**Article:**
Bestenlehner, J.M. orcid.org/0000-0002-0859-5139, Enßlin, T., Bergemann, M. orcid.org/0000-0002-9908-5571 et al. (3 more authors) (2024) Spectroscopic analysis of hot, massive stars in large spectroscopic surveys with de-idealized models. Monthly Notices of the Royal Astronomical Society, 528 (4). pp. 6735-6750. ISSN 0035-8711

https://doi.org/10.1093/mnras/stae298

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **528,** 6735–6750 (2024) https://doi.org/10.1093/mnras/stae298
Advance Access publication 2024 January 30

# Spectroscopic analysis of hot, massive stars in large spectroscopic surveys with de-idealized models

J. M. Bestenlehner [1,2★] T. Enßlin,[3] M. Bergemann [2] P. A. Crowther [1] M. Greiner[3] and M. Selig[4]

[1]*Department of Physics & Astronomy, University of Sheffield, Hounsfield Road, Sheffield S3 7RH, UK*
[2]*Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*
[3]*Max Planck Institute for Astrophysics, Karl-Schwarzschildstraße 1, D-85748 Garching, Germany*
[4]*DBFZ Deutsches Biomasseforschungszentrum gemeinnützige GmbH, Torgauer Straße 116, D-04347 Leipzig, Germany*

Accepted 2024 January 22. Received 2023 December 21; in original form 2023 August 3

## ABSTRACT

Upcoming large-scale spectroscopic surveys with e.g. WEAVE (William herschel telescope Enhanced Area Velocity Explorer) and 4MOST (4-metre Multi-Object Spectroscopic Telescope) will provide thousands of spectra of massive stars, which need to be analysed in an efficient and homogeneous way. Usually, studies of massive stars are limited to samples of a few hundred objects, which pushes current spectroscopic analysis tools to their limits because visual inspection is necessary to verify the spectroscopic fit. Often uncertainties are only estimated rather than derived and prior information cannot be incorporated without a Bayesian approach. In addition, uncertainties of stellar atmospheres and radiative transfer codes are not considered as a result of simplified, inaccurate, or incomplete/missing physics or, in short, idealized physical models. Here, we address the question of 'How to compare an idealized model of complex objects to real data?' with an empirical Bayesian approach and maximum a posteriori approximations. We focus on application to large-scale optical spectroscopic studies of complex astrophysical objects like stars. More specifically, we test and verify our methodology on samples of OB stars in 30 Doradus region of the Large Magellanic Clouds using a grid of FASTWIND model atmospheres. Our spectroscopic model de-idealization analysis pipeline takes advantage of the statistics that large samples provide by determining the model error to account for the idealized stellar atmosphere models, which are included into the error budget. The pipeline performs well over a wide parameter space and derives robust stellar parameters with representative uncertainties.

**Key words:** atomic data – methods: data analysis – methods: statistical – techniques: spectroscopic – stars: fundamental parameters – stars: massive.

## 1 INTRODUCTION

With the advent of large spectroscopic surveys using instruments such as WEAVE (William herschel telescope Enhanced Area Velocity Explorer, Jin et al. 2023) and 4MOST (4-metre Multi-Object Spectroscopic Telescope, de Jong et al. 2019), tens of thousands of spectra of massive stars ($\gtrsim$10 $M_\odot$) will be obtained, which will need to be analysed in a homogeneous and efficient way (e.g. Bensby et al. 2019; Chiappini et al. 2019; Cioni et al. 2019). Current pipelines of large spectroscopic surveys are largely designed for FGK stars, which are either data driven (e.g. Ness et al. 2015; Guiglion et al. 2020) or model driven (e.g. Allende-Prieto & Apogee Team 2015; Ting et al. 2019). Traditionally, and still widely performed today, massive stars have been analysed by 'eye', which limits the sample size to $<$100 massive stars. In addition, stellar parameters as well as uncertainties are estimated rather than determined. Larger samples of a couple of hundreds of stars are usually analysed with a $\chi^2$-minimization algorithm, where the final fit often needs to be visually verified depending on the goodness of fit.

Multidimensional probability distribution functions are obtained depending on the number of free parameters and uncertainties are then defined on confidence intervals rather than Gaussian standard deviations. Those uncertainties can be highly asymmetric and very large in the case of degenerated parameters. In the massive star community, there are two main flavours of $\chi^2$-minimization algorithms, grid based (e.g. Simón-Díaz et al. 2011; Castro et al. 2012; Bestenlehner et al. 2014) and genetic algorithms on the basis of natural selection (e.g. Mokiem et al. 2007; Brands et al. 2022). All those algorithms use a pre-defined selection of spectral line regions for their analysis.

Theoretical models of complex physical systems are necessarily idealizations. The implied simplifications allow us to focus the view on the essential physics, to keep the model computationally feasible, and to investigate systems for which not all of their components are perfectly known. In contrast to solar-like stars, massive stars have strong stellar winds, which influence the structure of the stellar atmospheres (line blanketing). In extreme cases, the stellar wind can be optically thick. The location, where the optical depth of the atmosphere is 2/3 and therefore where the effective radius and temperature are defined, can be well within the stellar wind

★ E-mail: j.m.bestenlehner@sheffield.ac.uk

volume. The inclusion of line-driven winds into stellar atmosphere models requires the assumption of spherical geometry in the co-moving frame of the star (expanding atmosphere). In addition to effective temperature and surface gravity, mass-loss rate, velocity law, terminal velocity, wind inhomogeneity, and line blanketing need to be included into the stellar atmosphere code. The stellar atmospheres of massive stars significantly depart from local thermal equilibrium (LTE) and therefore must be computed in fully non-LTE, which is computationally expensive (e.g. Santolaya-Rey, Puls & Herrero 1997; Hillier & Miller 1998; Gräfener, Koesterke & Hamann 2002). This limits state-of-the-art stellar atmosphere codes for hot, massive stars to 1D.

When faced with real data of the actual system, these models often perform insufficiently when judged on a goodness-of-fit basis. The difference between real and modelled data can easily exceed the error budget of the measurement. The reason is that the idealized models do not capture all aspects of the real systems, but they are still present on the data. To discriminate these missing aspects from measurement errors or noise, we use the term *model errors* to describe these imperfections of our theoretical description (Oberpriller & Enßlin 2018).

Often, one tries to determine the model parameters from the data via a likelihood-based methodology such as $\chi^2$-minimization, maximum likelihood, or Bayesian parameter estimation that uses the measurement uncertainties as a metric in data space. The resulting parameter estimates can be strongly distorted by the desire of the method to minimize all apparent differences between predicted and real data, indifferently if these are due to measurement or model errors.

Thus, the model errors should be included into the error budget of the parameter estimation. This would require that we have a model for the not yet captured aspects of the system or at least for the model errors these produce. To do this thoroughly, we would need to undertake a case-by-case analysis of the missing physics.

However, this would be quite impractical in cases of the spectroscopic analysis of complex astrophysical objects. Instead, we want to construct a plausible, but by no means perfect, effective description of the model errors. In the construction of the de-idealization model, we will follow a pragmatic route, but try to indicate the assumptions, approximations, and simplifications made.

In Section 2, we introduce the methodology, which is used in our spectroscopic analysis pipeline (Section 3). Using grids of stellar atmospheres (Section 3.3), the pipeline is applied to observational data (Section 3.4). The results are discussed in Section 4. We close with a brief conclusion and outlook (Section 5).

## 2 METHOD

### 2.1 Data model

We assume that we have a set of objects (e.g. stars, galaxies, etc., labelled by $i \in \{1, 2, \ldots, n\}$) with observable signals $s^{(i)} = (s_x^{(i)})_x$ over some coordinate $x$, e.g. the emitted spectral energy distribution $s^{(i)} = (s_\lambda^{(i)})_\lambda$ as a function of the wavelength $\lambda$. These signals are measured with a linearly responding instrument (response matrix $R$) with additive Gaussian noise ($n$) according to the measurement equation

$$d^{(i)} = R^{(i)} s^{(i)} + n^{(i)}. \tag{1}$$

The individual elements of the data vector ($d$) for the $i$th object are then given by

$$d_j^{(i)} = \int \mathrm{d}x \; R_{j\,x}^{(i)} s_x^{(i)} + n^{(i)}, \tag{2}$$

where in our spectroscopic cases $R_{j\,x}^{(i)}$ is the $j$th bandpass of our $i$th observation as a function of wavelength $x = \lambda$. Spectroscopic, colour filter, and bolometric measurements can thereby be treated with the same formalism and even combined into a single data vector and response matrix. In addition, we do not require that all objects are observed in the same way by keeping the response matrix dependent on the object index $i$. In this way, the formalism permits to combine heterogeneous observations.

For the measurement noise $n^{(i)}$ of the $i$th observation, we use the error spectrum from the data reduction that is assumed for simplicity to be Gaussian with zero mean and signal independent,

$$\begin{aligned}\mathcal{P}(n^{(i)}|s^{(i)}) &= \mathcal{G}(n^{(i)}, N^{(i)}) \\ &= \frac{1}{\sqrt{|2\pi N^{(i)}|}} \exp\left[-\frac{1}{2} n^{(i)\dagger} \left(N^{(i)}\right)^{-1} n^{(i)}\right]\end{aligned} \tag{3}$$

with assumed noise covariance $N^{(i)} = \langle n^{(i)} n^{(i)\dagger} \rangle_{(n^{(i)}|s^{(i)})}$. The dagger denotes the transposed and complex conjugated vector. The noise of the different observations is assumed to be independent as well, $\mathcal{P}(n|s) = \mathcal{G}(n, N) = \prod_i \mathcal{G}(n^{(i)}, N^{(i)})$ with $n = (n^{(i)})_i$ and $s = (s^{(i)})_i$ being the combined noise and signal vectors, respectively.

### 2.2 Model errors

Now, we assume that some idealized model for our objects exists that predicts a specific theoretical signal $t^{[p]}$ given a set of unknown model parameters $p$. These parameters should be physical quantities like surface gravity, radius, effective temperature, stellar wind properties, or chemical composition of the object, so that well-defined values $p^{(i)}$ exist for each object.[1] Those idealized models can be generated with stellar atmosphere and radiative transfer codes. Knowing these parameters for each object is the primary goal of the inference. In principle, the relation between parameters and signal could be stochastic, but for simplicity we concentrate on deterministic models. The de-idealization model we develop should serve as an effective description for the remaining stochasticity.

The idealized model hopefully captures the dominant properties of the system but certainly not all aspects. Therefore, the real signal $s^{(i)}$ of an object will deviate by an unknown stochastic component $u^{(i)}$, the model error, so that

$$s^{(i)} = t^{[p_i]} + u^{(i)}. \tag{4}$$

The aim of model de-idealization is to find an appropriate stochastic model $\mathcal{P}(u|p) = \prod_i \mathcal{P}(u^{(i)}|p^{(i)})$ for the model errors. With such, the likelihood becomes

$$\mathcal{P}(d|p) = \int \mathcal{D}u \; \mathcal{P}(d|s = t^{[p]} + u)\, \mathcal{P}(u|p), \tag{5}$$

where $\int \mathcal{D}u$ denotes a phase space integral for the model errors.

[1] Counterexample would be purely phenomenological parameters, describing aspects of the data that contain observation-dependent properties or such that only make sense within a specific object description methodology. Although the approach proposed here might be applicable to such phenomenological descriptions as well, we currently demand the physical existence of the used parameters in order to be on epistemologically firm grounds.

In our case, we want to restrict ourselves to using the simplest possible representation of the model uncertainties. This means that we take only the first and second moments of $u^{(i)}$ into account, $v^{(i)} = \langle u^{(i)} \rangle_{(u|p)}$ and $U^{(i)} = \langle (u-v)^{(i)} (u-v)^{(i)\dagger} \rangle_{(u|p)}$, and assume the fluctuations of different objects to be independent, $\langle u^{(i)} u^{(j)\dagger} \rangle_{(u|p)} = v^{(i)} v^{(j)\dagger} + \delta_{ij} U^{(i)}$. The probability distribution that represents mean and variance without further information on higher order correlations naturally is a Gaussian with this mean and variance. Among all possible probability distributions with given mean and variance, it has a maximal entropy (e.g. Jaynes & Bretthorst 2003; Caticha 2008). By adopting a Gaussian for the model errors,

$$\mathcal{P}(u|p) = \mathcal{G}(u - v,\, U), \tag{6}$$

the least amount of spurious information is entered into the inference system in case only $v$ and $U = \langle (u-v)\,(u-v)^{\dagger} \rangle_{(u|p)}$ are considered.

This does not mean that the model error statistics is a Gaussian in reality. It just means that higher order correlations are ignored for the time being. Taking such higher order correlations into account would most certainly improve the method, but is left for future work.

The Gaussianity of measurement noise and modelling error description permits us to integrate equation (5) analytically leading to

$$\mathcal{P}(d|p,\, v,\, M) = \mathcal{G}(d - R(v + t^{[p]}),\, M), \tag{7}$$

with $M = N + R\, U\, R^{\dagger}$ being the combined error covariance.

### 2.3 Implicit hyperprior

The de-idealization model requires that the auxiliary parameters $v$ and $U$ are determined as well, or better marginalized over. This requires that we specify our prior knowledge on these parameters, $\mathcal{P}(v, U|p)$, which is a very problem specific task. Using such a hyperprior, we could then derive an auxiliary parameter marginalized estimator for our desired model parameters $p$. A good, but numerically very expensive approach would be to sample over the joint space of model and auxiliary parameters, $p$, $v$, and $U$, for example using the Gibbs sampling method (e.g. Wandelt, Larson & Lakshminarayanan 2004; Jasche et al. 2010).

In order to have a generic, affordable, and pragmatic method, we introduce a number of approximations and simplifications. The first is that we replace the auxiliary parameter marginalization by an estimation using the following approximation:

$$\begin{aligned} \mathcal{P}(d,\, p) &= \int \mathcal{D}v \int \mathcal{D}U\, \mathcal{P}(d|p,\, v,\, U)\, \mathcal{P}(v, U|p)\, \mathcal{P}(p) \\ &\approx \mathcal{P}(d|p,\, v^{\star},\, U^{\star})\, \mathcal{P}(p) \end{aligned} \tag{8}$$

with $v^{\star}$ and $U^{\star}$ being suitable estimates of the auxiliary parameters and $\mathcal{P}(p)$ the parameter prior. Instead of constructing these point estimators using the so far unspecified and problem specific priors we propose to pragmatically specify them with an educated ad hoc construction.

The idea is to assume for a moment that a correct model parameter classification $p^{(i)}$ for any object exists, which has later on to be estimated self consistently with all the other estimates via iteration. The difference of the signals reconstructed from the data $m^{(i)} = \langle s^{(i)} \rangle_{(s^{(i)}|d^{(i)},\, p^{(i)})}$ and the one predicted from the model $t^{[p^{(i)}]}$ plus the current guesses for $v^{\star}$,

$$\delta^{(i)} = m^{(i)} - t^{[p^{(i)}]} - v^{(i)\star} \tag{9}$$

can be analysed to provide information on $v$ and $U$. The signal reconstruction can be done via a Wiener filter, since this is optimal in case of a linear measurement and Gaussian noise model equation (7). For the signal difference this is

$$\begin{aligned} \delta^{(i)} &= D^{(i)} R^{(i)\dagger} \left(N^{(i)}\right)^{-1} \left[ d^{(i)} - R^{(i)} \left( t^{[p^{(i)}]} + v^{(i)\star} \right) \right] \\ D^{(i)} &= \left[ \left(U^{(i)\star}\right)^{-1} + R^{(i)\dagger} \left(N^{(i)}\right)^{-1} R^{(i)} \right]^{-1}, \end{aligned} \tag{10}$$

where $D^{(i)}$ is the Wiener variance or uncertainty of the reconstruction. For the current case, we have used some guesses for $v^{\star}$ and $U^{\star}$ that need to be updated accordingly to the information contained in the statistics of the signal differences $\delta^{(i)}$. To do so, we introduce a suitable proximity measure in parameter space, $\omega_{ii'} = \mathrm{prox}(p^{(i)}, p^{(i')})$, that indicates for an object $i$ how much another objects $i'$ can be used to learn about the model error statistics of $i$. A naive choice would be $\omega_{ii'} = 1$ always, assuming that the model error statistics is everywhere the same in the model parameter space. A more sophisticated method would partition the parameter space into characteristic regions (e.g. corresponding to the different known star and galaxy classes in our spectroscopic example) and to set $\omega_{ii'} = 1$ or $\omega_{ii'} = 0$ in case $i$ and $i'$ belong to the same or different classes. Even more sophisticated proximity weighting schemes can be imagined with $\omega_{ii'} = 1/(1 + \mathrm{dist}(p^{(i)}, p^{(i')}))$ using some distance measure in parameter space. However, in our case, we group objects together with respect to their main line diagnostics and analyse them in the same batch.

Given such a scheme, the update operation for $v^{\star}$ and $U^{\star}$ are

$$v^{(i)\star} \rightarrow v^{(i)\star} + \sum_{i'} \frac{\omega_{ii'}}{\Omega_i} \delta^{(i')},$$

$$U^{(i)\star} = \sum_{i'} \frac{\omega_{ii'}}{\Omega_i} \left( \delta^{(i')} \delta^{(i')\dagger} + D^{(i')} \right) \text{ with} \tag{11}$$

$$\Omega_i = \sum_{i'} \omega_{ii'}. \tag{12}$$

The $v^{\star}$ update operation is a simple absorption of any systematic difference into the mean component of the model error $v^{\star}$. For an initial step, it is better to set $v^{(i)\star}_{\mathrm{iteration\#0}} = 0$ as it would absorb any differences between data and model, even though the model might be not representative of the objects. However, we find that the best choice for $v$ is to represent non-stellar features, which are not part of the model, like nebular lines, interstellar bands, or telluric lines.

The $U^{\star}$ update incorporates the variance in the signal difference reconstructions as well as their Wiener variances. The latter express the level of missing fluctuations of the Wiener filter reconstruction. The correction with the Wiener variance is done in analogy to the critical filter methodology developed in (Enßlin & Weig 2010; Enßlin & Frommert 2011), where a similar variance estimation was derived under a non-informative prior on the power spectrum of a statistical homogeneous random process. For this study, we intuitively extended this to non-diagonal covariance structures. This means that we adopt implicitly a non-informative hyperprior for the model error statistics as summarized by $v$ and $U$.

The logic behind this implicit prior is as follows. Assuming we would have managed to specify an appropriate non-informative hyperprior for $v$ and $U$. From this, we would derive some recipe for our point estimates $v^{\star}$ and $U^{\star}$ using some approximations. The resulting recipe should not incorporate hidden spurious assumptions. The $v^{\star}$ and $U^{\star}$ estimates therefore can only be build from elements like $\delta^{(i')}$, $\delta^{(i')}\delta^{(i')\dagger}$, and $D^{(i')}$, the latter being a summary of the former elements. Requiring the estimators to be unbiased and to enclose the mentioned critical filter as a limiting case, which is an unbiased power spectrum estimator, then fixes the numerical coefficients in front of $\delta^{(i')}$, $\delta^{(i')}\delta^{(i')\dagger}$, and $D^{(i')}$ in equation (11) to unity.

We admit that there are some frequentist elements in this derivation, since it postulates an estimator and argues for its appropriateness using bias arguments. We hope that it will be replaced by a more stringent calculation once a suitable non-informative hyperprior has been specified. For the time being, we use it in iteration with model parameter estimation.

### 2.4 Method summary

The combined de-idealized model parameter estimation method comprises the following steps:

(i) Specify the weighting scheme $\omega_{ii'} = \mathrm{prox}(p^{(i)}, p^{(i')})$ that determines how similar the model parameters of two objects have to be so that their model error statistics can be assumed to be similar.
(ii) Adopt some initial guess or default values for the model parameters $p^{(i)}$ and mode error parameters $v^{(i)\star}$ and $U^{(i)\star}$. A naive choice could be $p^{(i)} = p$ for some plausible central $p$ within the physically permitted parameter space, $v^{(i)\star} = 0$, and $U^{(i)\star} = \sum_{i'} \left(R^{(i')\dagger} d^{(i)} - t^{[p]}\right) \left(R^{(i')\dagger} d^{(i)} - t^{[p]}\right)^{\dagger}$.
(iii) Calculate $\delta^{(i)}$ and $D^{(i)}$ according to equation (10) for all objects.
(iv) Update $v^{(i)\star}$ and $U^{(i)\star}$ according to equation (11) for all objects.
(v) Update the model parameters $p^{(i)}$ of all objects using the combined likelihoods equation (7) that incorporates measurement and model errors.
(vi) Repeat steps 3–5 until convergence.

The resulting estimate will be similar to a joint maximum posteriori estimate. A posteriori estimation of the model and model error parameters are known to perform worse than a properly marginalized posterior with respect to an estimator based on the posterior mean. However, given the large number of degrees of freedom in the signal space (e.g. a highly resolved emission spectrum of a star or galaxy), such optimal estimators can be extremely expensive computationally. The proposed method might therefore be favourable in many circumstances, despite its approximative and partly ad hoc nature.

In order to be explicit about the assumptions and approximations adopted, we provide an overview below. This list should help to judge the range of applicability and to find possible improvements of the proposed method. In particular, we assume

(i) that the measurement noise is independent for the different objects and independent of their signals, it has Gaussian statistics with known covariance.
(ii) a linear and known measurement response.
(iii) that the model error knowledge can be approximated by a multivariate Gaussian in signal space.
(iv) that regions in parameter space exist and are known that have similar model error statistics as parametrized by a mean and variance.
(v) no prior knowledge on the values of this model error mean and variances.
(vi) that an iterated point estimate of all involved parameters leads to a reasonable approximative solution of the joint inference problem.

## 3 SPECTROSCOPIC ANALYSIS PIPELINE

The pipeline has been developed using *python3* with commonly used libraries such as *numpy*, *scipy*, *pandas*, *multiprocessing*, and *ctypes* plus *astropy.io* to read fits files. Using commonly used and maintained libraries will ensure that the pipeline is easy to maintain and should be usable over a long period of time. The following section outlines the required pre-processing steps (Section 3.1), brief overview of the pipeline implementation (Section 3.2) and, description of the grid of synthetic model spectra (Section 3.3) and the observational data to verify the pipeline (Section 3.4).

### 3.1 Pre-processing

The pipeline requires that all observational data are read in at the beginning as the spectra are required to be analysed all at once to determine iteratively the stellar parameters and model uncertainties (Section 2.4). Radial velocities shifts of the spectra need to be provided as the misalignment of observation and synthetic spectrum would dominate the model error. After a spectrum is loaded, it is corrected for the potential radial velocity shift, transformed to the wavelength sampling of the synthetic spectra grid (Section 3.3) and decomposed into principal components using the decomposition matrix calculated from the synthetic spectra to reduce the memory usage and speed up the analysis. This is essential, when analysing samples of spectra of more than a few hundred sources.

The decomposed grid of synthetic spectra is loaded into shared memory for parallelization purposes. The spectra are pre-convolved with combinations of varying broadening parameters of projected rotational velocity ($v_{\mathrm{eq}} \sin i$) and macro-turbulent velocity ($v_{\mathrm{mac}}$). The synthetic grid preparation is also a pre-processing step, which is laid out in Section 3.3. If the sample is small and/or sufficient random access memory of the computing system is available, the grid can be convolved with the star specific broadening parameters. Even though the convolution is applied utilizing the fast Fourier transformation library of *scipy*, this could increase the pre-processing time-scale up to a few hours per star depending on the size of the synthetic spectra grid ($\gg 100\,000$).

On a standard Desktop computer, 1000 spectra can be analysed in less than half a day. For larger data sets, we would advise to sort them into groups of similar objects, which will also lead to more representative model error at a specific parameter space when testing implemented physics or verifying assumptions in the theoretical model.

### 3.2 Analysis pipeline

After pre-processing the observational data including observational error spectra and loading the synthetic grid of model spectra, the pipeline is set up according to equation (1) with model de-idealization equation (4) assuming a Gaussian noise model for the observational error (3). We are interested in the posterior distribution of the signal given the data ($\mathcal{P}(s|d)$) and apply the Bayes theorem

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s)\mathcal{P}(s)}{\mathcal{P}(d)} \tag{13}$$

with likelihood $\mathcal{P}(d|s)$, prior $\mathcal{P}(s)$, and evidence $\mathcal{P}(d)$ to use, first, the likelihood $\mathcal{P}(d|s)$ and, secondly, apply the Wiener filter (Section 2.3) to reconstruct the signal ($P(d|s) \rightarrow P(s|d)$). We modified the likelihood as described in Section 2.3 ($\mathcal{P}(d|s) \Rightarrow \mathcal{P}(d|p)$), probability of the data $d$ given the stellar parameters $p$.

The best model is determined from a $\chi^2$-minimization Ansatz with model error variance matrix $U$ (equation 7) to maximize the modified likelihood $\mathcal{P}(d|p)$:

$$\chi^2 = \left(d^{(i)} - \mathrm{R}^{(i)} s^{(i)}\right)^{\mathrm{T}} \mathrm{N}^{(i)^{-1}} \left(d^{(i)} - \mathrm{R}^{(i)} s^{(i)}\right). \tag{14}$$

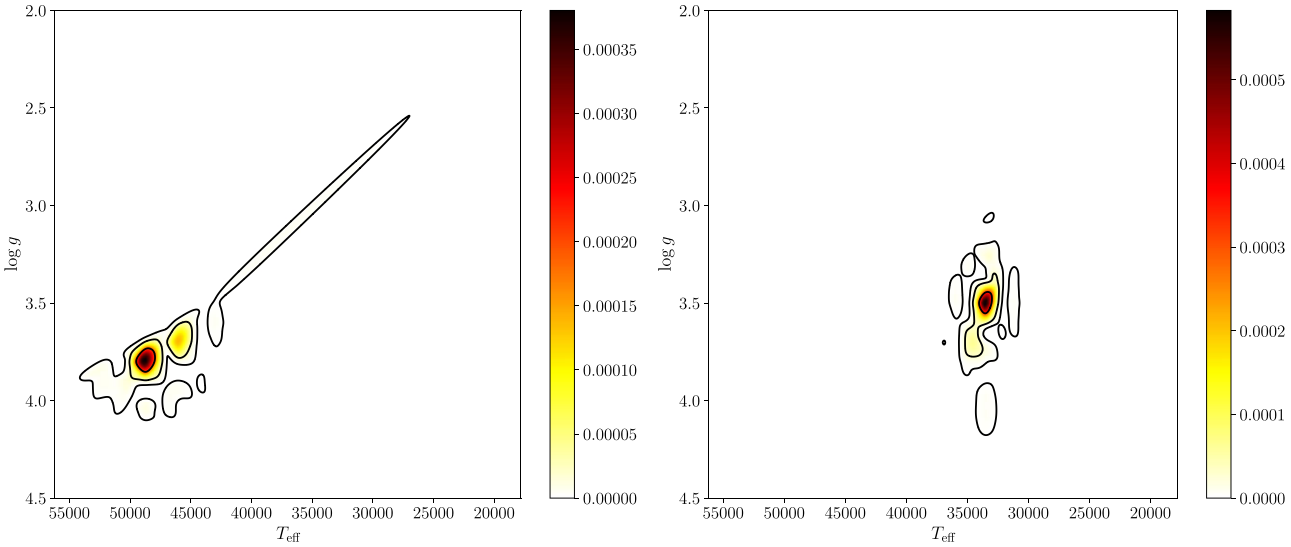


**Figure 1.** Probability heatmap of surface gravity versus effective temperature for VFTS-072 (left) and VFTS-076 (right). Contours indicate 2D standard-deviational ellipse confidence intervals of 39.4, 86.5, and 98.9 per cent (e.g. Wang, Shi & Miao 2015). Spectroscopic fits of those stars are shown in Fig. 3.

For the mean and model error variance matrix $v$ and $U$, we find that it is better to set $v = 0$ and multiply $U$ by a factor $\alpha = [10^{-5}, 0.35, 0.7, 1.0, 1.0]$, which increases after each iteration, to avoid that bad spectroscopic fits have significant impact on $v$ and $U$ and therefore the determination of the best model. $v$ could be also set equal to non-stellar features such as telluric bands, diffuse interstellar bands, and prominent interstellar lines like Ca II H and K. However, interstellar contribution can significantly vary with the line of sight while telluric bands change with atmospheric conditions and airmass. In cases where the non-stellar features vary on a star by star basis, those contributions can be combined with the observational error to give less weight to those spectral regions.

$\omega_{ii'}$ (equation 11) can contain any prior information $\mathcal{P}(p)$, e.g. parameter space of similar objects, stellar structure models, or population synthesis predictions. In the current implementation, we use a flat prior ($\omega_{ii'} = 1$).

The pipeline returns a multidimensional PDF for each star while $U$ is the same for all sources analysed in one batch. Parameters and their uncertainties are determined by defining confidence intervals (Fig. 1). To increase the accuracy, the PDF can be multiplied by an appropriate prior. More details on the implementation can be found in the source code of the pipeline.[2]

### 3.3 Stellar atmosphere grid

The grid of synthetic model spectra was computed with the non-LTE stellar atmosphere and radiative transfer code FASTWIND v10.6 (Santolaya-Rey, Puls & Herrero 1997; Puls et al. 2005; Rivero González et al. 2012) including H, He, C, N, O, and Si as explicit elements. The FASTWIND LINES-list and FORMAL_INPUT file is well tested and verified in the wavelength range from 4000 to 7000 Å. In the FORMAL_INPUT file we included H, He I–II, C II–IV, N II–V, O II–V, and Si II–IV in the wavelength range from λ3500 to 10 000. On the basis of the Vienna Atomic Line Database[3] (VALD, Piskunov et al. 1995; Ryabchikova et al. 2015) version 3 and NIST data base[4] (Kramida et al. 2022), we added the following lines to the FASTWIND LINES-list: C II λ6784, C III λ7703 and λ9701–05–16, C IV λ4647 and λ6592–93, N III λ5321–27–52, λ3935–39 and λ9402–24, N IV λ3748, λ5737, λ5776–85, λ6212–15–29, λ7103–09–11–23–27–29, λ7425 and λ9182–223, O III λ3703, λ3707–15, λ3755–57–60–74–91, O IV λ3560–63, λ3729–37, λ7032–54, λ5769 and λ9454–88–92, O V λ5114 and λ6500, and Si II λ9413. A full list of included spectral lines can be found in the Appendix A1.

Some lines are located in the region of telluric bands, but could be of great value, if a careful telluric correction has been performed. With data from 4MOST (de Jong et al. 2019) and in particular 4MOST/1001MC (Cioni et al. 2019), we are going to verify the FASTWIND LINES-list beyond the well-tested wavelength range utilizing the pipeline of this study.

One of the pre-processing steps is to transform observations and models to a common wavelength sampling (Section 3.1). As we use a $\chi^2$-minimization algorithm, the strong Balmer lines with many wavelength points dominate the $\chi^2$. Even though helium lines are weaker than Balmer lines, they are more prominent than metal lines. To balance the weights of the spectral lines, we redefined the common reference wavelength grid for both the observational and synthetic spectra. First, we initialize a wavelength array with the wavelength range and spectral resolution of our observational data. Secondly, based on our compiled FASTWIND LINES-list, we remove every other wavelength point around the Balmer lines. Thirdly, we doubled the number of wavelength points for metal lines.

Continuum wavelength points contain little to no information about the star. If the normalization is not perfect, those points will add noise to the $\chi^2$ values. Unlike cooler stars, OB stars have significantly less spectral lines in the optical and infrared, so that the continuum is usually clearly defined, and normalization is in most case straightforward as long as there are no prominent emission lines. Therefore, we remove all wavelength points beyond ±5 Å to

[2] https://github.com/jbestenlehner/mdi_analysis_pipeline
[3] For example http://vald.astro.uu.se/~vald/php/vald.php.
[4] https://www.nist.gov/pml/atomic-spectra-database

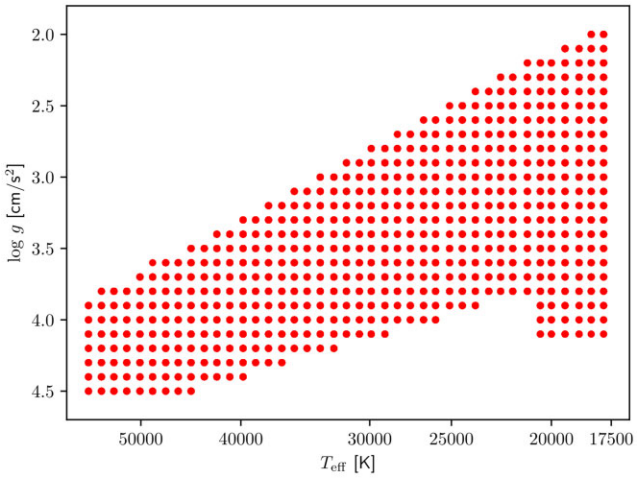


**Figure 2.** $T_{\rm eff}$–log $g$ plane of the computed grid of converged stellar atmospheres with the $x$-axis in log-scale. The high temperature low surface gravity regime (upper left area) is empty as those models are unstable due to the Eddington limit ($\Gamma_{\rm e} \approx 0.7$). Low temperature and high surface gravity models would be better calculated with a plane-parallel code without stellar winds.

the averaged line width of the spectral lines to reduce the contribution of the continuum to the overall $\chi^2$.

The grid covers the parameter space for the effective temperature from $T_{\rm eff} = 17\,800$ K ($\log T_{\rm eff} = 4.25$) to 56 200 K ($\log T_{\rm eff} = 4.75$), surface gravity $\log g/({\rm g\,cm^{-2}}) = 2.0$–4.5, transformed mass-loss rate (e.g. Bestenlehner et al. 2014) from $\log \dot{M}_{\rm t}/({\rm M_\odot\ yr^{-1}}) = -6.5$ to $-5.0$ assuming a constant radius and helium abundances by number from $Y = 0.07$ to 0.15. Three combinations of Carbon, Nitrogen and Oxygen (CNO) abundances representing LMC baseline abundances plus semi- and fully processed CNO composition due to the CNO cycle according to a 60 $\rm M_\odot$ evolutionary track by Brott et al. (2011). Fig. 2 shows the parameter space of the grid with respect to $\log g$ and $T_{\rm eff}$.

The high-temperature, low surface gravity regime (upper left area) is unpopulated as those models are unstable as they exceed the Eddington limit at an Eddington parameter $\Gamma_{\rm e} \approx 0.7$ considering only the electron opacity $\chi_{\rm e}$. Low temperature and high surface gravity models can be computed with FASTWIND, $T_{\rm eff}$ between 17 800 K ($\log T_{\rm eff} = 4.25$) and 21 400 K ($\log T_{\rm eff} = 4.33$) and $\log g > 4.0$, but a significantly larger number of depth points or high mass-loss rates would be required to make them converge. The computational timescale exceeds 1 d in contrast to less than an hour. However, enhanced mass-loss rates are only observed, if the star is in close proximity to of the Eddington limit (Bestenlehner et al. 2014). Therefore, those low-temperature and high surface gravity stellar atmosphere models are better calculated with a plane-parallel geometry code without stellar winds (e.g. TLUSTY; Hubeny & Lanz 1995; Lanz & Hubeny 2007).

The clumping factor was set to $f_{\rm cl} = 1$, i.e. a homogeneous stellar wind is adopted. We assumed a wind acceleration parameter of $\beta = 1.0$ and a fixed micro turbulence velocity of 10 km s$^{-1}$. The terminal velocity was calculated based on $\log g$ and stellar radius of the model using the escape-terminal velocity relation of $v_{\rm esc}/v_\infty = 2.6$ for models hotter than 21 000 K and $v_{\rm esc}/v_\infty = 1.3$ for cooler models (Lamers, Snow & Lindholm 1995). In total, we computed of the order $\lesssim 150\,000$ stellar atmospheres. For around $\sim 20$ per cent of those models the atmospheric structure, ionization balance and/or radiation field failed to converge properly leading to negative fluxes, discontinuities or even failed when the spectral lines were synthesized.

The grid was then convolved with a macro-turbulent velocity of $v_{\rm mac} = 20$ km s$^{-1}$ and varying projected rotational velocity $v \sin i = [0, 20, 50, 100, 150, 200, 250, 300, 400]$ km s$^{-1}$ assuming $v_{\rm eq} \sin i$ is the dominant broadening mechanism, which is a reasonable assumption given the spectral resolution of the observational data (Section 3.4) and that typical $v_{\rm mac}$ are of the order of a few 10s km s$^{-1}$. This results in a grid of $\lesssim 1\,100\,000$ synthetic spectra, which has been used to compute the decomposition matrix to decompose the grid into its principal components reducing the size by a factor $\sim$200. The decomposition matrix is also used to decompose the observational data (Section 3.1).

### 3.4 Observational data

To validate the methodology of the pipeline, we used the VLT/FLAMES data of VLT/FLAMES Tarantula survey (VFTS; Evans et al. 2011) in the traditional blue-optical wavelength regime with the LR02 (λ3964–4567, $\lambda/\Delta\lambda = 6000$), LR03 (λ4501–5078, $\lambda/\Delta\lambda = 7500$) and HR15N (λ6470–6790, $\lambda/\Delta\lambda = 19\,000$) gratings. We selected 240 O-type stars employing the same data and normalization plus radial velocities from (Sana et al. 2013) as used in Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014, 2017), and Ramírez-Agudelo et al. (2017) to avoid the introduction of biases.

The second data set we used is the VLT/MUSE observation of $\sim$250 OB stars plus radial velocities from Castro et al. (2018). The VLT/MUSE data cover the wavelength range between 4600 and 9300 Å at spectral resolution of 2000–4000. The normalization of the spectra was fully automated simulating the work flow of the spectroscopic analysis of large data sets. 35 stars are in common with VFTS, which will allow us to test the reliability of line diagnostics towards the red of H $\alpha$ (Section 3.3) and the automated normalization routine. The normalization is fairly consistent within and between the VFTS and VLT/MUSE sample (Fig. A1).

It is advisable to perform a homogeneous and automated spectra normalization, so that these are chosen consistently for all stars and all wavelength regions are normalized in a similar fashion for every star. We note that the choice of the spectrum normalization has an impact on the model error matrix, e.g. it affects in particular the edge of the detector spectral coverage or spectral regions with overlapping Balmer/Paschen lines and jumps.

## 4 RESULTS AND DISCUSSION

### 4.1 VLT/FLAMES: Evans et al. (2011)

We analysed all 240 O-type stars including stars with low SNR and/or strong nebular contamination (e.g. Fig. A2), while Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014, 2017), and Ramírez-Agudelo et al. (2017) provided reliable results for 173 out of 240 sources. In Fig. 3, we show the spectroscopic fits of a representative early and late O star (VFTS-072 and 076). The shaded error area in Fig. 3 reveals where a general mismatch between model and observations occurs and where is the square root of the diagonal elements of the model-error uncertainty matrix. In particular, the line centres of the Balmer and He I–II lines seemed to be poorly fitted as a result of nebular lines, inaccurately determined line broadening, insufficient grid resolution and range for helium abundances, fixed micro-turbulent velocity, or shape of line profiles (e.g. stellar wind parameters or spectroscopic binaries). We are also able to locate spectral lines, which are potentially not included in the FASTWIND

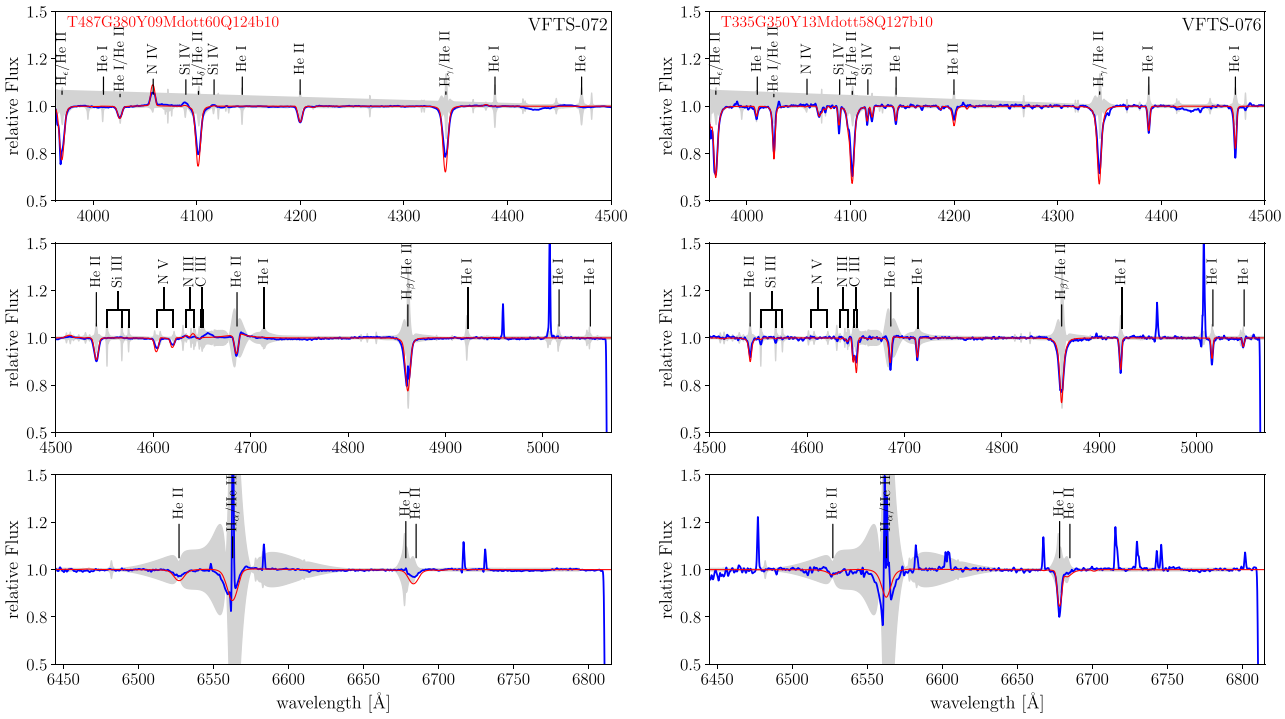


**Figure 3.** Left, spectroscopic fit of an fast rotating early O2 V–III(n)((f∗)) star VFTS-072 and, right, a late O9.2 III star VFTS-076 (right). Blue solid line is the observation, red solid line the synthetic spectrum, and the grey shaded area is the square root of the diagonal elements of the model-error uncertainty matrix calculated by the pipeline.

LINES-list or require improved atomic data (e.g. Si III λ4552.6, 4567.8, and 4574.7). Overall, the spectroscopic fit is good for the synthetic spectra (red solid line) to the observations (blue solid line).

Fig. 1 visualizes the probabilities in the log $g$–$T_{\rm eff}$ plane. VFTS-072 (left panel) shows within $2\sigma$ a dual solution (∼49 000 and ∼46 000 K). At ∼ 45 000 K, the He I lines disappear, but the N IV and V lines sufficiently contributed to the $\chi^2$ so that the correct solutions around 49 000 K have also the highest probabilities. By looking at the $3\sigma$ contour, we notice a degeneracy between log $g$–$T_{\rm eff}$ due to the proximity of VFTS-072 to the Eddington limit. In contrast, the heat map of VFTS-076 (right panel) is well centred on a specific log $g$–$T_{\rm eff}$ region. However, a slightly higher surface gravity up to 0.1–0.2 dex could be probable within $2\sigma$, which is the result of a degeneracy between surface gravity and mass-loss rates. High mass-loss rates fill in the wings of the Balmer lines mimicking a lower surface gravity.

### *4.1.1 Model-error uncertainty matrix*

The model-error uncertainty matrix is symmetric ($U^{\rm T} = U$) and shows correlation between wavelength or pixel regions. An example is given in the appendix (Table A2), where we reduced the rank of the matrix from 11 840 × 11 840 to 37 × 37 for visualization purposes. The strongest correlations are between the Balmer lines, which are the most prominent lines in O-type stars. On the other hand, He I is present in mid to late O stars, while He II lines are only strong in early O stars. Therefore, to visualize the model-error matrix and its correlations, we plot in Fig. 4 the model uncertainties for wavelengths of H $\alpha$, He I λ4471, and He II λ4686.

H $\alpha$ and He II λ4686 are anticorrelated with each other, although we amplified the uncertainties for He II λ4686 by a factor of 10.

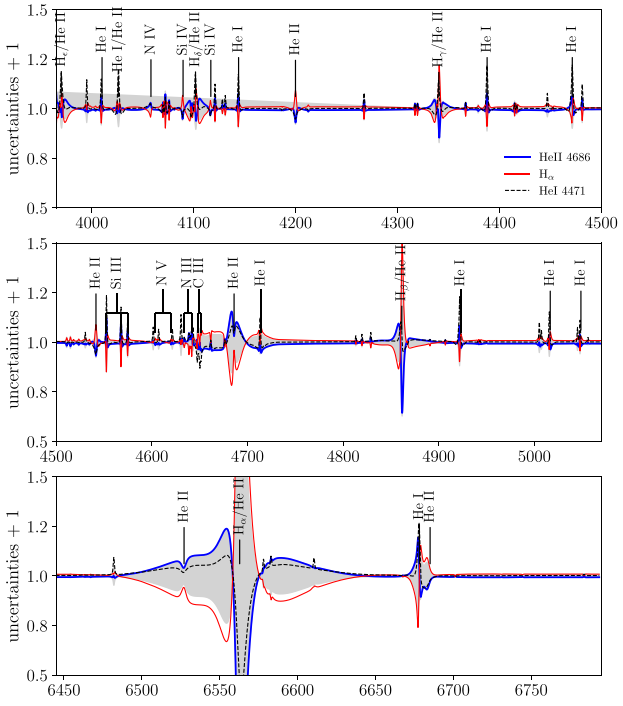


**Figure 4.** Model uncertainties extracted from model-error matrix as a function of wavelength for H $\alpha$ (solid red), He I λ4471 (black dashed), and He II λ4686 (solid blue). For better visualization, uncertainties for He I λ4471 and He II λ4686 are multiplied by a factor 25 and 10, respectively.

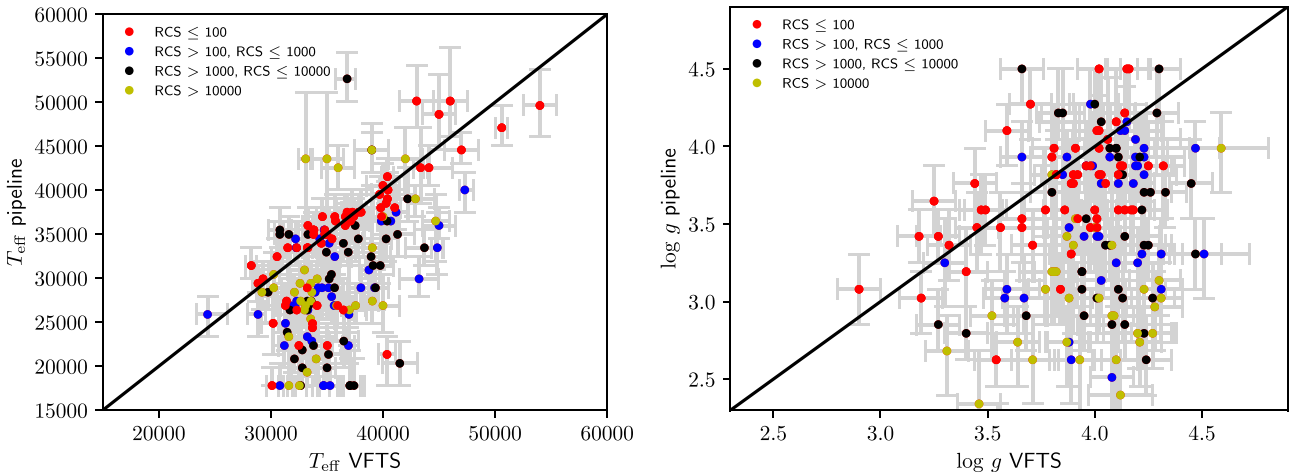


**Figure 5.** Effective temperatures (left) and surface gravities (right) determined by the pipeline versus the results from Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014, 2017), and Ramírez-Agudelo et al. (2017).

Wavelength regions of Balmer lines are a blend of hydrogen and He II lines. With increasing temperature, the He II lines are stronger while the Balmer lines become weaker. In addition, the line strength between hydrogen and helium also determines their abundances, e.g. overall stronger helium lines with respect to hydrogen lines mean lower hydrogen abundances.

He I λ4471 (amplified by a factor of 25) is correlated with He II λ4686 for helium lines, but anticorrelated for H $\delta$ and higher order Balmer lines following the trend of H $\alpha$. He I λ4471 shows stronger correlations with Si III and C III, which are only present in late O stars, where He I lines are strongest as well. Under this supposition, we would expect that we observe a strong correlation between He II and the higher ionized N IV and V, which seems not to be the case. A reason might be that the number of early O stars is too small due to the stellar mass function to significantly contribute to the model error (<5 per cent). Grouping similar objects together is advisable when testing model assumptions in stellar atmosphere codes.

### 4.1.2 *Challenges*

The examples shown in Fig. 3 show low and modest nebular contamination and the pipeline derives result in good agreement with VFTS. However, the pipeline has difficulties when spectra show strong nebular lines. In the case of VFTS-142 (Fig. A2), the temperature is still reasonably well reproduced, but the surface gravity is by ∼0.3 dex too low. If the spectra is dominated by nebular lines, the pipelines will fail, e.g. VFTS 410 (Fig. A3). Often nebular lines are clipped, but in the case of VFTS-410 only few diagnostic lines would remain. Clipping nebular lines could increase the accuracy for log $g$ as the latter is derived from the wings of the Balmer lines due to pressure broadening, but would reduce the capability to derive He abundances or mass-loss rates, which partially depend on the absorption depth and line profile.

Even though we perform a single star analysis, for double-lined spectroscopic binaries (SB2s) the pipeline is able to fairly fit the primary component, but struggles with the mass-loss rate and helium abundances due to the contribution of the colliding wind region of VFTS-527 (Fig. A4; Taylor et al. 2011).

The goodness of fit is usually evaluated by calculating the reduced chi-square (RCS), which uses in our case the diagonal of the error covariance matrix. Due to the nebular contamination and diffuse interstellar bands, none of our fits had an RCS close to 1. To visualize how well the pipeline performs, we compare our results versus tailored analysis of VFTS targets in Fig. 5. Our results agree well with Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014, 2017), and Ramírez-Agudelo et al. (2017) for fits with RCS <100. Effective temperatures show a tighter relation than the surface gravity. The determination of surface gravity is based on the wings of the Balmer lines, which is influenced by the line broadening and therefore how well $\upsilon_{\rm mac}$ and $\upsilon \sin i$ are determined. If the spectroscopic fit is poor, we derive systematically lower temperatures and surface gravities. Low temperature and gravity models have H $\alpha$ in emission to somehow fit the none-stellar H $\alpha$ nebular line while higher order Balmer lines remain in absorption. Overall, there is good agreement considering that our analysis took less than 30 min while the VFTS analysis involved the effort of three PhD theses.

Looking at the error bars, the pipeline obtains systematically larger uncertainties in part due to the inclusion of the model error but mostly as a result of the interpretation of the 4D posterior distribution function (PDF), which includes the degeneracies between $T_{\rm eff}$, log $g$, $\dot{M}_{\rm t}$, and $Y$. The derived errors might be larger, but they are a complete representation of the true uncertainties. A representative prior could increase the accuracy, but might introduce additional biases (Bestenlehner et al. 2020). Temperature uncertainties are systematically larger in the region around 45 000 K as a result of the weakness of He I lines and the ionization balance is not based on He I–II, but on the metal ions N III–IV–V. Nitrogen lines in early O star are relatively weak compared to He lines and therefore contributed little to the global $\chi^2$ without specific weighting of spectral lines. This can lead to an overall lower RCS, but inaccurate temperature determination (red outliers in Fig. 5). A similar behaviour occurs in the transition from late O to early B stars due to the weakness of He II lines, where the main temperature diagnostics are Si III and IV lines. So a careful weighting scheme should be a very promising way for optimizing the pipeline by increasing the accuracy while at the same time reducing the degeneracies between parameters.

A similar plot is shown in Fig. A5 for $\upsilon \sin i$ in the appendix. The projected rotational velocity is only estimated by the pipeline. It consists of minimizing the $\chi^2$ value, thereby finding the best-fitting model while reducing the model error. Stellar parameters can be found in the supplementary online material (Table S1).

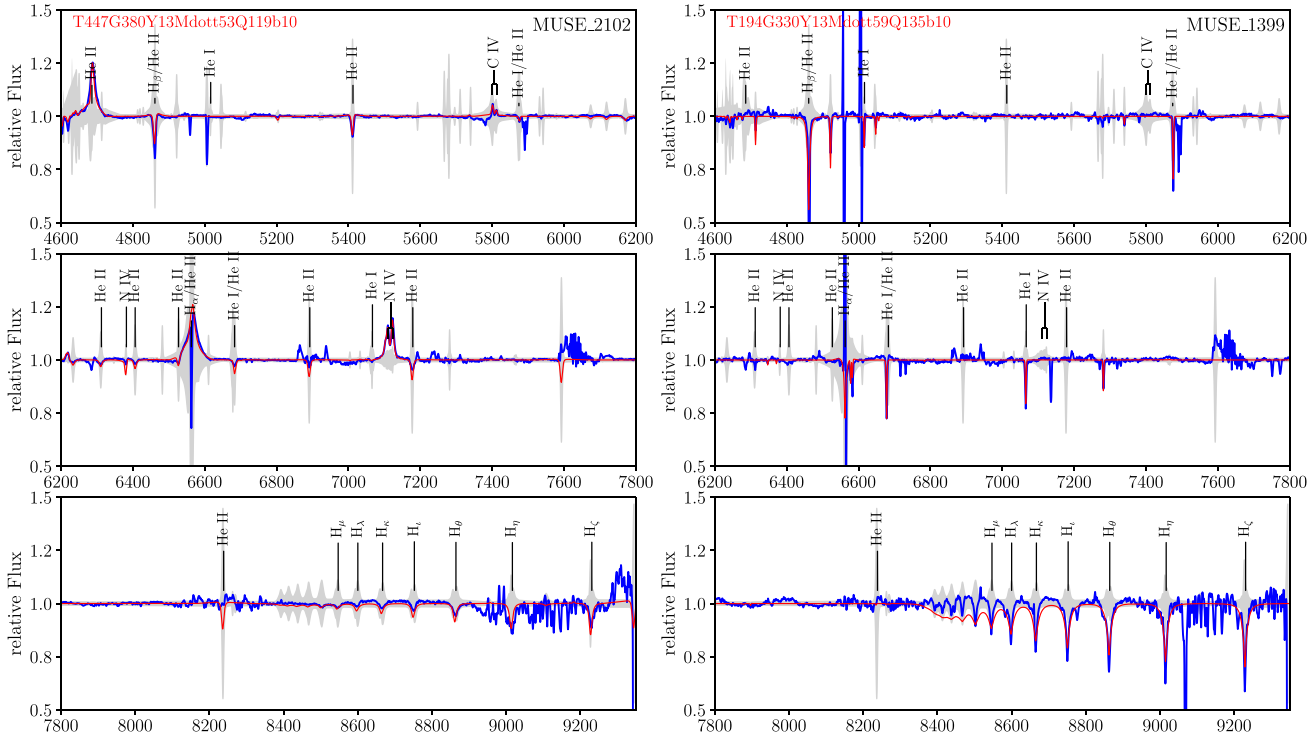


**Figure 6.** Spectroscopic fit of an O2 If (Mk42, left) and a B2 Ib (VFTS-417, right). Blue solid line is the observation, red solid line the synthetic spectrum, and the grey shaded area is the square root of the diagonal elements of the covariant matrix calculated by the pipeline. In the left panel, the newly added N IV mulitplet at λ7103–29 is able to reproduce the observed line.

### 4.2 VLT/MUSE: Castro et al. (2018)

Fig. 6 shows the spectroscopic fit of an of supergiant and a B supergiant with $\Delta T_{\rm eff} \approx 25\,000$ K and $\Delta \log \dot{M} \gtrsim 2.5$ dex. This highlights that stars covering a large spectral-type range can be successfully and reliably analysed with a single pipeline set-up at the same time. However, both stars would not be considered as similar, which has implication on the model error $U$. Such a model error is averaged over a wide parameter space and is probably not very helpful, when testing specific physics (e.g. stellar wind physics or atomic data) in the model. Similar to the VFTS data, the pipeline performs not well for low-signal-to-noise spectra (S/N $\lesssim$ 10–15), spectra with strong nebular lines and spectroscopic binaries/multiples.

In Fig. 7, we compare our results with those from Castro et al. (2021), which is based on the ionization balance of selected He I and He II features and the wings of H$\beta$. In contrast, we used all H, He plus CNO, and Si metal lines available in the VLT/MUSE wavelength range. The left panel compares effective temperatures, which shows a large scatter, but mostly agrees within their large uncertainties. Above 45 000 K, when He I lines disappear or is weakly present in the spectra, the temperature needs to be derive based on the ionization balance of metal lines. In the wavelength range of MUSE, we have C III–IV and N III–IV. While C IV and N IV are located in relatively clean areas of the spectra, the C III and N III are often found in the range of telluric bands or near the Paschen jump, where we have issues with the normalization (Fig. 6). B-type stars have as per definition no He II present in their spectra and the temperature is based in the case of early B stars on the ionization balance of Si III–IV lines. There is a reasonable number of lines in the MUSE range, but the temperature determination suffers with the presence of nebular lines or low-SNR spectra.

The right panel of Fig. 7 compares surface gravities that show an even larger scatter and uncertainties. Even though we utilized the Paschen lines as well, there are two potential caveats: first, the normalization near the Paschen jump is not straightforward as the lines overlap and therefore no continuum, and secondly, the degree of overlapping depends not only on $\log g$ but also on the line broadening due to the narrowness of the higher order Paschen lines, which is only approximately determined during the fitting process. This results in a degeneracy between $\log g$ and $v_{\rm eq}\ \sin i$. Surface gravities cluster in the range of $\log g = 3.5$–$4.5$, which is expected for a young stellar population of 1–2 Myr largely consisting of dwarfs and giants in the proximity of R136 (Bestenlehner et al. 2020).

To better quantify how reliable the analysis based on the VLT/MUSE data is, we compare our results to VFTS. 35 VLT/MUSE targets are in common with VFTS (Bestenlehner et al. 2014; Sabín-Sanjulián et al. 2014; Ramírez-Agudelo et al. 2017; Sabín-Sanjulián et al. 2017) and the comparison is shown in Fig. 8. Uncertainties are systematically larger. Effective temperatures agree within their 1$\sigma$ uncertainties for 26 out 35 stars (left panel) with differences largely as a result of the cleanliness and quality of spectra. Surface gravities are in agreement for only half of the sample, which is expected due to the challenges of the Paschen lines. Overall, the agreement is reasonable considering the difficulties in the analysing of the MUSE data.

To place the stars into the Hertzprung–Russell diagram (HRD), we derived bolometric luminosity on the basis of optical *UBV* from Selman et al. (1999) and near-infrared $JK_{\rm s}$ photometry from 2 Micron All Sky Survey and Vista Magellanic Clouds survey (Cioni et al. 2011) using the methodology of Bestenlehner et al. (2020, 2022). The HRD (Fig. 9) shows that most stars are populated near and to

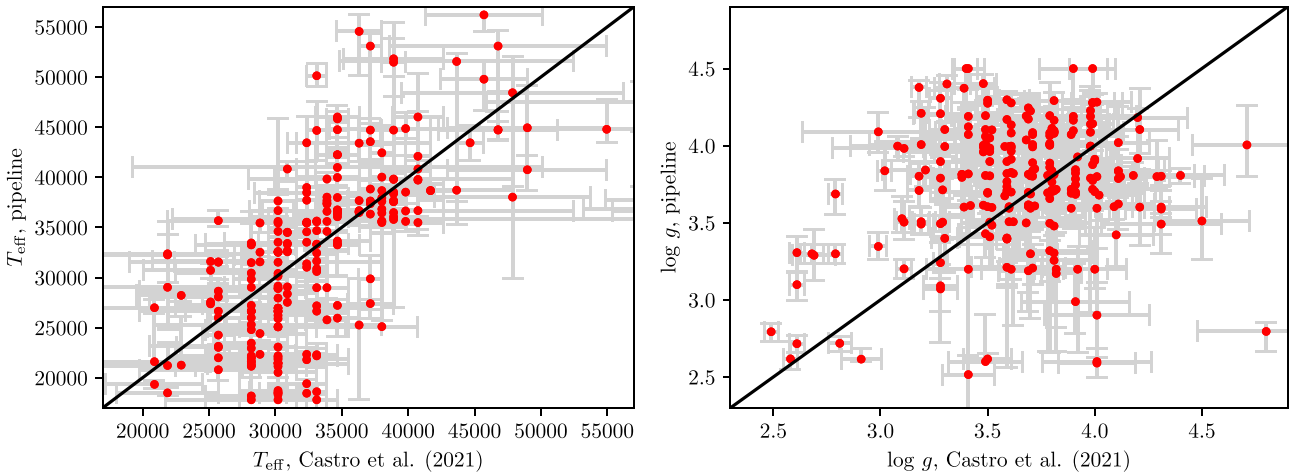


**Figure 7.** Effective temperatures (left) and surface gravities (right) determined by the pipeline versus the results from Castro et al. (2021).

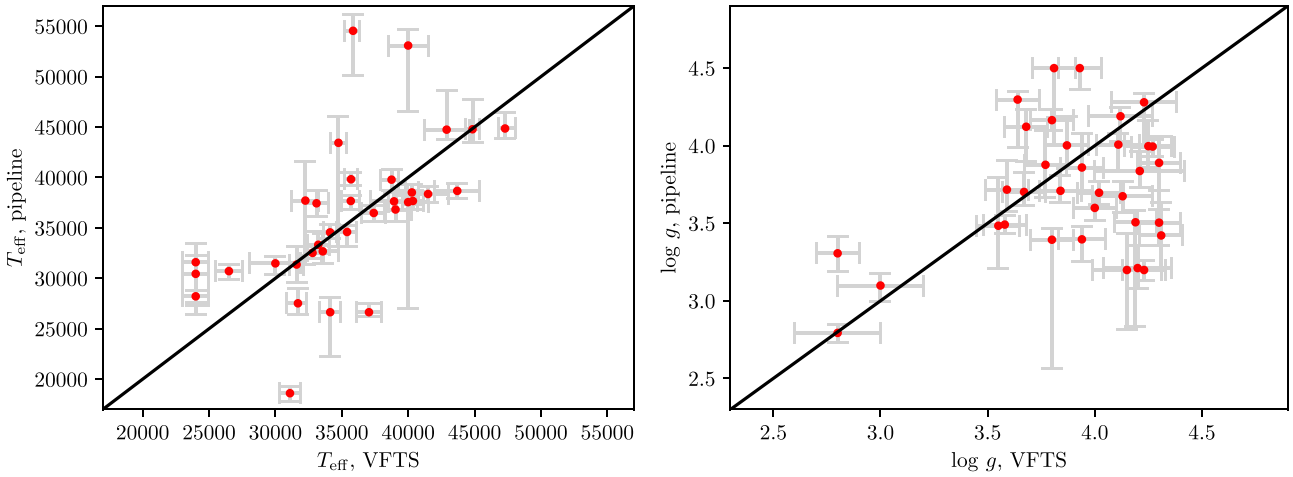


**Figure 8.** Comparison of MUSE targets in common with VFTS: effective temperatures (left) and surface gravities (right) determined by the pipeline using VLT/MUSE data versus the results from Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014, 2017), McEvoy et al. (2015), and Ramírez-Agudelo et al. (2017) using VLT/FLAMES data.

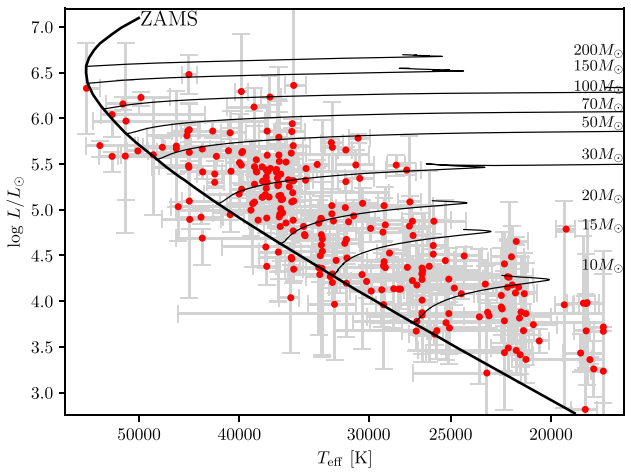


**Figure 9.** Hertzsprung–Russell diagram of the analysed stars using the VLT/MUSE data from Castro et al. (2018). Thin black lines are stellar evolutionary tracks by Brott et al. (2011) and Köhler et al. (2015).

the cool side of the zero-age main sequence (ZAMS). There are a couple of exceptions but their uncertainties do not exclude a cooler location in agreement with the majority of the sources. This can be improved by including a meaningful prior into the analysis, e.g. based on evolutionary tracks, and could increase the accuracy of the results, as we used only a flat prior ($\omega_{ii'} = 1$). Only hydrogen-deficient stars are found to be on the hot side of the ZAMS, e.g. self-stripping through stellar winds or binary evolution. A prior would give the star a higher probability to be found either on the hot or on the cool side of the ZAMS depending on its helium composition. Stellar parameters can be found in supplementary online material (Table S2) and individual spectroscopic fits for visual inspection are also in the supplementary online material. Mass-loss rates, and He, C, N, and O abundances are not included. Optical data can only provide an upper limit for most stars in our sample (flat PDF towards lower mass-loss rates) while the PDF for helium is cut-off at the primordial abundance of 25 per cent by mass. CNO abundances are too coarsely sampled and linked to the predicted chemical evolution of 60 $M_\odot$ star (Section 3.3).

### 4.3 Alternative spectroscopic pipeline for large data sets

Xiang et al. (2022, THE HOTPAYNE) applied the FGK methodology of THE PAYNE (Ting et al. 2019) to OBA stars to derive two stellar labels/parameters ($T_{\rm eff}$ and $\log g$) plus 15 chemical abundances.They used the plane-parallel LTE atmospheric models calculated with ATLAS12 (Kurucz 1970, 1993, 2005) and were able to analyse 330 000 spectra. While $T_{\rm eff}$ and $\log g$ are sensible for A and mid–late B dwarfs, the derived chemical abundances suffer from non-negligible systematics due to non-LTE effects. AB supergiants require spherical geometry (stellar radius $\gg$ scale height) including stellar winds as these effects cannot be neglected. In hotter and more luminous stars (early B and O stars), a non-LTE treatment is necessary (Mihalas & Auer 1970) such that THE HOTPAYNE results for $T_{\rm eff}$ and $\log g$ are not reliable including stars with weak winds. Even the inclusion of a model error will not improve much due to the fundamental missing physics in the ATLAS12 models.

To make THE HOTPAYNE usable for OBA stars, the underlying stellar models must be replaced with models computed with more sophisticated and fully non-LTE stellar atmosphere codes designed for hot, massive stars with stellar winds, e.g. CMFGEN (Hillier & Miller 1998), FASTWIND (Santolaya-Rey, Puls & Herrero 1997; Puls et al. 2005), or PoWR (Hamann & Gräfener 2003; Gräfener et al. 2012). The approach of the THE HOTPAYNE needs to be expanded: to include the model uncertainties into the error budget (e.g. this work) to account for the assumptions and parametrizations utilized in those complex stellar atmosphere codes. Additional stellar labels/parameters need to be incorporated, such as mass-loss rate, velocity law, wind inhomogeneity, and terminal and wind-turbulent velocity plus helium abundances. The helium abundance increases due to the CNO cycle, which in turn increases the mean molecular weight ($\mu$), and impacts the mass–luminosity relation ($L \propto \mu^4 M^3$), electron density, and therefore the structure and ionization balance of the stellar atmosphere. When analysing optical spectra, the wind parameters can be merged into a wind strength parameter $Q$ (Puls et al. 1996), transformed radius $R_{\rm t}$ (Schmutz, Hamann & Wessolowski 1989; Gräfener et al. 2002; Hamann & Gräfener 2004), or transformed mass-loss rate $\dot{M}_{\rm t}$ (Bestenlehner et al. 2014, Section 3.3).

## 5 CONCLUSIONS AND OUTLOOK

Large spectroscopic surveys with WEAVE and 4MOST will observe tens of thousands of massive stars, which need to be analysed in a homogeneous and efficient way. The pipeline presented in this work takes advantage and utilizes the information that large data sets provide by determining the model uncertainties, which are included into the error budget. This methodology could also be applied to galaxies or other domains like biology and geophysics for which approximate and incomplete theoretical models exist as well.

The runtime of the pipeline scales exponentially with the number of spectra, because all stars are simultaneously analysed and the error-model uncertainty matrix is iteratively updated (Section 2.4). However, once a converged error-model uncertainty matrix is obtained, we can limit the matrix operations to the $\chi^2$-minimization and switch to a star by star analysis. In this case, we are able to analyse one star in less than a second.

The fully automated spectroscopic analysis tool of this work reduces the human interaction to a minimum to cope with the amount of data. It is able to process ∼250 stars in less than half an hour (∼6 CPU hours) delivering results comparable to Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014), Ramírez-Agudelo et al. (2017), and Sabín-Sanjulián et al. (2017) over a decade. Overall, the quality of the spectroscopic fits is good, but around 15 per cent of the stars need additional attention as a result of strong nebular contamination, low S/N, multiplicity, etc. The pipeline performs well over a wide parameter space that is supported by the spectroscopic analysis of three benchmark stars by several groups within the X-Shooter and ULLYSES collaboration (XShootU; Vink et al. 2023) of optical VLT/X-Shooter data (Sander et al., in preparation).

Weights of spectral lines could increase the accuracy, but need to be adjusted depending on the parameter space that would then require human interaction. Determining weights for features (spectral lines) is a typical machine learning problem and often solved with neural networks (deep learning). However, to really take advantage of our statistical approach and optimize the pipeline, we will require much larger data sets, which will be soon provided by WEAVE and 4MOST. Future advances of our pipeline will be released on the pipeline repository.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for detailed and helpful comments that improved the clarity and content of the manuscript. JMB and PAC were supported by the Science and Technology Facilities Council research grant ST/V000853/1 (PI: V. Dhillon). MB was supported through the Lise Meitner grant from the Max Planck Society. We acknowledge support by the Collaborative Research Centre SFB 881 (projects A5 and A10), Heidelberg University, of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 949173). This work has made use of the VALD database, operated at Uppsala University, the Institute of Astronomy RAS in Moscow, and the University of Vienna and the Atomic Spectra Databases at the National Institute of Standards and Technology in Gaithersburg.

## DATA AVAILABILITY

The data underlying this article are available in the article and in its supplementary material. The spectroscopic analysis pipeline is available at https://github.com/jbestenlehner/mdi_analysis_pipeline. Spectroscopic data are available via the ESO archive facility, while grids of synthetic spectra can be requested from the lead author.

## REFERENCES

Allende-Prieto C., Apogee Team, 2015, American Astronomical Society Meeting Abstracts #225, p. 422.07
Bensby T. et al., 2019, The Messenger, 175, 35
Bestenlehner J. M. et al., 2014, A&A, 570, A38
Bestenlehner J. M. et al., 2020, MNRAS, 499, 1918
Bestenlehner J. M., Crowther P. A., Broos P. S., Pollock A. M. T., Townsley L. K., 2022, MNRAS, 510, 6133
Brands S. A. et al., 2022, A&A, 663, A36
Brott I. et al., 2011, A&A, 530, A115
Castro N. et al., 2012, A&A, 542, A79
Castro N. et al., 2021, A&A, 648, A65
Castro N., Crowther P. A., Evans C. J., Mackey J., Castro-Rodriguez N., Vink J. S., Melnick J., Selman F., 2018, A&A, 614, A147
Caticha A., 2008, preprint (arXiv:0808.0012)
Chiappini C. et al., 2019, The Messenger, 175, 30
Cioni M. et al., 2011, A&A, 527, A116
Cioni M. R. L. et al., 2019, The Messenger, 175, 54

de Jong R. S. et al., 2019, The Messenger, 175, 3
Enßlin T. A., Frommert M., 2011, Phys. Rev. D, 83, 105014
Enßlin T. A., Weig C., 2010, Phys. Rev. E, 82, 051112
Evans C. J. et al., 2011, A&A, 530, A108
Gräfener G., Koesterke L., Hamann W. R., 2002, A&A, 387, 244
Gräfener G., Vink J. S., Harries T. J., Langer N., 2012, A&A, 547, A83
Guiglion G. et al., 2020, A&A, 644, A168
Hamann W.-R., Gräfener G., 2003, A&A, 410, 993
Hamann W.-R., Gräfener G., 2004, A&A, 427, 697
Hillier D. J., Miller D. L., 1998, ApJ, 496, 407
Hubeny I., Lanz T., 1995, ApJ, 439, 875
Jasche J., Kitaura F. S., Wandelt B. D., Enßlin T. A., 2010, MNRAS, 406, 60
Jaynes E. T., Bretthorst G. L., 2003, Probability Theory, Cambridge University Press, Cambridge, UK
Jin S. et al., 2023, MNRAS
Köhler K. et al., 2015, A&A, 573, A71
Kramida A., Ralchenko Yu., Reader J., NIST ASD Team, 2022, NIST Atomic Spectra Database (ver. 5.10), National Institute of Standards and Technology, Gaithersburg, MD USA
Kurucz R. L., 1970, SAO Special Report, 309
Kurucz R. L., 1993, SYNTHE spectrum synthesis programs and line data
Kurucz R. L., 2005, Mem. Soc. Astron. Ital. Suppl., 8, 14
Lamers H. J. G. L. M., Snow T. P., Lindholm D. M., 1995, ApJ, 455, 269
Lanz T., Hubeny I., 2007, ApJS, 169, 83
McEvoy C. M. et al., 2015, A&A, 575, A70
Mihalas D., Auer L. H., 1970, ApJ, 160, 1161
Mokiem M. R. et al., 2007, A&A, 465, 1003
Ness M., Hogg D. W., Rix H. W., Ho A. Y. Q., Zasowski G., 2015, ApJ, 808, 16
Oberpriller J., Enßlin T. A., 2018, preprint (arXiv:1812.08194)
Piskunov N. E. et al. 1995, Astronomy and Astrophysics Supplement Series, Astronomy and Astrophysics Supplement Series, 112, p.525
Puls J. et al., 1996, A&A, 305, 171
Puls J., Urbaneja M. A., Venero R., Repolust T., Springmann U., Jokuthy A., Mokiem M. R., 2005, A&A, 435, 669
Ramírez-Agudelo O. H. et al., 2017, A&A, 600, A81
Rivero González J. G., Puls J., Najarro F., Brott I., 2012, A&A, 537, A79
Ryabchikova T. et al., 2015, A major upgrade of the VALD database, Physica Scripta, 90, 054005
Sabín-Sanjulián C. et al., 2014, A&A, 564, A39
Sabín-Sanjulián C. et al., 2017, A&A, 601, A79
Sana H. et al., 2013, A&A, 550, A107
Santolaya-Rey A. E., Puls J., Herrero A., 1997, A&A, 323, 488
Schmutz W., Hamann W., Wessolowski U., 1989, A&A, 210, 236
Selman F., Melnick J., Bosch G., Terlevich R., 1999, A&A, 341, 98
Simón-Díaz S., Castro N., Herrero A., Puls J., Garcia M., Sabín-Sanjulián C., 2011, in Journal of Physics Conference Series. p. 012021
Taylor W. D. et al., 2011, A&A, 530, L10
Ting Y.-S., Conroy C., Rix H.-W., Cargile P., 2019, ApJ, 879, 69
Vink J. S. et al., 2023, A&A, 675, A154
Wandelt B. D., Larson D. L., Lakshminarayanan A., 2004, Phys. Rev. D, 70, 083511
Wang B., Shi W., Miao Z., 2015, PLoS ONE, 10, e0118537
Xiang M. et al., 2022, A&A, 662, A66

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**MN-23-2976-MJ_som.pdf**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: ADDITIONAL MATERIAL

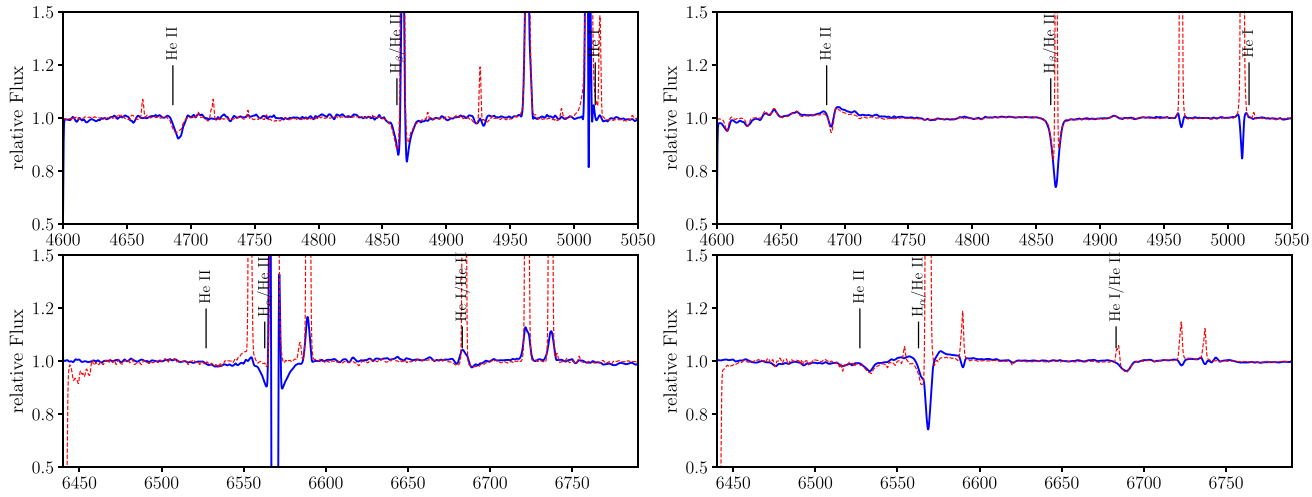


**Figure A1.** Comparison of the normalized MUSE (solid blue line) and VFTS (dashed red line): The normalization is fairly consistent between both sample. As VLT/MUSE is an integral field unit, Castro et al. (2018) was able to reasonable well remove nebular lines in some cases, but in some instance nebular lines where over corrected or still present.

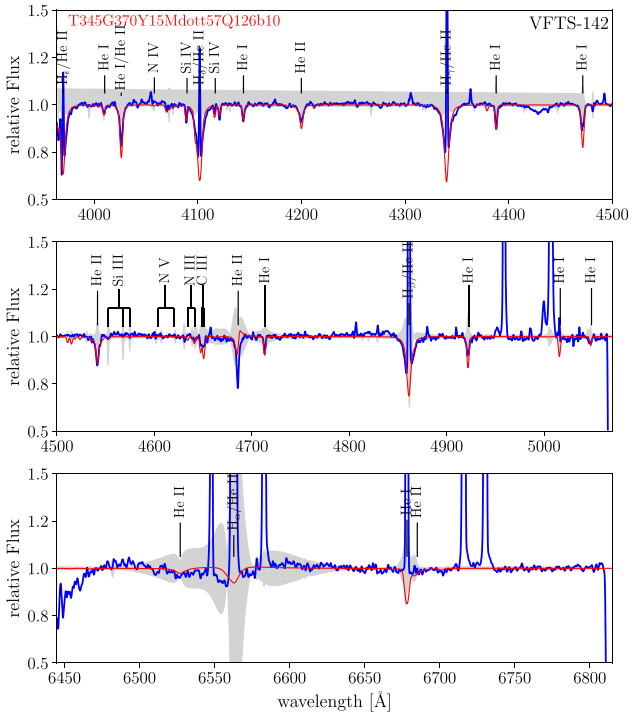


**Figure A2.** VFTS 142: moderate nebular contamination. Blue solid line is the observation, red solid line the synthetic spectrum, and the grey shaded area is the square root of the diagonal elements of the covariant matrix calculated by the pipeline. Effective temperature is well reproduce while the surface gravity is 0.2–0.3 dex too low.

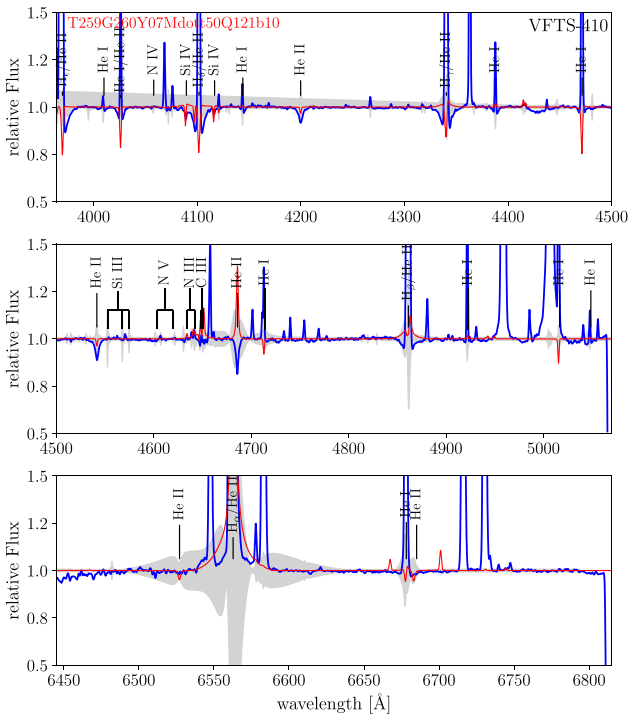


**Figure A3.** VFTS 410: strong nebular contamination. Blue solid line is the observation, red solid line the synthetic spectrum, and the grey shaded area is the square root of the diagonal elements of the covariant matrix calculated by the pipeline. The pipeline is unable to reproduce the stellar parameters.
bottom

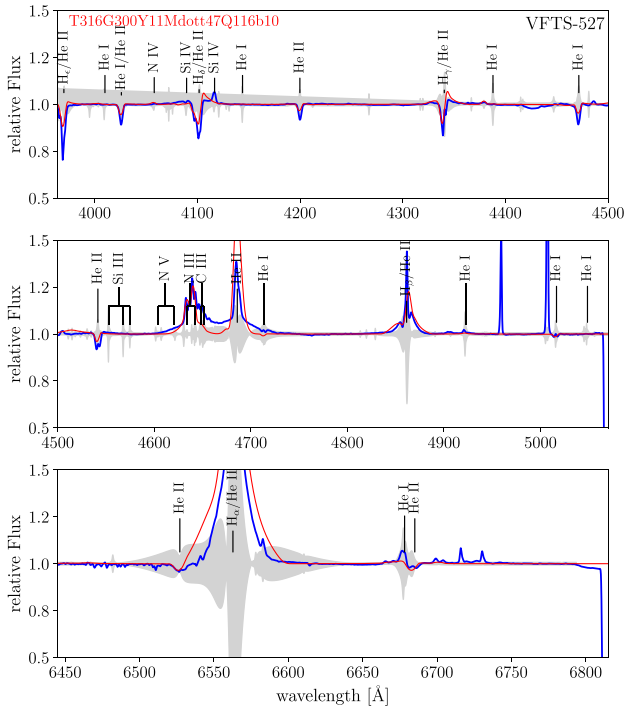


**Figure A4.** VFTS 527: double-lined spectroscopic binary VFTS-527 (Taylor et al. 2011). Blue solid line is the observation, red solid line the synthetic spectrum, and the grey shaded area is the square root of the diagonal elements of the covariant matrix calculated by the pipeline. Effective temperature and surface gravity of the primary are reproduced, but mass-loss rate is too high while helium abundance is too low due to the contribution of the colliding wind region largely effecting H$\alpha$ and H$\beta$.

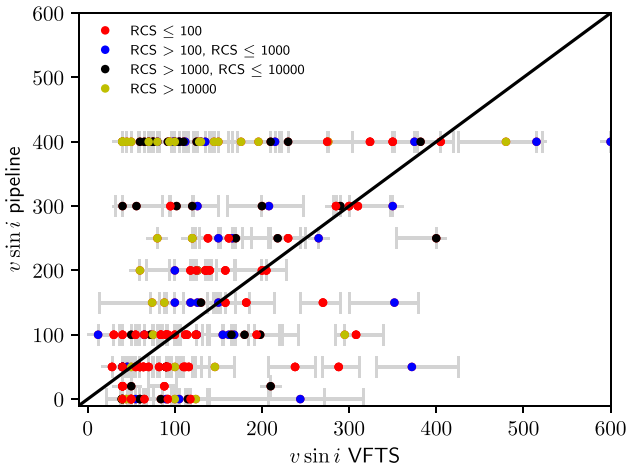


**Figure A5.** Projected rotational velocity estimated by the pipeline versus the results from Bestenlehner et al. (2014), Sabín-Sanjulián et al. (2014, 2017), and Ramírez-Agudelo et al. (2017) plus colour coded RCS. The grid has been pre-convolved with the following $v \sin i$ = [0, 20, 50, 100, 150, 200, 250, 300, 400] km s$^{-1}$, which does not covered the entire observed range up to 600 km s$^{-1}$ and is too coarse for an accurate determination. However, it contributes to reduce the $\chi^2$ and the model error.

**Table A1.** List of lines synthesized when calculating the formal integral. Wavelength ranges are multiplets with diverging central wavelengths.

| Ion | Wavelength (Å) | Ion | Wavelength (Å) | Ion | Wavelength (Å) | Ion | Wavelength (Å) | Ion | Wavelength (Å) | Ion | Wavelength (Å) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| H I | 3835.4 | He I | 3888.6 | C II | 3919.0–3920.7 | N II | 3995.9 | O II | 3945.0–3954.4 | Si II | 3853.7–3862.6 |
| H I | 3889.1 | He I | 3964.7 | C II | 4267.0–4267.3 | N II | 4447.0 | O II | 4069.6–4075.9 | Si II | 4128.1–4130.9 |
| H I | 3970.1 | He I | 4009.3 | C II | 4637.6–4639.1 | N II | 4530.4 | O II | 4317.1–4366.9 | Si II | 5041.0–5056.3 |
| H I | 4101.7 | He I | 4026.2 | C II | 5132.9–5151.1 | N II | 4552.5 | O II | 4414.9–4452.4 | Si II | 5957.6–5978.9 |
| H I | 4340.5 | He I | 4120.8 | C II | 5648.1–5662.5 | N II | 4601.5–4643.1 | O II | 4638.9–4676.2 | Si II | 6347.1–6371.4 |
| H I | 4861.4 | He I | 4143.8 | C II | 6151.3–6151.5 | N II | 5005.2 | O II | 4890.9–4906.8 | Si II | 9412.7–9412.8 |
| H I | 6562.8 | He I | 4387.9 | C II | 6461.9 | N II | 5007.3 | O II | 4941.1–4943.0 | Si III | 3791.4–3806.8 |
| H I | 8392.2 | He I | 4471.5 | C II | 6578.1–6582.9 | N II | 5045.1 | O III | 3703.4 | Si III | 4552.6–4574.8 |
| H I | 8413.1 | He I | 4713.1 | C II | 6783.9 | N II | 5666.6–5710.8 | O III | 3707.3–3715.1 | Si III | 4716.7 |
| H I | 8437.8 | He I | 4921.9 | C III | 4056.1 | N II | 5931.9–5941.7 | O III | 3754.7–3791.3 | Si III | 4813.3–4829.0 |
| H I | 8467.0 | He I | 5015.7 | C III | 4068.9–4070.3 | N II | 6482.1 | O III | 3961.6 | Si III | 5739.7 |
| H I | 8502.3 | He I | 5047.7 | C III | 4152.5–4162.9 | N II | 6610.6 | O III | 4072.6–4089.3 | Si III | 7461.9–7466.3 |
| H I | 8545.2 | He I | 5875.6 | C III | 4186.9 | N III | 3934.5–3938.5 | O III | 4366.5–4375.9 | Si III | 8262.6–8271.9 |
| H I | 8598.2 | He I | 6678.2 | C III | 4647.4–4651.5 | N III | 3998.6–4003.6 | O III | 4799.8 | Si III | 9799.9 |
| H I | 8664.8 | He I | 7065.2 | C III | 4663.6–4665.9 | N III | 4097.4–4103.4 | O III | 5268.3 | Si IV | 4088.9–4116.1 |
| H I | 8750.3 | He I | 7281.4 | C III | 5249.1 | N III | 4195.8–4200.1 | O III | 5508.2 | Si IV | 4212.4 |
| H I | 8862.6 | He II | 3796.3 | C III | 5253.6–5272.5 | N III | 4332.9–4345.7 | O III | 5592.3 | Si IV | 4950.1 |
| H I | 9014.7 | He II | 3813.5 | C III | 5695.9 | N III | 4379.1 | O III | 7711.0 | Si IV | 6667.6–6701.3 |
| H I | 9228.8 | He II | 3833.8 | C III | 5826.4 | N III | 4510.9–4547.3 | O IV | 3560.4–3563.3 | Si IV | 7047.9–7068.4 |
| H I | 9545.7 | He II | 3858.1 | C III | 6731.0–6744.3 | N III | 4527.9–4546.3 | O IV | 3729.0–3736.9 | Si IV | 8957.3 |
| | | He II | 3887.5 | C III | 7707.4 | N III | 4634.1–4641.9 | O IV | 3995.1 | Si IV | 9018.1 |
| | | He II | 3923.5 | C III | 8500.3 | N III | 4858.7–4867.2 | O IV | 4654.1 | | |
| | | He II | 3968.4 | C III | 9701.1–9715.1 | N III | 5320.8–5352.5 | O IV | 4813.2 | | |
| | | He II | 4025.6 | C IV | 4646.6–4647.0 | N III | 6445.3–6487.8 | O IV | 7004.1 | | |
| | | He II | 4100.1 | C IV | 5016.6–5018.4 | N III | 9402.5–9424.5 | O IV | 7032.3–7053.6 | | |
| | | He II | 4199.8 | C IV | 5801.3–5812.0 | N IV | 3747.5 | O IV | 9453.9– 9492.4 | | |
| | | He II | 4338.7 | C IV | 6591.5–6592.6 | N IV | 4057.8 | O V | 5114.1 | | |
| | | He II | 4541.6 | | | N IV | 5200.4–5205.1 | O V | 6500.2 | | |
| | | He II | 4685.7 | | | N IV | 5736.9 | | | | |
| | | He II | 4859.3 | | | N IV | 5776.3–5784.8 | | | | |
| | | He II | 5411.5 | | | N IV | 6212.4–6219.9 | | | | |
| | | He II | 6074.2 | | | N IV | 7103.2–7129.2 | | | | |
| | | He II | 6118.3 | | | N IV | 7425.3 | | | | |
| | | He II | 6170.7 | | | N IV | 9182.2–9223.0 | | | | |
| | | He II | 6233.8 | | | N V | 4603.7–4620.0 | | | | |
| | | He II | 6310.9 | | | | | | | | |
| | | He II | 6406.4 | | | | | | | | |
| | | He II | 6527.1 | | | | | | | | |
| | | He II | 6560.1 | | | | | | | | |
| | | He II | 6683.2 | | | | | | | | |
| | | He II | 6890.9 | | | | | | | | |
| | | He II | 7177.5 | | | | | | | | |
| | | He II | 7592.8 | | | | | | | | |
| | | He II | 8236.8 | | | | | | | | |
| | | He II | 9011.2 | | | | | | | | |
| | | He II | 9108.6 | | | | | | | | |
| | | He II | 9225.3 | | | | | | | | |
| | | He II | 9344.9 | | | | | | | | |
| | | He II | 9367.1 | | | | | | | | |
| | | He II | 9542.1 | | | | | | | | |

**Table A2.** Error-model uncertainty matrix based on the VFTS analysis (Section 4.1). The rank of the matrix has been reduced by merging elements together for visualization purposes. The Balmer lines showing the strongest correlations have been highlighted.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | H $\delta$ | | | | | | | | H $\gamma$ | | | | | | | | | | | | | | H $\beta$ | | | | | | | | | | H$_\alpha$ | |
| λ (Å) | 3987 | 4016 | 4045 | 4072 | 4089 | 4115 | 4136 | 4160 | 4192 | 4233 | 4310 | 4325 | 4346 | 4378 | 4424 | 4474 | 4511 | 4532 | 4559 | 4613 | 4638 | 4651 | 4669 | 4717 | 4792 | 4824 | 4875 | 4914 | 4987 | 5008 | 5025 | 6447 | 6472 | 6520 | 6582 | 6643 | 6728 |
| 3987 | **0.03** | 0.02 | 0.01 | −0.0 | 0.01 | 0.02 | 0.02 | 0.0 | −0.01 | 0.0 | 0.01 | 0.0 | 0.01 | 0.01 | 0.0 | 0.03 | 0.0 | −0.01 | 0.02 | 0.02 | 0.05 | −0.07 | −0.02 | 0.01 | 0.0 | 0.02 | 0.0 | 0.01 | 0.0 | 0.03 | 0.0 | 0.01 | 0.01 | 0.06 | −0.13 | 0.03 | −0.0 |
| 4016 | 0.02 | **0.02** | 0.01 | −0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | −0.0 | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | −0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.04 | 0.01 | 0.01 | 0.03 | 0.02 | 0.02 | 0.02 | −0.04 | 0.17 | 0.0 | 0.01 |
| 4045 | 0.01 | 0.01 | **0.01** | −0.01 | 0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.0 | −0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.02 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | −0.03 | 0.13 | −0.0 | 0.01 |
| 4072 | −0.0 | −0.0 | −0.01 | **0.03** | −0.03 | −0.01 | 0.0 | −0.0 | −0.01 | −0.0 | 0.0 | −0.0 | −0.01 | −0.0 | 0.01 | 0.0 | −0.01 | −0.02 | 0.01 | 0.0 | 0.0 | −0.03 | −0.0 | −0.01 | −0.0 | −0.0 | 0.02 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | −0.0 | −0.01 | 0.1 | −0.01 | −0.0 |
| 4089 | 0.01 | 0.0 | 0.0 | −0.03 | **0.03** | 0.02 | −0.0 | −0.0 | 0.0 | −0.0 | −0.01 | 0.0 | 0.01 | 0.0 | −0.01 | 0.0 | 0.01 | 0.01 | −0.01 | −0.01 | 0.0 | 0.01 | −0.01 | 0.0 | −0.0 | 0.01 | −0.04 | −0.0 | −0.01 | −0.01 | −0.01 | −0.01 | 0.0 | 0.05 | −0.25 | 0.03 | −0.01 |
| 4115 | 0.02 | 0.01 | 0.0 | −0.01 | 0.02 | **0.02** | 0.01 | −0.01 | −0.01 | −0.01 | −0.0 | 0.01 | 0.01 | 0.0 | −0.01 | 0.01 | −0.0 | −0.01 | 0.01 | −0.0 | 0.03 | −0.06 | −0.03 | 0.0 | −0.01 | 0.03 | −0.04 | 0.0 | −0.01 | −0.0 | −0.01 | −0.01 | 0.0 | 0.1 | −0.34 | 0.04 | −0.01 |
| 4136 | 0.02 | 0.01 | 0.01 | 0.0 | −0.0 | 0.01 | **0.02** | 0.0 | −0.01 | 0.0 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.02 | 0.0 | −0.0 | 0.02 | 0.01 | 0.03 | −0.05 | −0.01 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.0 | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 | −0.01 | 0.02 | 0.0 |
| 4160 | 0.0 | 0.01 | 0.01 | −0.0 | −0.0 | −0.01 | 0.0 | **0.02** | 0.02 | 0.02 | 0.01 | −0.01 | −0.0 | 0.01 | 0.02 | 0.0 | 0.02 | 0.02 | 0.0 | 0.01 | −0.01 | 0.04 | 0.03 | 0.01 | 0.01 | −0.03 | 0.05 | 0.01 | 0.02 | 0.03 | 0.02 | 0.02 | 0.01 | −0.1 | 0.35 | −0.02 | 0.02 |
| 4192 | −0.01 | 0.01 | 0.01 | −0.01 | 0.0 | −0.01 | −0.01 | 0.02 | **0.02** | 0.02 | 0.0 | −0.01 | −0.0 | 0.01 | 0.01 | −0.01 | 0.02 | 0.03 | −0.01 | 0.01 | −0.03 | 0.08 | 0.04 | 0.01 | 0.01 | −0.03 | 0.05 | 0.0 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | −0.11 | 0.36 | −0.03 | 0.02 |
| 4233 | 0.0 | 0.01 | 0.01 | −0.0 | −0.0 | −0.01 | 0.0 | 0.02 | 0.02 | **0.02** | 0.01 | −0.01 | −0.0 | 0.01 | 0.02 | 0.0 | 0.02 | 0.02 | 0.0 | 0.01 | −0.01 | 0.04 | 0.03 | 0.01 | 0.02 | −0.03 | 0.06 | 0.01 | 0.02 | 0.03 | 0.02 | 0.02 | 0.01 | −0.1 | 0.37 | −0.02 | 0.02 |
| 4310 | 0.01 | 0.01 | 0.0 | 0.0 | −0.01 | −0.0 | 0.01 | 0.01 | 0.0 | 0.01 | **0.0** | −0.0 | −0.0 | 0.01 | 0.01 | 0.01 | 0.0 | 0.0 | 0.01 | 0.01 | 0.01 | −0.0 | 0.01 | 0.01 | 0.01 | −0.01 | 0.03 | 0.0 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | −0.03 | 0.14 | −0.01 | 0.01 |
| 4325 | 0.0 | −0.0 | −0.0 | −0.0 | 0.0 | 0.01 | 0.0 | −0.01 | −0.01 | −0.01 | −0.0 | **0.01** | 0.0 | −0.0 | −0.01 | 0.0 | −0.01 | −0.01 | 0.0 | −0.01 | 0.01 | −0.04 | −0.02 | −0.01 | −0.01 | 0.02 | −0.03 | −0.0 | −0.01 | −0.01 | −0.01 | −0.01 | −0.01 | 0.07 | −0.24 | 0.02 | −0.01 |
| 4346 | 0.01 | 0.01 | 0.0 | −0.01 | 0.01 | 0.01 | 0.01 | −0.0 | −0.0 | −0.0 | −0.0 | 0.0 | **0.01** | 0.0 | −0.0 | 0.01 | −0.0 | −0.0 | 0.0 | 0.0 | 0.01 | −0.02 | −0.01 | 0.0 | −0.0 | 0.01 | −0.01 | 0.0 | −0.0 | 0.0 | −0.0 | −0.0 | 0.0 | 0.04 | −0.13 | 0.02 | −0.0 |
| 4378 | 0.01 | 0.02 | 0.01 | −0.0 | 0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | −0.0 | 0.0 | **0.01** | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | −0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.04 | 0.01 | 0.01 | 0.03 | 0.02 | 0.02 | 0.01 | −0.04 | 0.17 | −0.0 | 0.01 |
| 4424 | 0.0 | 0.01 | 0.01 | 0.01 | −0.01 | −0.01 | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | −0.01 | −0.0 | 0.01 | **0.02** | 0.01 | 0.02 | 0.02 | 0.01 | 0.02 | −0.0 | 0.03 | 0.03 | 0.01 | 0.02 | −0.03 | 0.06 | 0.01 | 0.02 | 0.03 | 0.03 | 0.02 | 0.01 | −0.1 | 0.4 | −0.03 | 0.02 |
| 4474 | 0.03 | 0.02 | 0.01 | 0.0 | 0.0 | 0.01 | 0.02 | 0.0 | −0.01 | 0.0 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | **0.02** | 0.0 | −0.01 | 0.02 | 0.02 | 0.04 | −0.06 | −0.02 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.0 | 0.03 | 0.01 | 0.01 | 0.01 | 0.03 | −0.03 | 0.02 | 0.0 |
| 4511 | 0.0 | 0.01 | 0.01 | −0.01 | 0.01 | −0.0 | 0.0 | 0.02 | 0.02 | 0.02 | 0.0 | −0.01 | −0.0 | 0.01 | 0.02 | 0.0 | **0.02** | 0.03 | −0.0 | 0.01 | −0.01 | 0.05 | 0.03 | 0.01 | 0.02 | −0.03 | 0.05 | 0.01 | 0.02 | 0.03 | 0.02 | 0.02 | 0.01 | −0.09 | 0.31 | −0.02 | 0.02 |
| 4532 | −0.01 | 0.01 | 0.01 | −0.02 | 0.01 | −0.01 | −0.0 | 0.02 | 0.03 | 0.02 | 0.0 | −0.01 | −0.0 | 0.01 | 0.02 | −0.01 | 0.03 | **0.04** | −0.01 | 0.01 | −0.03 | 0.09 | 0.05 | 0.01 | 0.02 | −0.04 | 0.05 | 0.0 | 0.02 | 0.02 | 0.03 | 0.02 | 0.01 | −0.13 | 0.41 | −0.03 | 0.02 |
| 4559 | 0.02 | 0.01 | 0.0 | 0.01 | −0.01 | 0.01 | 0.02 | 0.0 | −0.01 | 0.0 | 0.01 | 0.0 | 0.0 | 0.01 | 0.01 | 0.02 | −0.0 | −0.01 | **0.02** | 0.01 | 0.04 | −0.06 | −0.02 | 0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.0 | 0.02 | 0.0 | 0.01 | 0.01 | 0.03 | −0.03 | 0.01 | −0.0 |
| 4613 | 0.02 | 0.02 | 0.01 | 0.0 | −0.01 | −0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.0 | 0.02 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | **0.02** | 0.02 | −0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.05 | 0.01 | 0.01 | 0.04 | 0.02 | 0.02 | 0.02 | −0.05 | 0.24 | −0.01 | 0.01 |
| 4638 | 0.05 | 0.02 | 0.01 | 0.0 | 0.0 | 0.03 | 0.03 | −0.01 | −0.03 | −0.01 | 0.01 | 0.01 | 0.01 | 0.02 | −0.0 | 0.04 | −0.01 | −0.03 | 0.04 | 0.02 | **0.09** | −0.14 | −0.06 | 0.01 | −0.0 | 0.05 | −0.02 | 0.01 | −0.01 | 0.03 | −0.01 | 0.01 | 0.01 | 0.15 | −0.37 | 0.06 | −0.01 |
| 4651 | −0.07 | −0.01 | 0.01 | −0.03 | 0.01 | −0.06 | −0.05 | 0.04 | 0.08 | 0.04 | −0.0 | −0.04 | −0.02 | −0.01 | 0.03 | −0.06 | 0.05 | 0.09 | −0.06 | −0.01 | −0.14 | **0.32** | 0.14 | 0.01 | 0.03 | −0.12 | 0.11 | −0.01 | 0.04 | −0.0 | 0.05 | 0.02 | 0.0 | −0.38 | 1.1 | −0.12 | 0.06 |
| 4669 | −0.02 | 0.01 | 0.01 | −0.0 | −0.01 | −0.03 | −0.01 | 0.03 | 0.04 | 0.03 | 0.01 | −0.02 | −0.01 | 0.01 | 0.03 | −0.02 | 0.03 | 0.05 | −0.02 | 0.01 | −0.06 | 0.14 | **0.08** | 0.02 | 0.03 | −0.07 | 0.1 | 0.0 | 0.03 | 0.02 | 0.04 | 0.03 | 0.01 | −0.23 | 0.75 | −0.07 | 0.04 |
| 4717 | 0.01 | 0.01 | 0.01 | −0.01 | 0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.02 | **0.01** | 0.01 | −0.01 | 0.04 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.01 | −0.05 | 0.21 | −0.01 | 0.01 |
| 4792 | 0.0 | 0.01 | 0.01 | −0.0 | −0.0 | −0.01 | 0.0 | 0.01 | 0.01 | 0.02 | 0.01 | −0.01 | −0.0 | 0.01 | 0.02 | 0.0 | 0.02 | 0.02 | 0.0 | 0.01 | −0.0 | 0.03 | 0.03 | 0.01 | **0.01** | −0.02 | 0.05 | 0.01 | 0.01 | 0.03 | 0.02 | 0.02 | 0.01 | −0.09 | 0.31 | −0.02 | 0.02 |
| 4824 | 0.02 | −0.01 | −0.01 | −0.0 | 0.01 | 0.03 | 0.01 | −0.03 | −0.03 | −0.03 | −0.01 | 0.02 | 0.01 | −0.01 | −0.03 | 0.01 | −0.03 | −0.04 | 0.01 | −0.01 | 0.05 | −0.12 | −0.07 | −0.01 | −0.02 | **0.06** | −0.09 | −0.0 | −0.03 | −0.02 | −0.04 | −0.02 | −0.01 | 0.21 | −0.69 | 0.06 | −0.03 |
| 4875 | 0.0 | 0.04 | 0.02 | 0.02 | −0.04 | −0.04 | 0.01 | 0.05 | 0.05 | 0.06 | 0.03 | −0.03 | −0.01 | 0.04 | 0.06 | 0.01 | 0.05 | 0.05 | 0.01 | 0.05 | −0.02 | 0.11 | 0.1 | 0.04 | 0.05 | −0.09 | **0.2** | 0.02 | 0.05 | 0.09 | 0.08 | 0.06 | 0.04 | −0.33 | 1.25 | −0.08 | 0.06 |
| 4914 | 0.01 | 0.01 | 0.01 | 0.0 | −0.0 | 0.0 | 0.01 | 0.01 | 0.0 | 0.01 | 0.0 | −0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | −0.01 | 0.0 | 0.01 | 0.01 | −0.0 | 0.02 | **0.01** | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | −0.02 | 0.11 | 0.0 | 0.01 |
| 4987 | 0.0 | 0.01 | 0.01 | 0.0 | −0.01 | −0.01 | 0.0 | 0.02 | 0.02 | 0.02 | 0.01 | −0.01 | −0.0 | 0.01 | 0.02 | 0.0 | 0.02 | 0.02 | 0.0 | 0.01 | −0.01 | 0.04 | 0.03 | 0.01 | 0.01 | −0.03 | 0.05 | 0.01 | **0.02** | 0.03 | 0.02 | 0.02 | 0.01 | −0.09 | 0.35 | −0.02 | 0.02 |
| 5008 | 0.03 | 0.03 | 0.02 | 0.0 | −0.01 | −0.0 | 0.02 | 0.03 | 0.01 | 0.03 | 0.02 | −0.01 | 0.0 | 0.03 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 | 0.04 | 0.03 | −0.0 | 0.02 | 0.02 | 0.03 | −0.02 | 0.09 | 0.02 | 0.03 | **0.06** | 0.04 | 0.04 | 0.03 | −0.1 | 0.46 | −0.02 | 0.03 |
| 5025 | 0.0 | 0.02 | 0.01 | 0.0 | −0.01 | −0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.01 | −0.01 | −0.0 | 0.02 | 0.03 | 0.01 | 0.02 | 0.03 | 0.0 | 0.02 | −0.01 | 0.05 | 0.04 | 0.02 | 0.02 | −0.04 | 0.08 | 0.01 | 0.02 | 0.04 | **0.03** | 0.03 | 0.02 | −0.14 | 0.51 | −0.03 | 0.03 |
| 6447 | 0.01 | 0.02 | 0.01 | 0.0 | −0.01 | −0.01 | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | −0.01 | −0.0 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 | 0.02 | 0.03 | 0.02 | 0.02 | −0.02 | 0.06 | 0.01 | 0.02 | 0.04 | 0.03 | **0.02** | 0.02 | −0.09 | 0.36 | −0.02 | 0.02 |
| 6472 | 0.01 | 0.02 | 0.01 | −0.0 | 0.0 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | −0.01 | 0.0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.0 | 0.01 | 0.01 | 0.01 | −0.01 | 0.04 | 0.01 | 0.01 | 0.03 | 0.02 | 0.02 | **0.01** | −0.05 | 0.2 | −0.0 | 0.01 |
| 6520 | 0.06 | −0.04 | −0.03 | −0.01 | 0.05 | 0.1 | 0.02 | −0.1 | −0.11 | −0.1 | −0.03 | 0.07 | 0.04 | −0.04 | −0.1 | 0.03 | −0.09 | −0.13 | 0.03 | −0.05 | 0.15 | −0.38 | −0.23 | −0.05 | −0.09 | 0.21 | −0.33 | −0.02 | −0.09 | −0.1 | −0.14 | −0.09 | −0.05 | **0.73** | −2.46 | 0.21 | −0.12 |
| 6582 | −0.13 | 0.17 | 0.13 | 0.1 | −0.25 | −0.34 | −0.01 | 0.35 | 0.36 | 0.37 | 0.14 | −0.24 | −0.13 | 0.17 | 0.4 | −0.03 | 0.31 | 0.41 | −0.03 | 0.24 | −0.37 | 1.1 | 0.75 | 0.21 | 0.31 | −0.69 | 1.25 | 0.11 | 0.35 | 0.46 | 0.51 | 0.36 | 0.2 | −2.46 | **8.74** | −0.7 | 0.41 |
| 6643 | 0.03 | 0.0 | −0.0 | −0.01 | 0.03 | 0.04 | 0.02 | −0.02 | −0.03 | −0.02 | −0.01 | 0.02 | 0.02 | −0.0 | −0.03 | 0.02 | −0.02 | −0.03 | 0.01 | −0.01 | 0.06 | −0.12 | −0.07 | −0.01 | −0.02 | 0.06 | −0.08 | 0.0 | −0.02 | −0.02 | −0.03 | −0.02 | −0.0 | 0.21 | −0.7 | **0.07** | −0.03 |
| 6728 | −0.0 | 0.01 | 0.01 | −0.0 | −0.01 | −0.01 | 0.0 | 0.02 | 0.02 | 0.02 | 0.01 | −0.01 | −0.0 | 0.01 | 0.02 | 0.0 | 0.02 | 0.02 | −0.0 | 0.01 | −0.01 | 0.06 | 0.04 | 0.01 | 0.02 | −0.03 | 0.06 | 0.01 | 0.02 | 0.03 | 0.03 | 0.02 | 0.01 | −0.12 | 0.41 | −0.03 | **0.02** |

This paper has been typeset from a TEX/LATEX file prepared by the author.