashed red lines show the locations of the median and the central 68% distribution limits. The same quantities are shown as thick black horizontal lines in the colour bar.

**Fig. 10.** Uncertainty on astrocomb line centres from fitting our empirical IP models to the data, $\sigma_{\lambda_\ast,E}$. Main panel: $\sigma_{\lambda_\ast}$ as a function of position on the detector. The colour bar zero point is set to the sample median, such that the red and the blue colours correspond to values larger and smaller than the median, respectively. Lines with larger uncertainties generally have larger $\chi^2$ (c.f. Fig. 9) or low S/N (at 500 nm). Top panel: Histogram of the values plotted in the main panel. The vertical dashed red lines show the median and the central 68% distribution limits, which are also shown as thick black horizontal lines in the colour bar. The x-axis on the top shows $\sigma_{\lambda_\ast}$ in units m s$^{-1}$ (1 pix = 820 m s$^{-1}$).
Fig. 11. Differences between astrocomb line position measured by fitting two different IP models to the data, $\Delta x_\ast = x_{G,E} - x_{E,E}$. The subscripts indicate which IP model was used to determine the centre during line fitting (G for the Gaussian and E for our empirical IP models). Main panel: $\Delta x_\ast$ as a function of position on the detector. The zero point of the colour bar is set to $\Delta x_\ast = 0$ pix, such that the red and blue colours correspond to positive and negative differences. Contours show the spatial distribution of values more clearly, with contour levels corresponding to the median and the central 68% distribution limits. Top panel: The histogram of the values plotted in the main panel. The bottom $x$-axis is in units pixel and the top axis in units m s$^{-1}$ (1 pix $\approx$ 820 m s$^{-1}$). The vertical dashed red lines show the median and the central 68% distribution limits. The same quantities are shown as horizontal thick black lines in the colour bar to the right of the main panel.

Fig. 12. Differences between wavelengths of astrocomb lines measured using our empirical IP models and using Gaussian IPs, expressed as a velocity shift between the two (Eq. [27]). Main panel: $\Delta v_\ast$, as a function of position on the detector. The zero point of the colour bar is set to $\Delta v_\ast = 0$ m s$^{-1}$, such that the red and blue colours correspond to positive and negative velocity differences. Black contours more clearly visualise the distribution of $\Delta v_\ast$, values across the detector. Contour limits correspond to the median and the central 68% distribution limits. A full line is used for positive contour levels and a dashed line for negative ones. Top panel: The histogram of the values plotted in the main panel. The vertical dashed red lines show the median and the central 68% region limits. The same quantities are shown as horizontal thick black lines in the colour bar to the right of the main panel.

5.4. Wavelength scale distortions

Another important question is whether choosing a particular IP produces correlated errors (distortions) in the wavelength scale of the instrument. Avoiding distortions is important whenever relative wavelength shifts between different spectral features are measured, as is done for fundamental constant studies and sometimes also for studies of isotopic abundances of chemical elements. The existence of wavelength scale distortions in UVES (Dekker et al. 2000), HIRES (Vogt et al. 1994), and HDS (Noguchi et al. 2002) spectrographs is well documented and their impact on fundamental constant measurements is understood (Rahmani et al. 2013 | Evans et al. 2014 | Whitmore & Murphy 2015 | Dumont & Webb 2017 | Milaković et al. 2020). In fact, removing wavelength scale distortions is one of the major expected benefits from using astrocombs for instrument wavelength calibration and a major reason for their deployment on current and future spectrographs. Yet, previous studies found some small level of distortions can remain even when astrocombs are used (Milaković et al. 2020 | Schmidt et al. 2021).

Errors on $\lambda_\ast$, measurements, i.e. $\epsilon_\lambda = \lambda_\ast - \lambda_{\ast,C}$, are shown in Fig. [13] for all 10576 lines, expressed in m s$^{-1}$. The empirical IP case (left) shows no evidence of wavelength scale distortions at any scale and no dependence on position on the detector. The larger scatter at in the order with a 500 nm central wavelength below 2048 pix is due to low S/N in data. The histogram of the errors (top panel, left) approximately follows a normal distribution centred at zero. After removing 78 lines (0.7% of the total) with errors falling outside of the $\pm 20$ m s$^{-1}$ range, the remaining 10498 error values have a mean of $-0.17$ m s$^{-1}$ and a standard deviation of 3.75 m s$^{-1}$. The latter value is comparable to the standard deviation of the wavelength calibration residuals (3.72 m s$^{-1}$) but is $\approx 60\%$ larger than the average $\sigma_{\ast,pn}$ (2.39 m s$^{-1}$) and the average $\sigma_{\ast,E}$ (2.30 m s$^{-1}$). The median error is $-0.21$ m s$^{-1}$, and the central 68% of the error values are in the range $-3.47$ m s$^{-1}$ to 3.18 m s$^{-1}$. 95% of the lines have absolute errors smaller than 7.54 m s$^{-1}$.
Fig. 13. Wavelength measurement error, $\epsilon_1$, as a function of detector position for our empirical IP models (left) and for the Gaussian IP approximation (right). Colour schemes in both panels are identical and centred at 0 m s$^{-1}$ (white) to highlight the vast improvements associated with the former. Contours for the empirical IP case (left) were not informative so are not shown. Plotted contour levels for the Gaussian IP case (right) correspond to the 16th and 50th value percentiles (~14.06 m s$^{-1}$ and ~3.59 m s$^{-1}$, respectively) while values above the 68th percentile (i.e. above 0.93 m s$^{-1}$) do not cluster at any specific location so that contour is not shown. Smaller panels show the histogram of values shown under them. The median and the central 68% distribution limits are shown as vertical dashed red lines. The same quantities are shown as thick black lines in the colour bar to the right of each main panel.

Fig. 14. Wavelength measurement error, $\epsilon_1$, as a function of wavelength for our empirical IP models (top) and for the Gaussian IP approximation (bottom). Grey dots are the same 10576 error values shown in Fig. 13 but now plotted as a function of $\lambda_{LFC}$. The data were binned such that each bin contains all astrocmb lines appearing in the same echelle order, after which the mean and standard deviation were calculated for each bin (black dots with error bars). In calculating the above quantities, we ignored values falling outside of ±20 m s$^{-1}$ (±50 m s$^{-1}$) for the empirical (Gaussian) IP cases. The red line in each panel is the best-fit straight line going through the binned values (see text for parameter values). The red hashed histogram and the red labels on the right spine of the Figure show the number of grey dots falling in each bin. Note that the bin edges are approximate because of the partial overlap between spectral ranges of adjacent orders. Thin cyan lines in the bottom panel are the best fit second order polynomials to the grey dots. Their parameters are tabulated in Tab. E.1.

The right side of Fig. 13 shows $\epsilon_1$ when the Gaussian IP approximation is used instead. In this case, the histogram of the errors (top panel) shows a strong tail extending towards ~50 m s$^{-1}$. 30 values (0.3% of the total) falling outside of ±50 m s$^{-1}$ range were removed from further consideration. The median of the remaining values is ~3.59 m s$^{-1}$ and the central 68% of the values fall within ~14.06 m s$^{-1}$ and 0.93 m s$^{-1}$. Positioning the errors on the detector (main panel), clear evidence of wavelength scale distortions is revealed. The distortions appear both in the main dispersion direction (horizontally) and in the cross-dispersion.
direction (vertically), with the detector divided into approximately equal halves vertically following the 50th percentile contour level. This division is consistent with the patterns already seen in Figs. 3 and 12. The largest errors generally appear at the blue order edges of orders and the largest absolute errors are in the reddest echelle orders (see the contour tracing the 16th percentile in the same panel).

We further examined the long-range and short-range distortions by plotting $\epsilon_i$ as a function of wavelength, shown as grey dots in Fig. 14. To better examine long-range trends, $\epsilon_i$ values were binned according to the echelle order that they appear in, after which the mean and the variance of the values in each bin were calculated (large black dots with error bars in the same Figure). In calculating the means and the variances, we removed values falling outside of $\pm 20 \text{ km s}^{-1}$ ($\pm 50 \text{ km s}^{-1}$) for the empirical IP (Gaussian IP) case. The red hashed histograms in the two panels of Fig. 14 show how many values remained after outlier removal. At the end, a straight line of the form $m(\lambda - \lambda_0) + b$ was fitted to the mean errors with inverse variances as weights, and the result is shown as a red line in each panel. For $\lambda_0 = 600 \text{ nm}$, the top panel of Fig. 14 has $m = (−0.04 \pm 0.02) \text{ m s}^{-1}$ per 100 nm and $b = (−0.17 \pm 0.01) \text{ m s}^{-1}$ per 100 nm, and $m = (−2.8 \pm 0.3) \text{ m s}^{-1}$ per 100 nm and $b = (−6.2 \pm 0.2) \text{ m s}^{-1}$ for the bottom panel.

We found no evidence for short-range distortions associated with the empirical IP case (grey dots in the top panel of Fig. 14). Concentrating on the Gaussian IPs case, that is on the grey dots in the bottom panel of Fig. 14 we observed strong inter-order distortions. In the worst case, the total distortion within one order can be as large as 60 m s$^{-1}$. The distortion within a single echelle order seemed to be well described by a second order polynomial (cyan lines going through the grey dots in the bottom panel). Their best fit coefficients are reported in Tab. 1.

5.5. Examination of astrocomb modes appearing twice on the detector

Some astrocomb modes appear twice on the detector due to the spectral range overlap between two adjacent orders, offering another powerful tool to examine the wavelength measurement accuracy and the existence of any biases. If the IP models used to fit astrocomb lines are correct and if the wavelength calibration derived from astrocomb centres is accurate, the measured $\Delta \lambda$ for the two observations of the same mode should be identical. We identified 1472 such modes in the spectrum and calculated the relative shift between the pair of wavelength measurements:

$$\frac{\Delta \nu_{\text{pair}}}{c} = \frac{\lambda_{\nu,2} - \lambda_{\nu,1}}{\lambda_{\text{F}C}}. \quad (28)$$

In the formula above, $\lambda_{\nu,1}$ and $\lambda_{\nu,2}$ are the two measured wavelengths for the same mode appearing at two locations of the detector, and $c$ is the speed of light.

The distribution of $\Delta \nu_{\text{pair}}$ associated with the empirical IP case is plotted as the red histogram in Fig. 15 and shows that there is no bias associated with the use of our empirical IP models. In comparison, using Gaussian IPs results in a strong bias (dashed black histogram in Fig. 15). The statistical properties of these distributions can be found in Tab. 1. Interpreted as probability distributions, the histograms can be used to quantify accuracy associated with the use of a specific IP shape. For the empirical IP case, the probability of achieving accuracy better than $20 \text{ m s}^{-1}$, $10 \text{ m s}^{-1}$, and $5 \text{ m s}^{-1}$ is 99.9%, 92.0%, and 64.1%. The same levels of accuracy for the Gaussian IP case have probabilities of 84.3%, 28.8%, and 8.4% (respectively).

Fig. 15. Histogram of measured velocity shifts (Eq. (28)) between 1472 astrocomb modes appearing twice in the spectrum. Line wavelengths were measured by fitting the most likely empirical IP model (full red histogram) or a Gaussian IP model (dashed black histogram) to the data in a wavelength calibrated spectrum. Zero velocity means that the same wavelength was measured for both lines of the pair.

Fig. 16 shows $\Delta \nu_{\text{pair}}$ as a function of wavelength, analogously to Fig. 14. Grey dots show individual values and black dots with error bars show the mean and the standard deviation in bins, where each bin contains all lines falling in the same echelle order. The empirical IP case shows no wavelength dependency at any scale (top panel). The best fitting straight line going through the binned $\Delta \nu_{\text{pair}}$ values in the top panel has a slope $m = (0.0 \pm 0.2) \text{ m s}^{-1}$ per 100 nm and offset $b = (0.1 \pm 0.1) \text{ m s}^{-1}$ at $\lambda_0 = 600 \text{ nm}$. The Gaussian IP case shows a long-range linear trend of $\Delta \nu_{\text{pair}}$ with wavelength, with $m = (−12.1 \pm 0.7) \text{ m s}^{-1}$ per 100 nm and $b = (−18.0 \pm 0.4) \text{ m s}^{-1}$. At the two extremes, the observed $\Delta \nu_{\text{pair}}$ is $−5.9 \text{ m s}^{-1}$ at $\lambda = 500 \text{ nm}$ and $−28.9 \text{ m s}^{-1}$ at $\lambda = 690 \text{ m s}^{-1}$. Extrapolating the straight line outside of the fitted range, $\Delta \nu_{\text{pair}} = 0 \text{ m s}^{-1}$ is expected at $\lambda = 451 \text{ nm}$ and the blue end of the HARPS wavelength range ($\lambda = 380 \text{ m s}^{-1}$) should have $\Delta \nu_{\text{pair}} = 8.6 \text{ m s}^{-1}$. Although a straight line explains the data sufficiently well, the smooth variation of the binned values with wavelength may hint towards a more complex, non-linear, trend.

6. Results

We conducted a detailed study of the HARPS IP using its astrocomb calibration system. Applying the methods described in Sec. 3 to a single astrocomb calibration frame taken in December 2018, we reconstructed the IP in the wavelength range 500 nm to 690 nm (Sec. 4). The main results of this work are:

1. HARPS IP significantly differs from a Gaussian shape in the entire range covered by the astrocomb (Figs. 10, 13, 14). The IP asymmetry varies quickly across the detector in both the...
3. Using the Gaussian IP assumption is associated with errors on wavelength measurements as large as 50 m s\(^{-1}\) (Fig. 13 right). The errors reveal both intra-order and inter-order wavelength scale distortions which are generally larger at locations where the empirical IP shapes are more strongly asymmetric. The intra-order distortions associated with Gaussian IP use are described well by a second order polynomial and the long-range inter-order distortion by a linear function with a slope \(m = (−2.8 ± 0.3)\) m s\(^{-1}\) per 100 nm and a normalisation \(b = (−6.2 ± 0.2)\) m s\(^{-1}\) at \(λ_0 = 600\) nm (Fig. 14 bottom).

4. No bias was associated with the use of our empirical IP models. Errors on wavelength measurements for 10576 astrocomb lines using empirical IP models follow an approximately normal distribution with a −0.21 m s\(^{-1}\) median and 2.21 m s\(^{-1}\) median absolute deviation. No correlations of errors with position on the detector (Fig. 13 left) or with wavelength (Fig. 14 top) were observed. Examination of 1472 modes appearing twice on the detector (Sec. 5) confirmed this further. Using empirical IP models, wavelength accuracy was better than 10 m s\(^{-1}\) (5 m s\(^{-1}\)) for 92% (64%) of the lines. For comparison, only 29% (8%) of lines were measured with the same accuracy using the Gaussian IP.

5. Distortions associated with the Gaussian IP assumption change smoothly across the detector area, suggesting a connection to spectrograph optics. Similar distortions may therefore be seen in other cross-dispersed spectrographs when the Gaussian assumption is used for wavelength calibration.

### Table 1. Statistical properties of the distributions shown in Fig. 15

<table>
<thead>
<tr>
<th>Quantity (m s(^{-1}))</th>
<th>Empirical IP</th>
<th>Gaussian IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>−0.16</td>
<td>−13.67</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>8.24</td>
<td>8.08</td>
</tr>
<tr>
<td>Median</td>
<td>−0.05</td>
<td>−13.48</td>
</tr>
<tr>
<td>MAD†</td>
<td>3.59</td>
<td>4.34</td>
</tr>
<tr>
<td>Percentiles:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16(^{\text{th}})</td>
<td>−5.52</td>
<td>−19.97</td>
</tr>
<tr>
<td>84(^{\text{th}})</td>
<td>5.65</td>
<td>−7.18</td>
</tr>
<tr>
<td>98(^{\text{th}})</td>
<td>12.91</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Notes. † Median absolute deviation (around the median).

---

7. Discussion

Astrocombs are widely believed to be the ideal absolute wavelength calibration sources, providing accuracy and precision that is limited only by photon counting statistics. This is why several existing and planned high-resolution spectrographs are (expected to be) equipped with astrocombs for wavelength calibration and instrument monitoring and characterisation. While the first statement is likely true, this work demonstrated that the knowledge of the spectrograph’s instrument profile is crucial in unlocking the full potential of astrocombs. By modelling its empirical IP, we achieved a large step forward in terms of achieving accurate wavelength measurements with HARPS.

The IP reconstruction made use of the flexible nature of GP regression, an advantage independently recognised by Hirano et al. (2020), who first used it to reconstruct the IP of a high-resolution astronomical spectrograph. While both we and Hirano et al. (2020) modelled the IP in small sections of the detector, the details of this process differed between us in one major way: we normalised the astrocomb lines and stacked all lines falling in one segment on top of their centres before performing the regressions whereas Hirano et al. (2020) fitted the fluxes of all lines in the segment simultaneously. Another approach to IP reconstruction was put forward by Hao et al. (2020), who used modified Gaussian functions to determine the dominant, smoothly changing IP shape, and used cubic splines to model residuals of this dominant component. The inclusion of a Gaussian mean function in our GP (Sec. 3) renders the two approaches similar. The normalised residuals shown in their figure A.1 bear close resemblance to \(ψ - m\) for our models. One such example is shown in Fig. 4 as a dashed magenta line and in the second panel of Fig. C.1. Unfortunately, we could not compare our IP models to the previous investigation of the HARPS IP using astrocombs by
Zhao et al. (2021) because no IP models were produced as a part of that work.

Our IP models are described by set of 10 hyperparameters (Sec.5). Interestingly, albeit they were derived completely independently, the values of some hyperparameters vary smoothly across the detector area (Fig.7), suggesting that they trace changes in some physical quantities. This is most obviously seen for the hyperparameters describing the mean function ($\alpha$, $\mu$, $\tau$, and $y_0$), but can also be seen for $\alpha$, $\lambda$, and possibly other hyperparameters. Choosing a more physically motivated set of hyperparameters, based on considerations of instrument optics, or constraining them to smoothly change across the detector area similarly to Hao et al. (2020) may provide improvements in the future.

We found conclusive evidence that assuming a Gaussian IP shape introduces both intra-order and inter-order wavelength scale distortions (Secs.5.3 and 5.5). These distortions range between 5 m s$^{-1}$ to 50 m s$^{-1}$ and vary smoothly across the detector area (Fig.13). The latter suggest a connection with the optical design of HARPS. As such, similar distortions may appear in the wavelength calibrations of other cross-dispersed echelle spectrographs when their IP shape is assumed to be Gaussian.

Left uncorrected, these distortions negatively impact on some types of analyses made from the data calibrated using the Gaussian approximation. Their severity, of course, depends on where the spectral features used in the analysis fall with respect to the distortion and the specifics of the analysis being performed. The impact on $\Delta$$\alpha$/$\alpha$ measurements from quasar absorption systems may be important. For example, a 50 m s$^{-1}$ distortion between Fe II 2382 and Mg II 2796 (commonly used for $\alpha$ measurements), could be erroneously interpreted as evidence for $\Delta$$\alpha$/$\alpha$ $\approx$ 3 $\times$ 10$^{-6}$.

Spatial variations of the IP across the detector also affect velocity measurements via the Doppler effect. The spectra of objects (quasars or stars in the case of exoplanet studies) shift by $\pm$ 30 km s$^{-1}$ within 1 year due to changes in the Barycentric Earth Radial Velocity, corresponding to a maximal difference of $\approx$ 80 pixels on HARPS. Measuring line centres using a Gaussian IP when the correct IP is, in fact, significantly asymmetric (such as it is in the top left angle of the right panel of Fig.13), is expected to bias the result by $\sim$ 10 m s$^{-1}$. This is diminished by averaging over other lines used to measure the object velocity, with several thousands being a typical number used in exoplanet studies and the expected number of Lyman-$\alpha$ lines to be used for the redshift drift measurement (Liske et al. 2008). As an order of magnitude approximation, when averaging over 3000 lines, a 10 m s$^{-1}$ shift in a single line would result in a measured global shift of 18 cm s$^{-1}$ for the entire spectrum. This is similar to the expected signal from an Earth-Sun analogue ($\sim$9 cm s$^{-1}$ in a year) and 60x larger than the cosmological signal expected from redshift drift in $\Lambda$CDM (0.3 cm s$^{-1}$ per year). Should Gaussian IP approximation be used in such studies, a full simulation (considering the locations at which the lines appear on the detector) should be run to quantify instrumental uncertainties.

The presence of intra-order wavelength distortions raises important questions about the practice of order merging when combining object spectra as merging orders containing uncorrected distortions creates a complicated wavelength scale distortion in the combined spectrum. It may also smear out spectral features falling in the affected regions, complicating analysis. A quick (but possibly not very accurate) correction of intra-order distortions in wavelength calibrations produced by fitting Gaussian profiles to HARPS astrocomb spectra, is provided by the coefficients in Appendix E. We warn that we did not check whether the distortions are stable in time so these coefficients may not be applicable for data taken shortly before or after 2018-12-07.

In Sec.4.2 we presented our method for background modelling. To focus on our main aim of IP reconstruction, we did not elaborate discussion of the background further. However, we note that procedures relating to the spectral background determination had a noticeable impact on the final fit quality for astrocomb lines. The background is highly modulated and, looking at its power spectrum, we saw that it contains significant power at periods as short as 30 pix. Significant effort was therefore invested into determining the locations in the minima between the astrocomb modes as accurately as possible and into the background interpolation methods. Future attempts to model the IP from astrocomb spectra could investigate whether further improvements can be made by relating the modulation of the spectral background (connecting minima) to the modulation of the spectral envelope (connecting maxima). Furthermore, astrocombs based on technologies which do not require a high-power amplification stage (e.g. opto-electric LFCs, Zhuang et al. (2023)) may have smaller (or negligible) background levels so should be less sensitive to this issue.

Tables containing the numerical values describing empirical HARPS IP shapes are provided in the supplementary material in a machine readable format and also as a dataset on the following link: https://zenodo.org/doi/10.5281/zenodo.10492989. We hope that their use leads to interesting scientific discoveries.

8. Conclusions & future prospects

With some modification, our (or other similar) methods can be applied to any spectrograph equipped with an astrocomb, such as ESPRESSO (Pepe et al. 2021), EXPRES (Blackman et al. 2020), and Veloce (Gilbert et al. 2018). In the future, astrocombs will also be installed on the Keck Planet Finder (Wu et al. 2022), and two spectrographs to be mounted on 30-m class telescopes: ANDES (Marconi et al. 2022) and G-CLEF (Szentgyorgyi et al. 2018), and possibly also on novel and very interesting spectrograph designs fed by an array of smaller telescopes (Eikenberry et al. 2019; Angel et al. 2022). Most demanding projects planned with these instruments rely on these instruments, not only reaching wavelength calibration stability at the few cm s$^{-1}$ level, but also maintaining it over periods of years or decades. This presents several large technical and methodological challenges that are yet to be resolved, such as resolving the 60 cm s$^{-1}$ zero-point offset in the wavelength calibration when one astrocomb is replaced by another (Milakovitch et al. 2020). The impact of our empirical IP models on that result will be reported in a separate publication.

In the future, an astrocomb with tunable $f_0$ and/or tunable $f$, (see Savchenkov et al. 2008, Yan et al. 2023 e.g.) could be used to determine the IP with unprecedented detail. Changing $f$ by $\Delta f \sim$ 100 MHz would shift the astrocomb line centres by 5-10% of the pixel, allowing for the reconstruction of the IP locally without the need for stacking neighbouring lines or interpolation across large fractions of the detector area. Covering the full frequency range between lines, $f_i$, would then map out the IP at every pixel, as well as allow for the reconstruction of the intra-pixel response function $Q$.

A comparison with the IP derived from an $I_2$ absorption cell would also be interesting. $I_2$ cells have previously been used to determine the IP shape of optical echelle spectrographs directly from the science observations (Marcy & Butler 1992; Butler et al. 1996). HARPS was equipped with an $I_2$ absorption cell...
but it has been decommissioned in 2004 (Rodler & Lo Curto 2019), meaning that a direct comparison of the IP derived from astrometric calibrations and I$_2$ calibrations was not possible on HARPS. There is an ongoing project to compare the IP derived from astrometric against the IP derived from I$_2$ absorption cell for ESPRESSO as a part of preparations for ANDES.

Going from modelling the 1-dimensional IP (the line-spread function, as done here) to modelling the 2-dimensional IP (that is the point-spread function, PSF) would be very useful for improving spectral extraction procedures. Firstly, information on the PSF shape could be used to investigate and measure the effects of charge transfer efficiency and the intra-pixel sensitivity function by monitoring shifts in the centres of individual astrometric lines when flux levels or astrocomb frequencies are varied in a controlled way. These effects could then be modelled and their impact removed from the raw data before spectral extraction. Flux dependent shifts of the HARPS LFC in 2d raw images was already explored in Zhao et al. (2021), who found that they can be equivalent to radial velocity shifts of 15 m s$^{-1}$. Secondly, the PSF shapes and the 2d wavelength calibration of raw frames would produce important calibration information useful for advanced spectral extraction algorithms such as “spectro-perfectionism” (Bolton & Schlegel 2010).

Should all of the above be implemented to several high-resolution spectrographs equipped with astrometers, this could provide the ability to directly compare and even combine spectra obtained using different spectrographs and different telescopes with 1 cm$^{-1}$ accuracy. Currently, the spectra of the same objects obtained using different instruments are often not compatible, complicating result interpretation. The most obvious example relates to the studies of quasar absorption systems in the context of fundamental coupling constants’ variability (Webb et al. 1999). King et al. (2012). For example: the physical model parameters of the absorption system at z = 1.15 towards the quasar HE0515−4414 are different when derived from UVES, HARPS, or ESPRESSO observations (Kotuš et al. 2017; Milaković et al. 2021; Murphy et al. 2022).

Combining data from different instruments and telescopes may be necessary to detect Earth-Sun analogues via the radial velocity (RV) method. Comparing RV measurements of Sun as a star made by four extremely stable spectrographs, Zhao et al. (2023) finds that day-to-day variations in the mean daily RV measured by the four instruments agree within ± 50 cm s$^{-1}$ (see their figure 9). It would be interesting to see whether this number decreases when IP variations are considered in RV calculations, although other systematic effects considered by the authors may be dominant.

Acknowledgements

D. Milaković thanks Luca Pasquini, John Webb, and Paolo Molaro on discussions which improved this work. This research was funded in part by the Austrian Science Fund (FWF) SFB 10.55776/F68 “Tomography Across the Scales”, project F68111-N36 (Advancing Extragalactic Archeology through Novel Inversion Techniques). This work made use of Python programming language packages numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), matplotlib (Hunter 2007), and fitsio (Sheldon et al. 2021), in addition to the software already cited. We also thank the referee for useful suggestions.

References

Foreman-Mackey. D. 2023, tinygr: The tiniest of Gaussian Process libraries
Hänisch, T. W. 2006, Reviews of Modern Physics, 78, 1297
Hirano, T., Kuzuhara, M., Kotani, T., et al. 2020, PASJ, 72, 93
Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
Milaković, D. & Jethwa: A new method for instrumental profile reconstruction of high-resolution spectrographs
A&A proofs: manuscript no. output
Appendix A: Choosing a centre for an asymmetric profile

Our method requires that some measure of an IP centre is provided, so that line centres can be determined through fitting (Eq. 7-9). For asymmetric profiles, there is no meaningful definition of the line centre and, looking through the relevant literature, we found scarce information about this particular point.

A&K introduced their own centre estimator based on the considerations for the distribution of flux on the detector produced by an unresolved source. They posited that the IP is properly centred when an unresolved source located at the boundary between the two brightest pixels results in equal fluxes in the boundary pixels, i.e. \( \psi(\Delta x = -0.5) = \psi(\Delta x = 0.5) \). As a reminder, \( \Delta x \) is the distance of the pixel centre from the unresolved spectral feature. This definition is of the centre depends, of course, on the IP shape in the central region but its only assumption is that the IP is a smooth function peaking somewhere in between the two pixel centres. It was developed specifically to deal with asymmetric IP shapes but its performance with respect to other centre estimators has never before been quantified so we were unsure whether it was the most appropriate for our purposes.

We therefore performed a numerical exercise to determine what is the most appropriate centre estimator for asymmetric IPs. The criterion chosen for their ranking was the estimator’s invariability under small perturbations of the IP shape (robustness), where the best estimator would be the least variable. Four different estimators were considered: the mean, the median, the mode, and the definition of line centre from Anderson & King (2000). The numerical exercise made use of the MCMC models from Appendix B and posterior distributions for hyperparameter values in particular. The posterior distribution for relevant hyperparameters was sampled 1000 values to create 1000 IP models shown in Fig. A.1. Centres of these models were subsequently determined using each of the four estimators and their distributions were examined.

The mean values for all four estimators are different and they are not consistent with each other based on the widths of their distributions (see the inset in Fig. A.1). The largest difference between any two estimators is \( \approx 0.12 \) pix (between the median and A&K’s estimator) and the smallest difference is \( \approx 0.01 \) pix (between the median and the mean). Distributions for all four estimators show a single peak and are approximately symmetric around their central value (see Fig. A.2). Concentrating on the distribution width, we see significant differences across the estimators. The text in each of the four panels of the same figure gives the central 68% distribution limits, when the distribution median is set to zero. Ranking the four estimators according to their distribution widths (from the most narrow to the most broad), we obtain: the A&K’s estimator \( \psi(\Delta x = -0.5) = \psi(\Delta x = 0.5) \), the median, the mean, and the mode.

We therefore chose to centre all our empirical IP profiles \( \psi \) at the location which results in \( \psi(\Delta x = -0.5) = \psi(\Delta x = 0.5) \). This is also reflected in Fig. A.1 which is plotted with its centre at that location.

Appendix B: A full Markov Chain Monte Carlo approach

All of the analysis presented in the main body of this work optimises the GP hyperparameters. While our results show that this is sufficient to describe the data, we wanted to check whether...
uncertainty in the hyperparameter values may significantly affect our conclusions. To test this, we sampled the posterior distribution of the hyperparameters via Markov Chain Monte Carlo (MCMC) algorithm for several randomly chosen segments. Specifically, we used the Hamiltonian Monte Carlo (MCMC) algorithm as implemented in the nupype (Bingham et al. 2019; Phan et al. 2019) software.

The posterior comprises a likelihood identical to that used in section 4.5 (i.e. with one GP for the line profile and another for the excess variance) along with prior distributions on the hyperparameters. For these ‘hyperpriors’, we used normal distributions whose mean and variance were set to some reasonable number (see Fig. B.1). We ran four chains using 300 and 600 sampling steps. We assessed convergence of the chains via two diagnostics: the split Gelman-Rubin $\hat{R}$ (Vehtari et al. 2021) and $N_{\text{eff}}$ (Geyer 2011). The values of these diagnostics for all fitted parameters are shown in Tab. B.1.

The median IP model derived through MCMC is indistinguishable from the IP model derived through the L-BFGS-B method and the differences between the two models are consistent with the noise (Fig. B.2). The posterior probability distributions for the hyperparameters $A, \mu, \sigma, \gamma_0, \lambda, l$ for the instrumental profile, and $(\alpha_g, l_g)$ for the excess variance are shown in Fig. B.4. The hyperparameter values returned by the L-BFGS-B method for the same data are also shown as red lines and squares, and printed out in the top right corner of the figure. Hyperparameter values associated with the IP shape returned by the L-BFGS-B method all fall within the central 68% of the posterior distribution determined using MCMC.

On the other hand, hyperparameters controlling the variance modification all fall outside of the central 68% of their respective MCMC posterior distributions. However, overplotting the $g(\Delta x)$ from both methods in Fig. B.3 we see that the two functions agree very well so this disagreement in hyperparameter values is assumed not to be problematic.

Table B.1. The convergence of the MCMC chains was confirmed using two statistics: $\hat{R}$ (should be near 1) and $N_{\text{eff}}$ (should be over 50).

<table>
<thead>
<tr>
<th>Quantity</th>
<th>$\hat{R}$</th>
<th>$N_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>1.000</td>
<td>1262.332</td>
</tr>
<tr>
<td>$\mu$</td>
<td>1.001</td>
<td>913.540</td>
</tr>
<tr>
<td>$\log(\sigma)$</td>
<td>1.001</td>
<td>1587.845</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.999</td>
<td>1070.650</td>
</tr>
<tr>
<td>$\log(l_g)$</td>
<td>1.001</td>
<td>810.947</td>
</tr>
<tr>
<td>$\log(l_\lambda)$</td>
<td>1.014</td>
<td>93.589</td>
</tr>
<tr>
<td>$\log(\alpha_g)$</td>
<td>1.006</td>
<td>198.493</td>
</tr>
<tr>
<td>$\log(C_g)$</td>
<td>0.999</td>
<td>54.589</td>
</tr>
<tr>
<td>$\psi(\Delta x)$</td>
<td>1.005$^\star$</td>
<td>711.299$^\star$</td>
</tr>
<tr>
<td>$g(\Delta x)$</td>
<td>1.091$^\star$</td>
<td>82.120$^\star$</td>
</tr>
</tbody>
</table>

$^\star$ Average over the data points.

Fig. B.1. Prior and posterior distributions for $\psi(\Delta x)$ and $g(\Delta x)$. Top panel: The data correspond to the tenth segment of order 110. The dashed blue line is the median of the prior distribution for $\psi(\Delta x)$. The blue shaded bands around the blue line show the central 68% and 95% of the prior distribution. The red solid line and the corresponding red bands show the same quantities for the posterior distribution. Bottom panel: The lines and the bands have the same meaning as in the top panel but are for function $g(\Delta x)$ which modifies the variance array as per Eq. (24).

Fig. B.2. Comparison between IP models derived in two different ways. The model derived using the L-BFGS-B is shown as a solid red line (overlaid on top of black points) and the MCMC derived model is shown as a dashed blue line (overlaid on top of grey points). The two IP models are indistinguishable from one another. The two sets of lines and data points have been offset by 3 in y direction for clarity. The normalised residuals between the two models, that is the MCMC model minus the L-BFGS-B model divided by the error on the data, are shown as a black dots at the top of the panel. The two dotted horizontal lines bracketing the residuals indicate $\pm 1\sigma$, where $\sigma$ is the error on the data.
Fig. B.3. Comparison of $g(\Delta x)$ derived using MCMC and L-BFGS-B. The black squares and the solid black line show the method for variance modification based on L-BFGS-B presented in the main body of the paper and implemented in this work. The dashed red line shows the median of the MCMC posterior distribution for $g(\Delta x)$ and the thin solid red lines show the central 68% limits. The two models are in excellent agreement.
A&A proofs: manuscript no. output

L-BFGS-B parameters

\begin{align*}
A &= 100.985 \text{ (arb.)} \\
\mu &= -0.005 \text{ (pix)} \\
\sigma &= 1.406 \text{ (pix)} \\
y_0 &= -0.070 \text{ (arb.)} \\
a &= 1.419 \text{ (arb.)} \\
l &= 0.961 \text{ (pix)} \\
\log(\epsilon_0) &= -9.899 \\
a_g &= 0.715 \text{ (arb.)} \\
l_g &= 1.717 \text{ (pix)} \\
C_g &= 0.916 \text{ (arb.)}
\end{align*}

Fig. B.4. Posterior distribution of hyperparameters in MCMC calculations. The corresponding L-BFGS-B values are printed in the top right corner and are indicated by red lines in all panels.
Appendix C: Details of a single example model
Fig. C.1. Detailed view of the IP model for a single segment. First panel from the top: The black points are the 251 samples of the IP, $\hat{\psi}$, in the first segment of order 90 derived from flux normalised and stacked astrocomb lines (Eq. (10)). The thick red line is the empirical IP model, $\psi(x)$ in the first iteration. The model was derived by performing a GP regression on the black points (Sec. 2.4) with a secondary GP for variance estimation. The dashed blue line is the mean function of the GP, Eq. (17). The numbers to the right of the panel give additional information. The first seven numbers are the values of the parameters ($\theta, \phi, \sigma_0$) that minimise the loss function, $-\log L$, where $L$ is the model likelihood (also quoted). $N$ and $\nu$ are the number of data points and the number of degrees of freedom in the fit, respectively. $\chi^2$ is given by Equation (22) and $\chi^2/\nu = \chi^2/\nu$. The final number specifies the shift applied to $\Delta x$ to ensure proper centring, Eq. (20). Second panel: The black points show the difference between the data and the GP mean function (a Gaussian with parameters $\theta = \{A, \mu, \sigma, y_0\}$). The red line shows $\psi(x) - m(x)$, i.e. departures from the Gaussian shape. Third panel: Grey dots are the variances on the $\hat{\psi}$, i.e. Eq. (11). Orange points and bands are the variances inferred from the data (as per Eq. (24)) and the corresponding uncertainty. The larger black dots are the sum of modified $\varepsilon_{\text{data}}$ and $\sigma_0$. The red line is the variance of the GP model, i.e. the diagonal of $K_{i,j}$, Equation (18). Fourth panel: The grey dots show the normalised residuals. The grey shaded area shows the ±1σ range and the dashed black lines show the 5th and 95th percentiles. The small panel to the right shows the histogram of the residuals. The shaded area and the dashed lines are the same as in the panel immediately to the left. The number in the top right corner of the panel is the median, and the upper and lower limits correspond to the 5th and 95th percentiles.

Appendix D: IP models in velocity space
Fig. D.1. HARPS IP models in velocity space for orders 89 through 122. Each panel shows the IP in 16 segments of one order, whose central wavelength is printed above it. Line colour within the panel changes from blue to red with increasing wavelength. The same plots with background reference grids (which make it easier to see the profile changes) can be found on the arXiv version of this manuscript [https://doi.org/10.48550/arXiv.2311.05240].
### Appendix E: Table with coefficients for intra-order distortions

<table>
<thead>
<tr>
<th>Order</th>
<th>$\lambda_0$ (nm)</th>
<th>$c_2$ (m s$^{-1}$ nm$^{-2}$)</th>
<th>$c_1$ (m s$^{-1}$ nm$^{-1}$)</th>
<th>$c_0$ (m s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
<td>501.804</td>
<td>$-0.4 \pm 0.3$</td>
<td>$1.5 \pm 0.5$</td>
<td>$-2 \pm 1$</td>
</tr>
<tr>
<td>121</td>
<td>505.814</td>
<td>$-0.6 \pm 0.1$</td>
<td>$2.5 \pm 0.2$</td>
<td>$-2.4 \pm 0.4$</td>
</tr>
<tr>
<td>120</td>
<td>510.027</td>
<td>$-0.65 \pm 0.09$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.6 \pm 0.3$</td>
</tr>
<tr>
<td>119</td>
<td>514.319</td>
<td>$-0.67 \pm 0.09$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.5 \pm 0.3$</td>
</tr>
<tr>
<td>118</td>
<td>518.675</td>
<td>$-0.67 \pm 0.09$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-2.6 \pm 0.3$</td>
</tr>
<tr>
<td>117</td>
<td>523.098</td>
<td>$-0.7 \pm 0.1$</td>
<td>$2.6 \pm 0.2$</td>
<td>$-2.8 \pm 0.4$</td>
</tr>
<tr>
<td>116</td>
<td>527.614</td>
<td>$-0.68 \pm 0.09$</td>
<td>$2.8 \pm 0.1$</td>
<td>$-2.9 \pm 0.3$</td>
</tr>
<tr>
<td>115</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>113</td>
<td>541.623</td>
<td>$-0.64 \pm 0.07$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.6 \pm 0.3$</td>
</tr>
<tr>
<td>112</td>
<td>546.447</td>
<td>$-0.70 \pm 0.07$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-2.5 \pm 0.3$</td>
</tr>
<tr>
<td>111</td>
<td>551.377</td>
<td>$-0.72 \pm 0.07$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-2.4 \pm 0.3$</td>
</tr>
<tr>
<td>110</td>
<td>556.378</td>
<td>$-0.69 \pm 0.07$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.6 \pm 0.3$</td>
</tr>
<tr>
<td>109</td>
<td>561.489</td>
<td>$-0.67 \pm 0.07$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.6 \pm 0.3$</td>
</tr>
<tr>
<td>108</td>
<td>566.676</td>
<td>$-0.66 \pm 0.06$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-2.8 \pm 0.3$</td>
</tr>
<tr>
<td>107</td>
<td>571.979</td>
<td>$-0.68 \pm 0.06$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-2.6 \pm 0.3$</td>
</tr>
<tr>
<td>106</td>
<td>577.362</td>
<td>$-0.68 \pm 0.06$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.7 \pm 0.3$</td>
</tr>
<tr>
<td>105</td>
<td>582.868</td>
<td>$-0.65 \pm 0.06$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-2.8 \pm 0.3$</td>
</tr>
<tr>
<td>104</td>
<td>588.470</td>
<td>$-0.63 \pm 0.07$</td>
<td>$2.5 \pm 0.1$</td>
<td>$-2.9 \pm 0.3$</td>
</tr>
<tr>
<td>103</td>
<td>594.180</td>
<td>$-0.67 \pm 0.05$</td>
<td>$2.65 \pm 0.09$</td>
<td>$-3.0 \pm 0.3$</td>
</tr>
<tr>
<td>102</td>
<td>600.002</td>
<td>$-0.70 \pm 0.06$</td>
<td>$2.6 \pm 0.1$</td>
<td>$-2.9 \pm 0.3$</td>
</tr>
<tr>
<td>101</td>
<td>605.940</td>
<td>$-0.66 \pm 0.06$</td>
<td>$2.7 \pm 0.1$</td>
<td>$-3.1 \pm 0.3$</td>
</tr>
<tr>
<td>100</td>
<td>611.996</td>
<td>$-0.68 \pm 0.06$</td>
<td>$2.9 \pm 0.1$</td>
<td>$-3.2 \pm 0.3$</td>
</tr>
<tr>
<td>99</td>
<td>618.174</td>
<td>$-0.69 \pm 0.06$</td>
<td>$2.9 \pm 0.1$</td>
<td>$-3.5 \pm 0.3$</td>
</tr>
<tr>
<td>98</td>
<td>624.490</td>
<td>$-0.70 \pm 0.05$</td>
<td>$3.0 \pm 0.1$</td>
<td>$-3.6 \pm 0.3$</td>
</tr>
<tr>
<td>97</td>
<td>630.913</td>
<td>$-0.75 \pm 0.06$</td>
<td>$3.1 \pm 0.1$</td>
<td>$-3.6 \pm 0.3$</td>
</tr>
<tr>
<td>96</td>
<td>637.493</td>
<td>$-0.80 \pm 0.06$</td>
<td>$3.2 \pm 0.1$</td>
<td>$-3.8 \pm 0.3$</td>
</tr>
<tr>
<td>95</td>
<td>644.225</td>
<td>$-0.87 \pm 0.07$</td>
<td>$3.3 \pm 0.1$</td>
<td>$-3.7 \pm 0.4$</td>
</tr>
<tr>
<td>94</td>
<td>651.049</td>
<td>$-0.86 \pm 0.07$</td>
<td>$3.7 \pm 0.1$</td>
<td>$-4.3 \pm 0.4$</td>
</tr>
<tr>
<td>93</td>
<td>658.046</td>
<td>$-0.91 \pm 0.06$</td>
<td>$3.8 \pm 0.1$</td>
<td>$-4.6 \pm 0.4$</td>
</tr>
<tr>
<td>92</td>
<td>665.208</td>
<td>$-1.02 \pm 0.06$</td>
<td>$4.2 \pm 0.1$</td>
<td>$-4.9 \pm 0.4$</td>
</tr>
<tr>
<td>91</td>
<td>672.500</td>
<td>$-1.11 \pm 0.06$</td>
<td>$4.6 \pm 0.1$</td>
<td>$-5.3 \pm 0.4$</td>
</tr>
<tr>
<td>90</td>
<td>679.982</td>
<td>$-1.20 \pm 0.07$</td>
<td>$4.8 \pm 0.1$</td>
<td>$-5.8 \pm 0.5$</td>
</tr>
<tr>
<td>89</td>
<td>687.629</td>
<td>$-1.42 \pm 0.07$</td>
<td>$5.4 \pm 0.1$</td>
<td>$-5.8 \pm 0.5$</td>
</tr>
</tbody>
</table>

Table E.1. Coefficients for a second order polynomial describing intra-order wavelength scale distortions associated with Gaussian IP use, i.e. $\epsilon(\lambda, \lambda_0) = \sum_{i=0}^{2} c_i (\lambda - \lambda_0)^i$. Order 115 falls in between two CCDs comprising the HARPS detector so is not recorded.