Fundamental physics with ESPRESSO: Precise limit on variations in the fine-structure constant towards the bright quasar HE 0515−4414[*],[**]

Michael T. Murphy[1,2], Paolo Molaro[3,2], Ana C. O. Leite[4,5,6], Guido Cupani[3,2], Stefano Cristiani[3,2],
Valentina D'Odorico[3,2], Ricardo Génova Santos[7,8], Carlos J. A. P. Martins[4,6], Dinko Milaković[3,2,9],
Nelson J. Nunes[10,11], Tobias M. Schmidt[12,3], Francesco A. Pepe[12], Rafael Rebolo[7,8], Nuno C. Santos[4,5],
Sérgio G. Sousa[4,5], Maria-Rosa Zapatero Osorio[13], Manuel Amate[7], Vardan Adibekyan[4,5], Yann Alibert[14],
Carlos Allende Prieto[7,8], Veronica Baldini[3], Willy Benz[14], François Bouchy[12], Alexandre Cabral[10,11],
Hans Dekker[15], Paolo Di Marcantonio[3], David Ehrenreich[12], Pedro Figueira[4,16], Jonay I. González Hernández[7,8],
Marco Landoni[17], Christophe Lovis[12], Gaspare Lo Curto[16], Antonio Manescau[15], Denis Mégevand[12],
Andrea Mehner[16], Giuseppina Micela[18], Luca Pasquini[15], Ennio Poretti[17,19,20], Marco Riva[17],
Alessandro Sozzetti[20], Alejandro Suárez Mascareño[7,8], Stéphane Udry[4], and Filippo Zerbi[17]

1 Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia
  e-mail: mmurphy@swin.edu.au
2 Institute for Fundamental Physics of the Universe, Via Beirut 2, 34151 Miramare, Trieste, Italy
3 INAF – Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
4 Instituto de Astrofísica e Ciências do Espaço, CAUP, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
5 Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua Campo Alegre, 4169-007 Porto,
  Portugal
6 Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
7 Instituto de Astrofísica de Canarias (IAC), Calle Vía Láctea s/n, 38205 La Laguna, Tenerife, Spain
8 Departamento de Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain
9 Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Via Bonomea 265, 34136 Trieste, Italy
10 Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa,
  Portugal
11 Departamento de Física da Faculdade de Ciências da Universidade de Lisboa, Edifício C8, 1749-016 Lisboa, Portugal
12 Observatoire Astronomique de l'Université de Genève, Chemin Pegasi 51, Sauverny 1290, Switzerland
13 Centro de Astrobiología (CSIC-INTA), Crta. Ajalvir km 4, 28850 Torrejón de Ardoz, Madrid, Spain
14 Physics Institute, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
15 European Southern Observatory, Karl-Schwarzschild-Strasse 2, 85748 Garching b. München, Germany
16 European Southern Observatory, Alonso de Córdova 3107, Vitacura, Región Metropolitana, Chile
17 INAF – Osservatorio Astronomico di Brera, Via E. Bianchi 46, 23807 Merate, Italy
18 INAF – Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, 90134 Palermo, Italy
19 INAF – Osservatorio Nazionale Galileo, Rambla José Ana Férnandez Pérez, 7, 38712 San Antonio de Breña Baja, Spain
20 INAF – Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese, Italy

ABSTRACT

The strong intervening absorption system at redshift 1.15 towards the very bright quasar HE 0515−4414 is the most studied absorber for measuring possible cosmological variations in the fine-structure constant, α. We observed HE 0515−4414 for 16.1 h with the Very Large Telescope and present here the first constraint on relative variations in α with parts-per-million (ppm) precision from the new ESPRESSO spectrograph: $\Delta\alpha/\alpha = 1.3 \pm 1.3_{\rm stat} \pm 0.4_{\rm sys}$ ppm. The statistical uncertainty (1σ) is similar to the ensemble precision of previous large samples of absorbers and derives from the high signal-to-noise ratio achieved (≈ 105 per $0.4\,{\rm km\,s^{-1}}$ pixel). ESPRESSO's design, and the calibration of our observations with its laser frequency comb, effectively removed wavelength calibration errors from our measurement. The high resolving power of our ESPRESSO spectrum ($R = 145\,000$) enabled the identification of very narrow components within the absorption profile, allowing a more robust analysis of $\Delta\alpha/\alpha$. The evidence for the narrow components is corroborated by their correspondence with previously detected molecular hydrogen and neutral carbon. The main remaining systematic errors arise from ambiguities in the absorption profile modelling, effects from redispersing the individual quasar exposures, and convergence of the parameter estimation algorithm. All analyses of the spectrum, including systematic error estimates, were initially

---

blinded to avoid human biases. We make our reduced ESPRESSO spectrum of HE 0515−4414 publicly available for further analysis. Combining our ESPRESSO result with 28 measurements, from other spectrographs, in which wavelength calibration errors have been mitigated yields a weighted mean $\Delta\alpha/\alpha = -0.5 \pm 0.5_{\text{stat}} \pm 0.4_{\text{sys}}$ ppm at redshifts 0.6−2.4.

Key words. quasars: absorption lines – quasars: individual: HE 0515–4414 – intergalactic medium – cosmology: observations – cosmology: miscellaneous – instrumentation: spectrographs

1. Introduction

The standard model of particle physics is characterised by a set of physical interactions, whose strengths are described through fundamental dimensionless couplings. Historically, these have been assumed to be spacetime-invariant. However, fundamental couplings are known to run with energy, and in many extensions of the standard model they will also change in time and possibly in space. These, and other aspects of this field have been reviewed recently by Martins (2017).

The fine-structure constant, $\alpha \equiv e^2/\hbar c$, and the proton-to-electron mass ratio, μ, are two dimensionless fundamental constants that can be probed directly with spectroscopic techniques. In Earth-based laboratories, comparing atomic clocks based on different transition frequencies over $\sim$2−4 yr timescales has provided extraordinarily precise limits on local time variations (e.g. Rosenband et al. 2008). The most recent measurements constrain the time rate of change in α and μ to just $(1.0 \pm 1.1) \times 10^{-18}$ and $(-8 \pm 36) \times 10^{-18}$ yr^{-1}, respectively (Lange et al. 2021). While it may be tempting to linearly extrapolate such constraints to cosmological time (and distance) scales, it should be emphasised that how the fundamental constants may vary, and on what they may depend, is entirely unknown; such extrapolations would be just as simplistic as assuming no variation at all. Clearly, the variability or constancy of α and μ must be explicitly tested over the full range of time (and distance) scales available to experiments.

The most effective way to probe cosmological variations in α is by measuring any relative shifts between metal transitions produced in intervening quasar absorption systems. This technique was introduced by Dzuba et al. (1999a,b) and Webb et al. (1999) and is known as the many multiplet (MM) method because it utilises the wavelengths of many different transitions from several atomic species. Transitions from different multiplets have different dependencies on α, so measuring the velocity shifts between these transitions in a quasar absorption system provides a direct probe of $\Delta\alpha/\alpha$ – the relative difference between α in the absorber (α_{abs}) and its current laboratory value on Earth (α_{lab}):

$$\frac{\Delta\alpha}{\alpha} \equiv \frac{\alpha_{\text{abs}} - \alpha_{\text{lab}}}{\alpha_{\text{lab}}} \approx -\frac{\Delta v_i}{c}\frac{1}{2Q_i}, \qquad (1)$$

where Δv_i is the velocity shift caused by a small variation in α (i.e. $\Delta\alpha/\alpha \ll 1$; Dzuba et al. 2002). Here, $Q_i \equiv \omega_i/q_i$ where q_i is the sensitivity coefficient: the expected sensitivity of transition i's laboratory rest wavenumber (ω_i) to variations in α. These q coefficients have been calculated in numerous works, using several different many-body quantum mechanical techniques, starting with Dzuba et al. (1999a) and compiled in Murphy & Berengut (2014). For typical values of q, a 1 ppm variation in α would produce a shift of order $\Delta v \sim 20 \, \text{m s}^{-1}$ between different transitions.

The MM method has been widely used to measure $\Delta\alpha/\alpha$ at large redshifts (i.e. large look-back-times) with the largest samples obtained with archival spectra from two high-resolution spectrographs: UVES at the VLT and HIRES at the Keck Observatory. Webb et al. (2011) and King et al. (2012) combined the sample of 143 Keck/HIRES absorption systems in Murphy et al. (2003, 2004) with a sample of 154 from VLT/UVES to produce

a data set of $\Delta\alpha/\alpha$ measurements in 293 distinct absorption systems. This sample showed a statistical preference for dipolar spatial variation of α across the sky at the $\sim$10 ppm level with $>4\sigma$ statistical significance (Webb et al. 2011). As should be expected for such a surprising result, with potential impact on fundamental physics, many authors have questioned the data, analysis, assumptions and potential systematic errors underpinning these measurements, and presented alternative data sets and analyses (e.g. Chand et al. 2004; Quast et al. 2004; Levshakov et al. 2005, 2007; Molaro et al. 2008a). However, these did not provide compelling alternative explanations or results (e.g. Murphy et al. 2007a, 2008; Wilczynska et al. 2015). Constraints from higher quality spectra of individual absorbers were also obtained, but none of them directly supported or strongly conflicted with the α dipole evidence (e.g. Molaro et al. 2013a; Bainbridge & Webb 2017; Wilczynska et al. 2020).

It should be emphasised that a spatial variation is a theoretically unexpected result. While many well-motivated models allow redshift dependencies of α, possibly with additional dependencies on the local density (e.g. Martins 2017), a theoretical model which can account for a pure spatial dipole is more difficult to identify and requires significant fine-tuning (Olive et al. 2011). Testing this result with improved accuracy – that is, decreased systematic errors – is therefore essential.

One possibility is that the evidence for an α dipole may originate from different – possibly similar – systematic errors in the two spectrographs (Keck/HIRES and VLT/UVES). By observing solar spectra reflected from asteroids with UVES, and comparing them with a more accurately wavelength calibrated solar spectrum from a Fourier Transform Spectrometer, Rahmani et al. (2013) found that the usual thorium–argon (ThAr) lamp calibration produced long-range wavelength distortions. Gradients in these distortions were as large as $\sim$400 m s^{-1} over 600 Å ranges. Using 20 years of archival HIRES and UVES spectra of asteroids and solar twins, Whitmore & Murphy (2015) found that long-range distortions in their (ThAr-based) wavelength scales were ubiquitous and substantial, with gradients typically varying between $\pm$200 m s^{-1} per 1000 Å, potentially causing spurious shifts in $\Delta\alpha/\alpha$ at the $\sim$10 ppm level. Using a simple model of the distortions, they were able to substantially explain the previous evidence for α variations from the large UVES sample, and partially explain that of the HIRES sample. The smaller samples of HIRES and UVES spectra cited above would be similarly affected.

While it remains plausible that more detailed distortion modelling may not fully explain the HIRES and UVES results (Dumont & Webb 2017), the additional uncertainties introduced into the $\Delta\alpha/\alpha$ measurements by such models would greatly decrease the α dipole's statistical significance. Overall, the wavelength calibration distortions in slit-based spectrographs substantially undermine confidence in the evidence for spatial variations in α. Furthermore, recent quasar observations with UVES, HIRES, and Subaru/HDS, dedicated to measuring α and explicitly corrected for distortions (with contemporaneous solar twin and asteroid spectra), or insensitive to them (utilising transitions close together in wavelength), do not support spatial variations but are only inconsistent with the dipole at the 2σ

level (Evans et al. 2014; Murphy et al. 2016; Kotuš et al. 2017; Murphy & Cooksey 2017). On balance, there is currently no compelling evidence for variations in α over cosmological time or distance scales. Nevertheless, even when accurate corrections for distortions are possible, the uncertainty in the wavelength calibration of slit-based spectrographs dominates the total error budget.

The new high-resolution spectrograph at the VLT, ESPRESSO (Echelle SPectrograph for Rocky Exoplanet and Stable Spectroscopic Observations; Pepe et al. 2021), was specifically designed to suppress wavelength calibration errors in quasar absorption measurements of α (Molaro et al. 2006; Molaro 2009). ESPRESSO is fed by optical fibres, is sealed in a stable vacuum vessel with temperature control at the mK level, and can be calibrated with an 'astrocomb' – a femtosecond-pulsed laser frequency comb (LFC) with secondary mode filtering adapted to ESPRESSO's resolving power. In principle, the high frequency-space density of uniformly separated (filtered) comb modes, whose individual frequencies are known a priori to $\ll 1\,\mathrm{cm\,s^{-1}}$ accuracy, could enable $\sim 1\,\mathrm{cm\,s^{-1}}$ (photon-limited) calibration precision (Murphy et al. 2007b). The ESPRESSO astrocomb calibration, described in detail by Schmidt et al. (2021) and Lovis et al. (in prep.) and summarised below (Sect. 2.2), covers the wavelength range $\sim 4850-7000\,\text{Å}$ (cf. ESPRESSO's full range of $3780-7780\,\text{Å}$) and provides a $\lesssim 1\,\mathrm{m\,s^{-1}}$ accuracy for velocities measured between the transitions in the spectrum analysed here (Sect. 2.2). This is well below the ensemble statistical precision of the quasar absorption measurements ($\sim 30\,\mathrm{m\,s^{-1}}$), so ESPRESSO effectively removes wavelength calibration from the error budget. The fibre feed and highly stabilised environment ensures that drifts in ESPRESSO's wavelength calibration are far too slow ($>$days) and small ($\lesssim 1\,\mathrm{m\,s^{-1}}$) to affect quasar absorption measurements of α.

Here we make the first precise ESPRESSO measurement of $\Delta\alpha/\alpha$, with LFC calibration providing a highly accurate wavelength scale. We selected HE 0515−4414 as a high-value target because it is one of the brightest quasars in the southern sky, with an intervening absorption system at redshift $z_{\mathrm{abs}} = 1.1508$ which places the most important metal transitions within the LFC-calibrated wavelength range. The absorption profile also contains many separate, relatively narrow features which appear in many different transitions; in simple terms, this increases the number of centroid measurements that combine to constrain $\Delta\alpha/\alpha$. These advantages mean this absorber has been studied numerous times and provided the most precise cosmological constraint on $\Delta\alpha/\alpha$ from any individual system (Quast et al. 2004; Levshakov et al. 2006; Chand et al. 2006; Molaro et al. 2008a; Kotuš et al. 2017; Milaković et al. 2021).

2. Data acquisition and reduction

2.1. Observations

HE 0515−4414 (J2000 right ascension $05^{\mathrm{h}}17^{\mathrm{m}}07.6^{\mathrm{s}}$ and declination $-44°10'55.58''$) was identified as a very bright ($m_v \approx 15.2\,\mathrm{mag}$; Gaia $G = 14.90\,\mathrm{mag}$) quasar at redshift $z_{\mathrm{em}} = 1.71$ by Reimers et al. (1998). It was part of the ESPRESSO Consortium's Guaranteed Time Observations (GTO) in two main runs: a visitor-mode run on 4–7 November 2018, and service mode observations between November 2019 and March 2020. Table 1 provides a journal of the observations. The total integration of 57 916 s was obtained over 17 exposures: 32 400 s over 9 exposures in 2018, and 25 516 s over 8 exposures in 2019 and 2020. ESPRESSO is located in the incoherent combined Coudé facility, underneath the four Unit Telescopes (UTs), and can be fed by any UT or, indeed, all four simultaneously. The high-resolution,

Table 1. Journal of the observations under the ESPRESSO Consortium's GTO program (ESO Project ID 1102.A-0852).

Exposure start (UTC)	Exposure (s)	Airmass	IQ ('')	S/N
2018-11-04 07:00	3600	1.07	0.68	31
2018-11-05 03:36	3600	1.33	1.03	20
2018-11-05 04:37	3600	1.17	0.89	24
2018-11-05 05:38	3600	1.09	0.84	27
2018-11-05 06:39	3600	1.07	0.87	28
2018-11-05 07:40	3600	1.10	0.82	28
2018-11-07 05:21	3600	1.10	0.84	26
2018-11-07 06:22	3600	1.07	1.00	26
2018-11-07 07:27	3600	1.10	0.86	21
2019-11-26 06:53	3500	1.15	0.92	30
2020-02-28 01:14	1016	1.16	1.11	8
2020-02-28 01:33	3500	1.26	0.86	25
2020-02-28 02:34	3500	1.49	1.12	19
2020-03-03 00:34	3600	1.16	1.16	22
2020-03-19 02:34	3400	2.13	1.05	25
2020-03-20 01:41	3500	1.68	1.14	26
2020-03-21 01:39	3500	1.69	1.08	26

Notes. The airmass is the average value during the exposure, and IQ is the full-width-half-maximum image quality measured by the image analysis detector at the telescope's focus; this has been found empirically to correspond well with occasional image quality measurements taken with a detector at the end of the Coudé train. The signal-to-noise ratio (S/N) per $0.4\,\mathrm{km\,s^{-1}}$ is measured near 6000 Å as an average over 4 Å in the extracted spectrum (after combining both traces of the same echelle order).

single-UT mode with 2-pixel binning in the spatial direction (i.e. 'singleHR21') was selected for observing HE 0515−4414, proving a nominal resolving power of $R \sim 145\,000$ with a $1''0$ diameter fibre (in the red arm). All exposures were obtained with UT3-Melipal, except for those in November 2019 which were observed with UT1-Antu. It was later found that the atmospheric dispersion corrector of UT3 was not perfectly aligned during the 2018 run, causing a $\sim 30\%$ loss in throughput. Also, the fibre link was upgraded in June 2019 and this increased the total efficiency by $\sim 40\%$ (Pepe et al. 2021). The cloud and seeing conditions varied considerably during our two runs. These factors combined to produce variations in the signal-to-noise ratio (S/N) from ≈ 19 to 30 per $\approx 0.4\,\mathrm{km\,s^{-1}}$ pixel at 6000 Å in the extracted spectra. One exposure was terminated early, with $<1/3$ of the nominal 1 h exposure time, producing a S/N of only ≈ 8.

2.2. Data reduction and wavelength calibration

The data were reduced with version 2.2.3 of the standard ESPRESSO data reduction software (DRS), based on ESO's REFLEX environment[1]. The reduction process is detailed in Pepe et al. (2021). The DRS-extracted spectra are supplied in Murphy et al. (2021) online. Most important for this work is the LFC wavelength calibration of the quasar spectra; below we focus on this aspect of the data reduction.

In the 2018 visitor-mode run, the LFC spectra were taken at the beginning of each night; that is, each quasar spectrum was calibrated with a LFC exposure taken up to $\sim 9\,\mathrm{h}$ before. For the single 2019 service-mode exposure, the LFC calibration was taken $\sim 16\,\mathrm{h}$ prior. In the 2020 service-mode run, three

---

[1] http://www.eso.org/pipelines

LFC exposures were taken on 27 and 29 February and 3 March, bracketing the first four quasar exposures, but none were taken for the last three quasar exposures (19–21 March) due to technical problems with the LFC. Those problems may have affected the latter two LFC exposures. Indeed, a comparison between the wavelength solutions from the three LFC exposures showed differences of $\approx 3\,\mathrm{m\,s^{-1}}$ in the blue arm's calibration, while those in the red arm were $<1\,\mathrm{m\,s^{-1}}$. We therefore calibrated all the 2020 quasar exposures with the 27 February LFC exposure.

In all science exposures, the quasar was observed in Fibre A while Fibre B was used to record simultaneous sky spectra instead of recording a simultaneous wavelength calibration spectrum. Therefore, no information about possible instrumental drifts during a night is available. However, ESPRESSO is highly stabilised, and drifts are known to be less than $1\,\mathrm{m\,s^{-1}}$ over a night (Pepe et al. 2021). While it appears unlikely that the changes in the last three LFC exposures above are due to instrument drifts over the 1-week timescale, we cannot rule it out entirely. If we assume those drifts are real and extrapolate them over the full 3-week timescale of the 2020 service-mode run, they would create $\sim 9\,\mathrm{m\,s^{-1}}$ differences between exposures in the blue arm and, more importantly, a $\sim 5\,\mathrm{m\,s^{-1}}$ relative shift between the two arms. The latter could potentially cause a systematic, $\sim 0.3\,\mathrm{ppm}$ shift in $\Delta\alpha/\alpha$, though this would be significantly diluted by the better-calibrated 2018 exposures. However, few of the transitions in the $z_{\mathrm{abs}} = 1.1508$ absorber of HE 0515−4414 fall in the blue arm, reducing this effect's impact on $\Delta\alpha/\alpha$. Indeed, we explicitly test the effect of removing the blue-arm transitions from our analysis in Sect. 4.3.3 and find it to be negligible. We note that this only applies to the analysis of this specific absorber, with our specific set of observations, and should not be generalised in future analyses to mean that possible instrument drifts are never important. They should be checked in each specific analysis of $\Delta\alpha/\alpha$ with ESPRESSO.

The LFC wavelength calibration procedure is explained in Schmidt et al. (2021) and Lovis et al. (in prep.). Briefly, the LFC spectrum is imaged onto part of both ESPRESSO CCDs, covering the wavelength range $\approx 4700-7200\,\text{\AA}$ (the CCDs' wavelength ranges slightly overlap at $\approx 5200-5270\,\text{\AA}$). Given the drop in LFC intensity at both edges of the spectrum, a satisfactory calibration can be obtained in a slightly more restricted range from $\approx 4850-7000\,\text{\AA}$. Outside this range the wavelength solution is normally obtained from the daytime standard ThAr and Fabry–Pérot combined calibrations, which are treated in a similar way to those used by HARPS (Cersullo et al. 2019). The ESPRESSO DRS identifies the LFC lines automatically: the LFC mode frequencies are known a priori to very high accuracy, but the mode number of a single mode must first be established, and this is achieved using a 'rough' first guess at the wavelength scale established with a traditional ThAr lamp.

The LFC spectrum itself exhibits a significant background light contribution and also strong modulation of the line intensities which may originate in the Fabry–Pérot mode-filtering cavities (Milaković et al. 2020; Schmidt et al. 2021). These can be effectively removed with detailed modelling and subtraction of the background light contribution (Schmidt et al. 2021). Finally, Schmidt et al. (2021) identified an unexpected and, so far, unexplained high-frequency pattern in the LFC calibration residuals: the best-fit centroid wavelength for neighbouring LFC modes alternate either side of the final wavelength solution by $\approx 5-7\,\mathrm{m\,s^{-1}}$. However, for the absorption system studied here, the individual transitions span $\approx 750\,\mathrm{km\,s^{-1}}$, or ~ 80 LFC modes, so any high-frequency residuals would average down to $\lesssim 1\,\mathrm{m\,s^{-1}}$ per transition. Averaged over the >7 transitions that contribute the strongest constraints on $\Delta\alpha/\alpha$, the systematic error should be $<0.5\,\mathrm{m\,s^{-1}}$, or $\lesssim 0.02\,\mathrm{ppm}$ in $\Delta\alpha/\alpha$, which is negligible.

2.3. Combined spectrum

UVES_POPLER (version 1.05; Murphy 2016; Murphy et al. 2019) was used to combine the spectra extracted by the DRS to form a single spectrum for subsequent analysis. The S/N of the combined spectrum is ≈ 105 per $0.4\,\mathrm{km\,s^{-1}}$ pixel at 6000 Å, near the Mg II 2796/2803 doublet of the $z_{\mathrm{abs}} = 1.1508$ absorber, and ≈ 85 at 5000 Å near Fe II $\lambda 2344$, the bluest line that contributes strong constraints on $\Delta\alpha/\alpha$. It is one of the highest quality echelle quasar spectra available: its S/N per $\mathrm{km\,s^{-1}}$ is ≈ 170 at 6000 Å, compared with ≈ 230 for the UVES spectrum of the same quasar in the UVES SQUAD database (Murphy et al. 2019), but with a substantially higher resolving power: $R \approx 145\,000$ for ESPRESSO compared to 75 000 for the UVES spectrum.

There are two traces of each echelle order, so the 17 exposures (Table 1) provide at least 34 independent flux measurements to be considered for each pixel in the combined spectrum (68 for the overlapping regions of neighbouring orders). This allows for aggressive rejection of outlying flux values and a reliable final (inverse-variance) weighted mean flux value. Upward spikes in flux ('cosmic rays') are automatically flagged by UVES_POPLER on each extracted order trace, and discrepant values are further flagged via iterative σ-clipping during the combination (Murphy et al. 2019); for the ESPRESSO spectrum of HE 0515−4414, individual values more than 2.5σ above or below the weighted mean were rejected. However, UVES_POPLER is also designed for detailed visual inspection and comparison of individual echelle order spectra, and this revealed many instances of low-level 'cosmic rays' and/or 'hot pixels' that were not flagged by the techniques above. These features were typically spread over $3-10$ pixels; while most individual pixels escaped the automatic clipping approaches above, the broader features were clearly visible during manual inspection. The ESPRESSO DRS has very recently been improved, and UVES_POPLER will be improved in future, to remove most of these effects automatically. Nevertheless, for the present work, they were removed based on visual inspection only. The detailed record of these manual actions is contained within the UVES_POPLER Log (UPL) file in Murphy et al. (2021) online.

We also masked all relevant telluric emission and absorption features around the transitions of interest for the $z_{\mathrm{abs}} = 1.1508$ absorber, which left small gaps in coverage and narrower regions of lower S/N, in the Fe II $\lambda 2344$ and 2586 lines, the Mg II $\lambda\lambda 2796/2803$ doublet and Mg I $\lambda 2852$. The mask was created using a synthetic spectrum of the atmosphere for Cerro Paranal with the default, recommended input parameter values in ESO's SKYMODEL software[2] (Noll et al. 2012; Jones et al. 2013). The resolution was matched to our ESPRESSO observations (nominal $R = 145\,000$) and a unit airmass was assumed. Pixels in the synthetic spectrum whose transmission was reduced by more than 0.5% by telluric features, relative to the local maximum value (within $\pm 200\,\mathrm{km\,s^{-1}}$), were flagged. Given the S/N of the combined HE 0515−4414 spectrum (~ 100 per $0.4\,\mathrm{km\,s^{-1}}$ pixel), this ensures all features potentially causing a $>1\sigma$ suppression of flux, integrated over five pixels, will be removed, assuming the features are aligned in all exposures. In practice, the 17 different exposures have somewhat different barycentric corrections, and so the Earth-frame telluric features are shifted relative to each other by up to $17\,\mathrm{km\,s^{-1}}$ between the 2018, 2019 and 2020 runs. Emission features were flagged, in a similar way, as pixels with >2 times the background emission in the synthetic spectrum than the minimum in the surrounding $\pm 200\,\mathrm{km\,s^{-1}}$. This threshold was derived by visually inspecting the sky-subtraction residuals

---

[2] See http://www.eso.org/observing/etc/skycalc/skycalc.htm

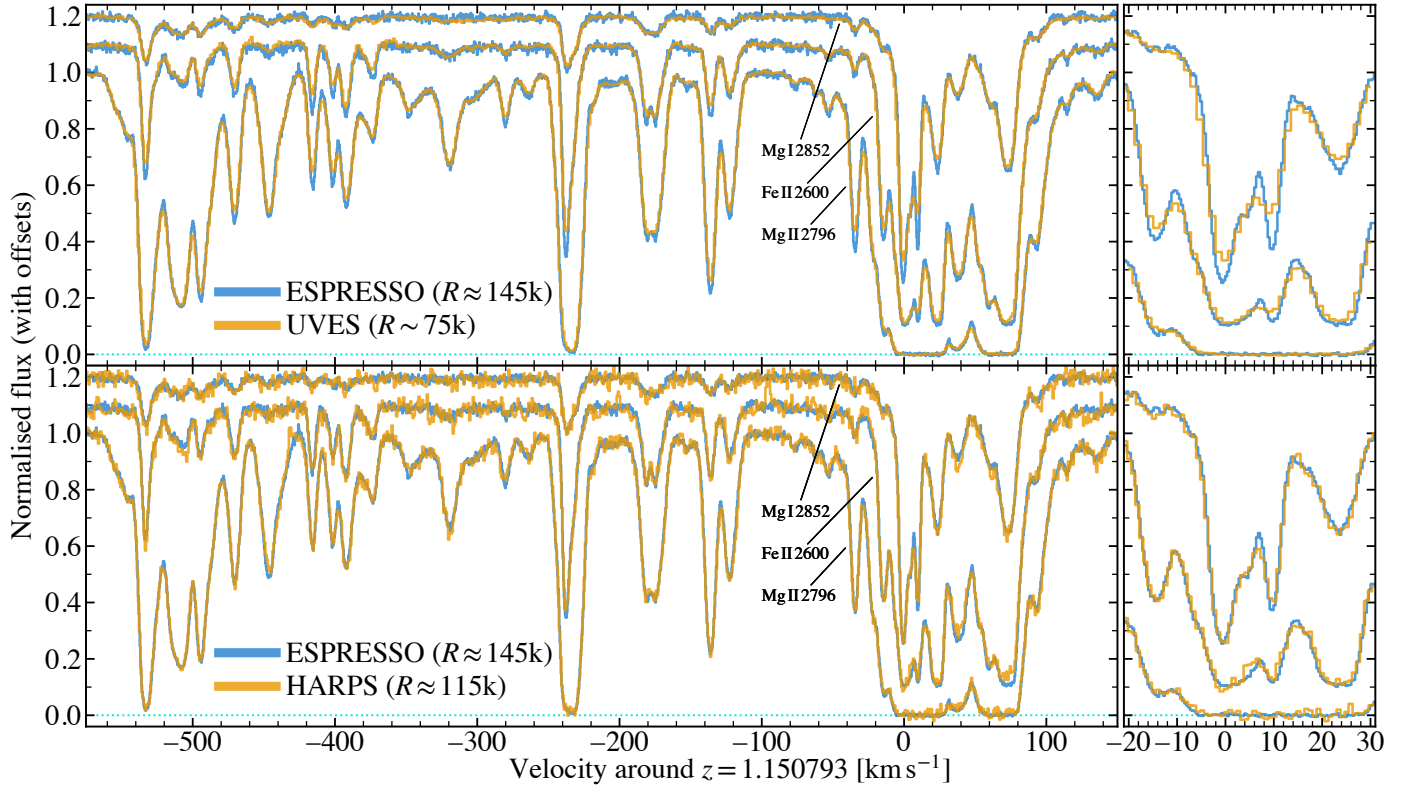**Fig. 1.** Comparison between continuum-normalised ESPRESSO, HARPS, and UVES spectra of the z_{abs} = 1.1508 absorber towards HE 0515−4414. *Upper panels*: compare the ESPRESSO (blue) and UVES (orange) spectra, while *lower panels*: compare the ESPRESSO (blue) and HARPS (orange) spectra. In each panel, the upper, middle, and lower spectra correspond to the three different transitions indicated – Mg I λ2852, Fe II λ2600 and Mg II λ2796, respectively – offset by 10% in normalised flux for clarity. We note that the resolving powers (R) of the spectra are substantially different, but the differences between the spectra are minor, except for the features near 0 and 9 km s^{-1} (see text). *Right-hand panels*: focus on this region, where the absorption is strongest (on a different velocity scale).

near very strong emission lines, and so should be very conservative. The mask was defined by excluding wavelength ranges within 3 km s^{-1} of flagged pixels, and applied to each exposure in UVES_POPLER after the combination step.

UVES_POPLER automatically constructs a nominal continuum for the combined spectrum using low-order polynomial fits. However, given the very broad absorption features of the z_{abs} = 1.1508 system, we manually re-fitted the continuum across each transition of interest. Again, this is a manual and visually guided process, but the log of these steps is transparently recorded in the UPL file, supplied in Murphy et al. (2021) online, which can easily be used to fully reproduce the combined spectrum from the DRS-extracted spectra (Murphy et al. 2019).

3. Analysis

This section presents the full details of the analysis procedure. Casual readers may wish to skip the intricacies of the absorption profile modelling in Sects. 3.5–3.7, which will be more important for those seeking to reproduce our analysis approach and for comparison with other techniques.

3.1. Comparison between ESPRESSO, HARPS, and UVES spectra

Figure 1 compares the new, combined ESPRESSO spectrum of HE 0515−4414 with those observed with HARPS (Milaković et al. 2021) and UVES (Kotuš et al. 2017) in three strong transitions of the z_{abs} = 1.1508 absorber. The resolving powers of these three spectra are substantially different –

$R \sim 75\,000$ for UVES, $\approx$115 000 for HARPS, and $\approx$145 000 for ESPRESSO – yet only small differences are apparent between the UVES and ESPRESSO spectra, and there is almost no discernable difference between the ESPRESSO and HARPS spectra. As expected, the largest differences between the ESPRESSO and UVES spectra appear in the cores of the sharpest, narrowest spectral features, but even those differences are not apparent between the ESPRESSO and HARPS spectra. This demonstrates that almost all the velocity structure of the z_{abs} = 1.1508 absorber is already resolved at $R \sim 100\,000$. This was anticipated by Kotuš et al. (2017): even the lower resolution UVES spectrum required a large number of closely-spaced (relative to their b parameters) velocity components to fit the spectral features of the velocity structure, so increasing the resolving power was predicted to reveal little unresolved structure.

However, two exceptions to the general conclusion above are the spectral features near 0 and 9 km s^{-1} in Fig. 1. When comparing the spectra in the unsaturated Mg I λ2852 transition of these features, the difference between the ESPRESSO and UVES spectra is larger than for other features; the 9 km s^{-1} feature even appears different in the ESPRESSO and HARPS spectra. This indicates the presence of particularly narrow (i.e. low b parameter) velocity components near 0 and 9 km s^{-1}. These correspond in redshift to the narrow C I absorption lines tracing very cold gas detailed by Quast et al. (2002), and are consistent with hosting the H$_2$ absorption discovered by Reimers et al. (2003). We discuss these features in Sect. 3.5 when constructing our model fit to this region of the absorber.
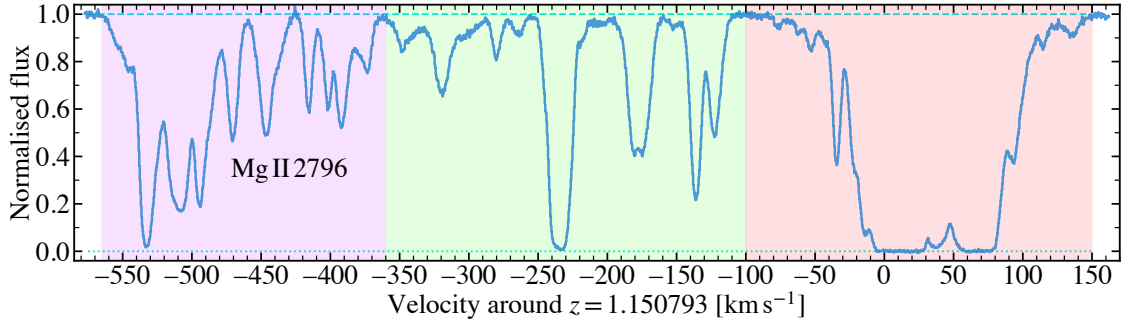
Fig. 2. Definition of 'regions' in the $z_{abs} = 1.1508$ absorber: left (purple shading; -565 to $-360\,\mathrm{km\,s^{-1}}$), central (green; -360 to $-100\,\mathrm{km\,s^{-1}}$) and right (red; -100 to $150\,\mathrm{km\,s^{-1}}$). The continuum-normalised flux in the strongest transition detected (Mg II λ2796) is plotted to illustrate that the flux returns to the fitted quasar continuum level at the region edges. While not strictly necessary in the fitting process, this helps avoid edge effects in the model when it is convolved with a (Gaussian) model of the ESPRESSO instrumental profile.

3.2. Absorption systems and transitions

Given its very complex absorption profile, we simplified the analysis of the $z_{abs} = 1.1508$ absorber – especially the profile fitting aspect – by following Kotuš et al. (2017)'s approach of splitting the velocity structure into three 'regions', as illustrated in Fig. 2, defined in velocity relative to $z_{abs} = 1.150793$: left (-565 to $-360\,\mathrm{km\,s^{-1}}$), central ($-360$ to $-100\,\mathrm{km\,s^{-1}}$) and right ($-100$ to $150\,\mathrm{km\,s^{-1}}$). The strongest absorption and narrowest features are in the right region, so it is not surprising that all previous measurements in this absorber have found the strongest constraints on $\Delta\alpha/\alpha$ from this part of the absorber. Therefore, while we analysed all three regions in the same level of detail, we describe the analysis of the right region more comprehensively here, and pay greater attention to systematic errors due to the smaller statistical uncertainty on $\Delta\alpha/\alpha$ it provides.

Figure 3 shows the transitions detected in our ESPRESSO spectrum and used to measure $\Delta\alpha/\alpha$ in the $z_{abs} = 1.1508$ absorber. The plot shows the sensitivity coefficient, Q, of each transition, as defined in Eq. (1). The main principle of the Many Multiplet method is to compare as many transitions, with as great a contrast of Q coefficients, as possible. In practice, to understand the relative importance of each transition for constraining $\Delta\alpha/\alpha$, the Q distribution in Fig. 3 has to be considered together with the absorption strength in each transition (including any saturation), and the narrowness of the features in the different parts of the velocity structure. The strongest available transitions in the $z_{abs} = 1.1508$ absorber are the Mg II doublet and the five Fe II lines at 2344–2600 Å. As Fig. 3 shows, this combination of transitions offers a high contrast in Q. Indeed, in the left and central regions, Fig. 1 shows that these transitions offer many narrow, unsaturated features to be compared. However, for the right region, the features that are strongest in Fe II are saturated in Mg II, so the Mg II doublet is less important. Instead, the Mg I λ2852 line offers unsaturated features and a very different Q value to the Fe II lines; it is these 6 transitions that most strongly constrain $\Delta\alpha/\alpha$ in the right region. We note that uncertainties in the Q values, which are $\lesssim 7\%$ for the Fe II transitions (Murphy & Berengut 2014), will cause $\Delta\alpha/\alpha$ to be systematically scaled from its true value by, at most, a similar proportion; these errors are reduced when using many transitions and are negligible for the purposes of this work in any case. We also note that the form of Eq. (1) means that systematic errors in the Q values cannot generate a spurious, non-zero $\Delta\alpha/\alpha$. Finally, as indicated in Fig. 3, many other transitions are detected, from Mg I, Fe I, Zn II, Cr II and Mn II. These are all very weak, even in the right region, and do not constrain $\Delta\alpha/\alpha$ strongly. How-
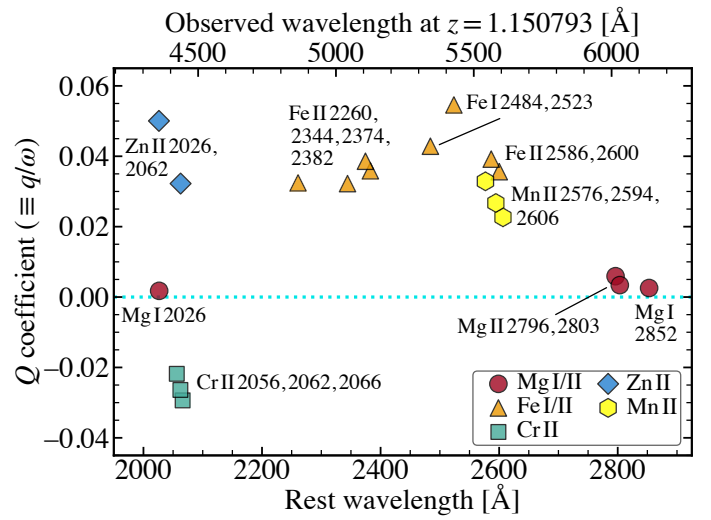

Fig. 3. Sensitivity (Q) coefficients for the transitions used here to constrain $\Delta\alpha/\alpha$ in the $z_{abs} = 1.1508$ absorber towards HE 0515$-$4414. The Q coefficients were compiled in Murphy & Berengut (2014) and are defined in Eq. (1). The lower horizontal axis shows the laboratory rest wavelengths of the transitions, while the upper axis shows the observed wavelength at z_{abs}.

ever, fitting the weak Mg I λ2026 transition simultaneously with the much stronger Mg I λ2852 line has important consequences for our model fit to the right region, as discussed in detail below (Sect. 3.5).

While we focus on measuring $\Delta\alpha/\alpha$ in the $z_{abs} = 1.1508$ absorber in this work, it is important to identify other intervening absorbers in case any of their transitions blend with those of the $z = 1.1508$ system. Given its relatively narrow wavelength range (3780–7780 Å), the ESPRESSO spectrum is not ideal for this purpose: clearly identifying any absorber between $z_{abs} = 0$ and $z_{abs} = z_{em} = 1.71$ requires a large wavelength range in which to confidently detect at least two transitions. By contrast, the wavelength coverage of the UVES spectrum of Kotuš et al. (2017) is much wider (3051–10430 Å) and better suited for this. Apart from the $z_{abs} = 1.1508$ absorber and Galactic absorption, Kotuš et al. (2017) identified 8 others in the UVES spectrum at redshifts $z_{abs} = 0.2223, 0.2818, 0.4291, 0.9406, 1.3849, 1.5145, 1.6737$ and 1.6971. For the transitions we fit, the only known blend is with the Fe II λ2344 transition from Ca II λ3934 in the $z_{abs} = 0.2818$ absorber (Kotuš et al. 2017). We therefore do not fit this Fe II transition between -4.5 and $4.5\,\mathrm{km\,s^{-1}}$ or redwards

of $80 \, \text{km s}^{-1}$ in the right region. However, during our profile fitting analysis of the right and central regions below (Sects. 3.5 and 3.6) we detected additional absorption in two transitions (Mg II λ2803 and Fe II λ2586). The identity of these possible blends remains unknown.

3.3. Voigt profile fitting approach

We fitted the absorption profiles using the same general approach as employed in many previous works (e.g. Murphy et al. 2003; King et al. 2012; Molaro et al. 2013b), particularly Kotuš et al. (2017) who studied UVES spectra of the same absorber. We therefore present only a brief summary here, leaving further details and discussion about the fits for the three different regions in Sects. 3.5–3.7.

For each region, model profiles of the multi-component velocity structure of each transition were fitted to the ESPRESSO spectrum using VPFIT (version 12.1; Carswell & Webb 2014). This is a non-linear least-squares χ^2 minimisation code designed specifically for fitting Voigt profiles to quasar absorption complexes. For a single ionic species, a velocity component is described by three free parameters – its column density, Doppler b parameter and redshift – plus several atomic parameters of the transition determined from laboratory measurements or theory: wavelength, oscillator strength, damping constant and Q coefficient. We use the atomic parameters compiled in Murphy & Berengut (2014), including the isotopic structures assuming terrestrial isotopic abundances (see further discussion in Sect. 4.3.6). To measure $\Delta\alpha/\alpha$, the corresponding components of different ionic species must be assumed to share the same redshift – see Eq. (1). While this assumption has been tested in many previous works, we do so again with our higher resolving power ESPRESSO spectrum in Sect. 4.3.5 below. The model profile is convolved with a Gaussian function representing the ESPRESSO instrumental line-shape. The Gaussian width is set to the mean resolution derived from fits to the LFC modes in the extracted wavelength calibration spectra: $2.10 \, \text{km s}^{-1}$ for the blue arm and $2.05 \, \text{km s}^{-1}$ for the red arm.

The velocity structure is not known a priori, so we take the same data-driven approach used in many recent studies: fitting as many velocity components as the data requires, as determined using χ^2 per degree of freedom, χ^2_ν, as the information criterion (e.g. Evans et al. 2014; Murphy et al. 2016; Murphy & Cooksey 2017; Kotuš et al. 2017). That is, we attempt to fit many different models, with different numbers of components, and accept the one with the lowest χ^2_ν as the fiducial model. In this process, we further assume that the b parameter of a velocity component is the same for all ionic species. This assumption of turbulent broadening has been discussed extensively in the previous works cited above, but we discuss it further in Sect. 5.4 below. Finally, all b parameters are limited to the range $0.15−30 \, \text{km s}^{-1}$ during the χ^2 minimisation process: the lower limit ensures components are not too much narrower than the $\approx 2 \, \text{km s}^{-1}$ ESPRESSO resolution element, while the upper limit ensures that weak components do not become proxies for local continuum level changes in some parts of the fitting region.

After the fiducial model is established, only then is $\Delta\alpha/\alpha$ introduced as a final free parameter (i.e. it is fixed at zero during the construction and comparison of models above). χ^2 is minimised for that model again, providing the final best-fit parameter values, including $\Delta\alpha/\alpha$, plus their 1σ uncertainties derived from the diagonal terms of the covariance matrix. We note that $\Delta\alpha/\alpha$ is not strongly covariant with any other parameter because

a single value applies to all components of all fitted transitions in a region: as α varies during the χ^2 minimisation process, it shifts all components of a transition together, by the same velocity. This is fundamentally different to how the model responds to variation in the column densities, b parameters or redshifts of individual components.

3.4. Blind analysis

Given the complexity of the $z_{\text{abs}} = 1.1508$ absorber's profile, it is not possible with the above, human-directed model construction procedure to explore every potential fit. Decisions about where to place new components, to create more detailed fits, will be somewhat subjective. Of course, measuring $\Delta\alpha/\alpha$ is our main goal, so we focussed more attention on fitting the narrowest features comprehensively, as these will most strongly constrain $\Delta\alpha/\alpha$; this reduces the number of possible fits worth exploring in detail. We then use an objective information criterion to select the fiducial model before $\Delta\alpha/\alpha$ is measured. Nevertheless, it may seem plausible that some human biases remain. In principle – though, we would argue, not in any practical way – this may lead, or even be perceived to lead, to a biased measurement of $\Delta\alpha/\alpha$. To avoid any potential for this, we conducted the entire analysis in a blind manner such that it – particularly the model construction step – could not have been influenced by the results obtained.

The procedure for blinding the analysis followed that of several previous works (Evans et al. 2014; Murphy & Cooksey 2017; Kotuš et al. 2017). Briefly, the ESPRESSO spectrum of HE 0515−4414 was 'blinded' in UVES_POPLER by introducing both long-range and intra-order distortions of the wavelength scale to each exposure. The size of the distortions varied randomly from exposure to exposure, but also contained some common features (also randomly assigned) to ensure the effects did not average away once the exposures were combined. Importantly, the size of these effects in the individual exposures and the final spectrum was not known, but their maximum possible size was limited so that they would lead to velocity shifts between the transitions corresponding to a spurious $|\Delta\alpha/\alpha| \lesssim 5$ ppm, about 4 times the final statistical uncertainty. This limit ensured that the model construction procedure was not unduly affected by the unknown shifts introduced between transitions.

All aspects of the analysis after the ESPRESSO DRS data reduction were conducted using the blinded spectra; that is, UVES_POPLER processing, profile model construction, χ^2 minimisation, parameter estimation (including $\Delta\alpha/\alpha$), systematic error analysis and consistency checks. Only once these steps were fully developed, refined, and finalised was the blinding lifted in UVES_POPLER and the entire analysis re-run, without any further changes or human intervention, to produce the final results reported here.

3.5. Model for the right region

Figures 4 and 5 show the fiducial absorption profile model for the right region. The full VPFIT input and output files, providing all the parameter values and uncertainties, are provided in Murphy et al. (2021) online. The strongest transitions, shown in Fig. 4, have much higher effective S/N – that is, the ratio of feature depth to S/N in the continuum – and so provide much stronger constraints on $\Delta\alpha/\alpha$ than the weak transitions in Fig. 5. The main statistics of the fit are summarised in Table 2, including the number of components fitted in each region. The model construction process was the most challenging for the
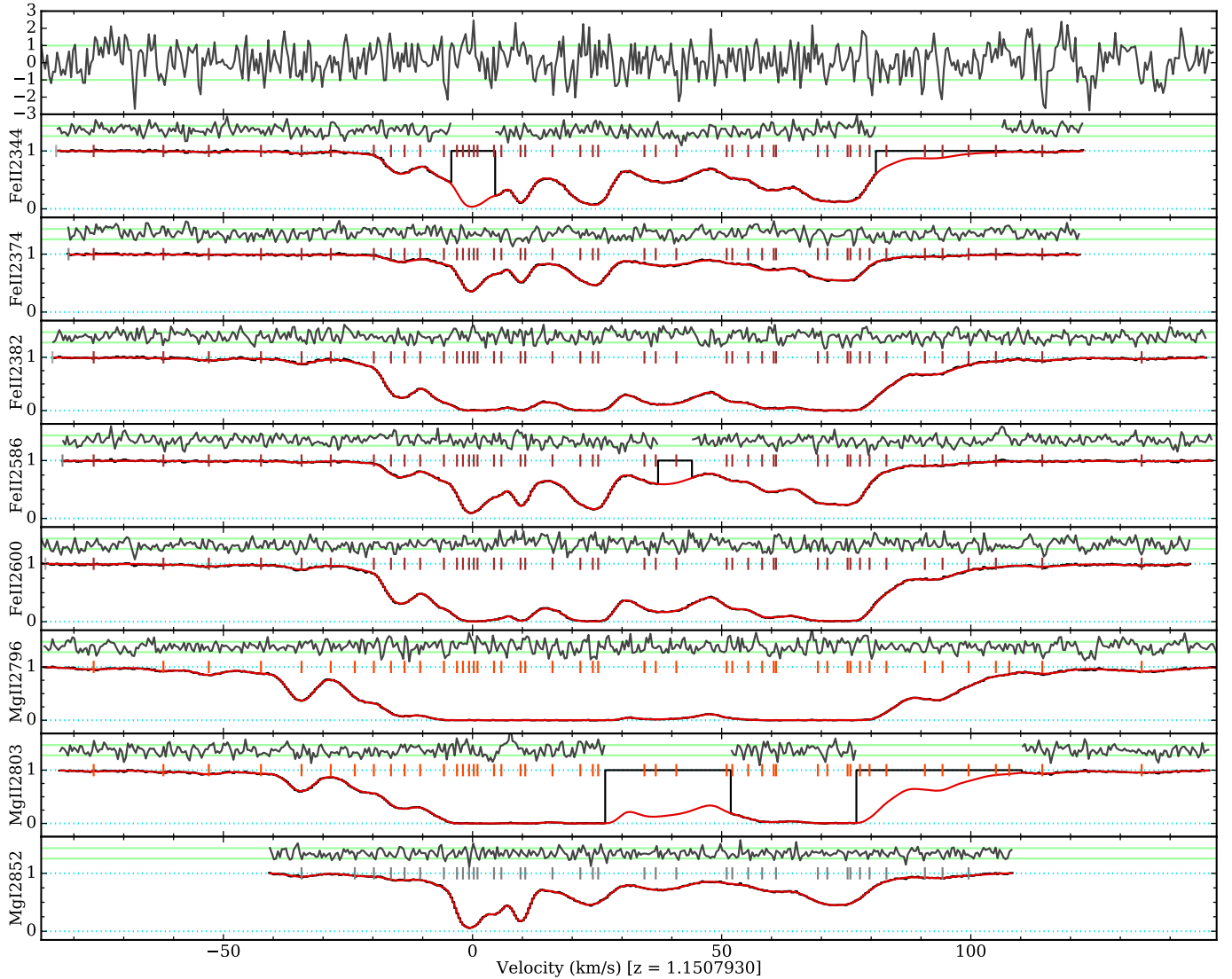
Fig. 4. Combined, normalised ESPRESSO spectrum (black histogram) in the right region of the $z_{abs} = 1.1508$ absorber, for the strongest transitions only (the weak transitions are shown in Fig. 5). Each panel shows a different transition, as labelled on the vertical axis. The red, solid line represents the fiducial model constructed from the Voigt profile components whose centroids are shown by tick-marks above each transition. We note that the fit traces the data so closely that the former almost obscures the latter in most places. To assist visualising any discrepancies, the normalised residuals (i.e. (data–model)/1σ uncertainty) are shown above each transition (black line) with the $\pm 1\sigma$ deviations marked by the green horizontal lines. *Top panel*: composite residual spectrum; that is, the average normalised residual spectrum of the transitions shown, in units of standard deviations. The data missing from Mg II λ2803, Fe II λ2586, and Fe II λ2344 were excised manually due to the presence of additional absorption, though only the latter is confirmed to arise from a known absorber; see text for discussion.

right region. Indeed, the combination of high S/N, high resolving power, large number of transitions, and the narrow-but-clearly multi-component absorption features, made it the most difficult fit we have ever attempted. A total of 48 velocity components were required to fit the strongest transitions (the Mg II doublet) in the right region. Table 2 shows that this is similar to the number used by Kotuš et al. (2017) to fit their UVES spectrum (49 components), but significantly more than used by Milaković et al. (2021) for their HARPS spectrum in the same region (26). This is unsurprising, to a first approximation, because the S/N per km s^{-1} for the ESPRESSO and UVES spectra are similar ($\approx$170 and 230, respectively) and much higher than for the HARPS spectrum ($\approx$58). However, we note that our fitting approach was the same as Kotuš et al. (2017)'s but quite different to Milaković et al. (2021)'s; this is discussed further in Sect. 5.4.

Despite the complexity and difficulty of the fitting process, the final χ^2 of the fit, per degree of freedom, is very close to unity, $\chi^2_\nu = 1.03$, indicating it is statistically acceptable. Importantly, we also do not observe any large outliers, or any remaining coherent structure in the residuals between the profile model and the data; these are normalised by the 1σ flux uncertainty spectra and plotted above each transition in Figs. 4 and 5. Finally, a 'composite residual spectrum' is plotted above each figure, showing the average of the normalised residual spectra for the transitions plotted. These were calculated in the same way described in Malec et al. (2010) and offer additional power to detect evidence of statistically significant unfitted structure. However, we do not find any such evidence in the composite residual spectra for the right region.

Several gaps are apparent in the strong transitions of the right region in Fig. 4. These were excised manually in UVES_POPLER
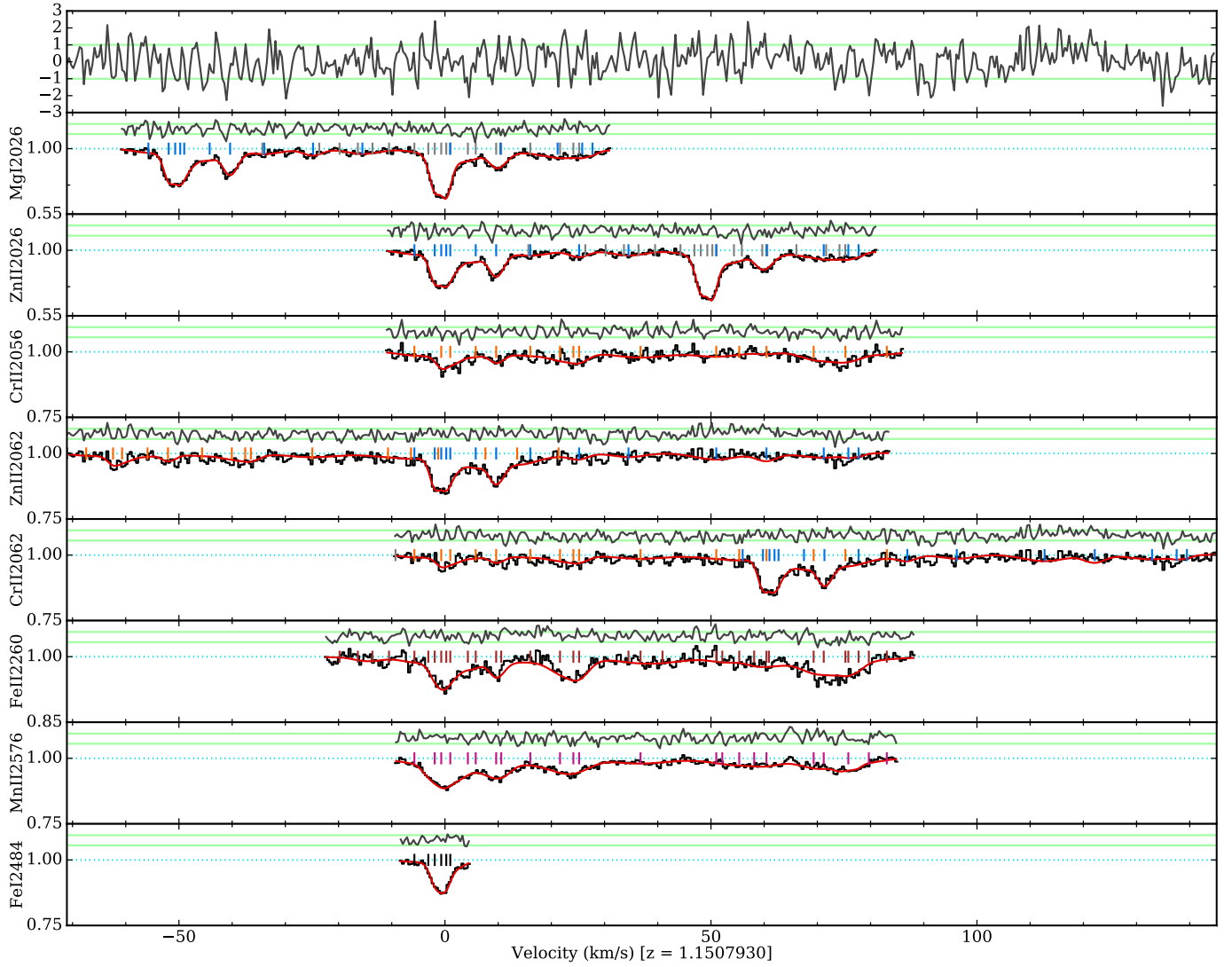
Fig. 5. Same as Fig. 4, but for the weak transitions in the right region of the absorber. For simplicity, the weakest fitted transitions of Fe I, Cr II, and Mn II are not plotted, but they were included in the fit: Fe I λ2523, Cr II λ2066, Mn II λ2594, and 2606.

because of apparent blending. However, as noted in Sect. 3.2 above, only the blend in Fe II λ2344 is associated with a known absorber (Ca II λ3934 at $z_{abs} = 0.2818$). We detected very subtle, but highly statistically significant additional absorption in both Mg II λ2803 and Fe II λ2586 in the ESPRESSO spectrum. We confirmed its presence, over the same velocity ranges, in the UVES spectrum of Kotuš et al. (2017, see their Fig. 4), indicating that this is not telluric absorption. However, despite extensive searches for additional absorbers, and even considering possible diffuse interstellar band absorption in our Galaxy, we have not been able to identify these putative blends.

The strongest, narrowest features at 0 and 9 km s⁻¹ in Fig. 4 clearly should offer excellent constraints on $\Delta\alpha/\alpha$, but were also very difficult to fit. For the 0 km s⁻¹ feature, this is because it contains at least one, and most likely two very narrow components, with Doppler parameters $b \lesssim 0.5$ km s⁻¹. Figure 6 shows the detailed velocity structure of Mg I λ2852 in this feature, with the Voigt profile of each velocity component in our fiducial fit shown explicitly; the two components at ≈-2 and 0 km s⁻¹ are both very narrow. At first, it may seem unlikely that b parameters narrower than 1/3 of the resolution element are required. Indeed, we found that a satisfactory fit of the Mg II and Fe II tran-

sitions, plus the strong Mg I λ2852 transition only, is possible without invoking these very narrow, or 'cold', components. However, this fit grossly under-predicted the absorption observed in the 0 km s⁻¹ feature of the weaker Mg I λ2026 transition. The b parameter of the strongest component in the 9 km s⁻¹ feature was similarly required to be quite narrow ($b \approx 1.6$ km s⁻¹) and some alternative fits required it to be even narrower ($b \approx 0.7$ km s⁻¹, though these fits were ultimately unstable). But this feature is very weak in Mg I λ2026, correspondingly weakening its constraint on the b parameter. The effect of the narrow components is also illustrated in Fig. 7. Compared to the other absorption features, those at 0 and 9 km s⁻¹ are particularly strong in Mg I relative to Fe II. The former also appear to be slightly shifted with respect to the corresponding Fe II feature. These are both manifestations of unresolved, cold velocity components blended with broader surrounding ones. We found no simpler velocity structure than that in Fig. 6 could fit these two features in all fitted transitions shown in Figs. 4 and 5. This illustrates the substantial benefit of the high resolving power of ESPRESSO ($R \approx 145\,000$) in this case: Fig. 1 clearly shows how, for example, UVES's lower resolving power did not reveal the effects of these narrow components, even in the very high S/N UVES spectrum used by Kotuš et al. (2017,

Table 2. Final $\Delta\alpha/\alpha$ measurements in the three regions, and statistics of the fits, compared with those of Kotuš et al. (2017) and Milaković et al. (2021).

	This work (ESPRESSO)				Kotuš et al. (2017) (UVES)				Milaković et al. (2021) (HARPS)			
Region	$\Delta\alpha/\alpha$	$\sigma(\Delta\alpha/\alpha)$	χ^2_ν	N_{comp}	$\Delta\alpha/\alpha$	$\sigma(\Delta\alpha/\alpha)$	χ^2_ν	N_{comp}	$\Delta\alpha/\alpha$	$\sigma(\Delta\alpha/\alpha)$	χ^2_ν	N_{comp}
Left	2.2	3.3	0.80	41	−6.2	3.4	1.09	32	7.2	3.1	0.99	19
Central	1.6	5.6	0.84	40	−4.7	1.7	1.17	25	−4.7	4.0	0.87	17
Right	1.1	1.5	1.03	48	−0.8	0.6	1.26	49	−2.6	3.4	0.99	26
All	1.3	1.3		129	−1.4	0.6		106	0.9	2.0		62

Notes. The best-fit values $\Delta\alpha/\alpha$ and their statistical uncertainties are in units of parts-per-million (ppm). We note that the full results of this work, with systematic uncertainties included, are presented later in Sect. 4 (Table 3). The χ^2 per degree of freedom (χ^2_ν) and number of fitted velocity components (N_{comp}) in each model are also presented. The final row provides the weighted mean $\Delta\alpha/\alpha$ and 1σ uncertainty for our work. Milaković et al. (2021) defined five regions, instead of the three used here and in Kotuš et al. (2017); the values reported below for Milaković et al. (2021) are representative only and combine their Table 2's LFC values for regions I and II for our left region, and their regions III and IV for our central region, using the inverse of their total $\Delta\alpha/\alpha$ uncertainties as weights.
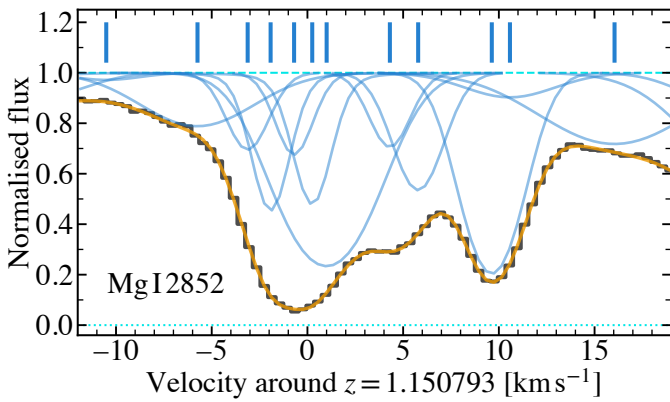

Fig. 6. Detailed Mg I $\lambda 2852$ velocity structure showing the very narrow components near $0\,km\,s^{-1}$ in the right region. The model fit (solid orange line) overlays the spectrum (black histogram) and the Voigt profiles of individual components are shown in thin blue lines, with tick marks above the spectrum indicating their centroids. In our fiducial fit shown here, the components at ≈ -2 and $0\,km\,s^{-1}$ have best-fit b parameters of just 0.29 ± 0.07 and $0.22 \pm 0.13\,km\,s^{-1}$.
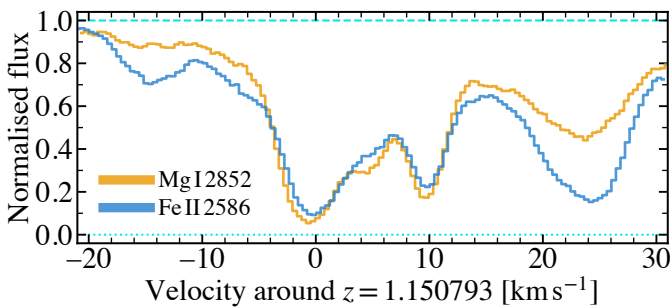

Fig. 7. Comparison between the Mg I and Fe II velocity structures in the features at 0 and $9\,km\,s^{-1}$ in the right region.

$S/N \approx 250$ per $1.3\,km\,s^{-1}$ pixel). These narrow components were also not required by the automatic fitting algorithm employed by Milaković et al. (2021), most likely because the Mg I $\lambda 2026$ transition was not covered by their HARPS spectra.

Corroborating the evidence for the cold components in the ESPRESSO spectrum is the detection of corresponding C I and H_2 absorption. Quast et al. (2002) detected two C I components at 0 and $9\,km\,s^{-1}$, finding their widths to be $b = 2.0$ and $3.5\,km\,s^{-1}$, respectively. Given the much lower resolving power

of their UVES spectra ($R \approx 55\,000$, i.e. $5.5\,km\,s^{-1}$ full-width-half-maximum resolution), it is possible, even likely these b parameters are overestimated. From the column densities in the different C I transitions, Quast et al. (2002) inferred the populations of the fine-structure levels, finding the kinetic temperature to be $\sim 240\,K$, corresponding to thermal broadening of $b_{ther} \approx 0.6\,km\,s^{-1}$. Reimers et al. (2003) discovered molecular hydrogen absorption covering the redshift range of these two cold C I components with *Hubble* Space Telescope spectroscopy ($R \sim 30\,000$), finding a total H_2 column density of $N(H_2) = 8.7^{+8.7}_{-4.0} \times 10^{16}\,cm^{-2}$. They also estimated the photodissociation rate to be more than 300 times the mean Galactic disk value, suggesting enhanced star-formation activity and a gas density of $\sim 100\,cm^{-2}$. This high density implies a very small cloud size, $L \sim 0.25\,pc$, along the line-of-sight for the molecular cloud. The relative column densities of the first two rotational levels ($J = 0$ and 1) of the ground electronic state provided an excitation temperature estimate of $T_{01} \sim 90\,K$. It is therefore unsurprising to find very narrow neutral and singly-ionised velocity components in our ESPRESSO spectrum, associated with the cold, small C I and H_2 gas clouds, which are likely embedded in warmer, more rarefied gas which gives rise to other, broader components at very similar velocities. Interestingly, complexes of clouds with precisely these properties have been observed in our Galaxy through $R \approx 10^6$ optical spectroscopy of interstellar Na I and Ca II absorption lines (Barlow et al. 1995).

The requirement to fit Mg I $\lambda 2026$ to strongly constrain the narrow components' parameters implies the need to fit the five weak transitions of Zn and Cr II. Firstly, it is important to fit the Zn II $\lambda 2026$ transition which partially blends with Mg I $\lambda 2026$ (the two transitions are separated by $50\,km\,s^{-1}$). This can be seen in the top two spectra of Fig. 5. Given this blending, constraining the Zn II parameters is greatly improved by fitting the Zn II $\lambda 2062$ transition as well, but this is blended with Cr II $\lambda 2062$ – see Fig. 5's middle two spectra. To more reliably constrain Cr II's parameters, we therefore need to fit Cr II $\lambda 2056$ and 2066 as well.

The remaining weak transitions fitted in the right region are the Mn II triplet and the two strongest transitions of Fe I: $\lambda 2484$ and 2523. To our knowledge, this is the first time Fe I transitions have been included in a $\Delta\alpha/\alpha$ measurement – they are rarely even detected in quasar absorbers. While they are very weak in our ESPRESSO spectrum, their strongest features appear just as narrow as in most other transitions so we include them in our fit mainly to help constrain the b parameters of the cold components discussed above. It is important to note that their laboratory wavelengths are only known with

$\sim$38−74 m s^{-1} precision (Nave et al. 1994; Nave & Sansonetti 2011), which is another reason they are not normally included in $\Delta\alpha/\alpha$ measurements. Nevertheless, they are only measured with effective velocity precisions of $\sim$300 m s^{-1} each in our ESPRESSO spectrum – see Sect. 4.3.5. This is much larger than the best constrained transitions in the right region ($\sim$30 m s^{-1}), and their laboratory wavelength precisions, so including them will not cause significant systematic errors in $\Delta\alpha/\alpha$. We fitted the Mn II triplet for completeness, because it was detected, but in practice it provides especially weak constraints on $\Delta\alpha/\alpha$: the absorption in all three transitions is very weak and their broad hyperfine structure ($\gtrsim$4 km s^{-1} wide; Blackwell-Whitehead et al. 2005; Aldenius 2009) effectively smooths out the sharper features.

The strong saturation in the Mg II doublet meant that the column densities of several Mg II velocity components were unconstrained. While the b parameters and redshifts of these components are, by assumption, directly linked to those of the other fitted ions, the column densities of the nine Mg II components between −6 and 18 km s^{-1}, and the three between 133 and 142 km s^{-1}, had to be fixed to reasonable values to ensure they did not destabilise the fit (e.g. become very high during the χ^2 minimisation, leading to other components being rejected by VPFIT). We estimated the fixed values by assuming the ratios of their column densities with nearby, unsaturated components matched the corresponding ratios in Fe II. This has a negligible impact on our results because the saturated regions do not directly constrain $\Delta\alpha/\alpha$. Almost identical results were obtained by artificially increasing the flux errors on the affected pixels, or by removing those pixels from the χ^2 minimisation process altogether. Similarly, fixing $\Delta\alpha/\alpha$ to zero for the saturated components, or allowing it to take a different value than for the other, unsaturated components, yielded negligible changes to the results.

Finally, we note that χ^2_ν for the right region (1.0) is slightly larger than for the central and left regions (0.8) – see Table 2. At first, it may seem that this implies that our models of the central and left regions contain too many components, and are over-fitting the data, especially since χ^2_ν in the right region is so close to the expected value of unity. However, in our combined ESPRESSO spectrum the root-mean-square flux deviations around unabsorbed sections of continuum are, in general, $\approx$10% smaller than the average formal statistical flux uncertainty in the same sections. At present, this slight overestimation of the flux error spectrum appears to arise due to a combination of effects in the DRS and UVES_POPLER (e.g. rebinning leading to small correlations between the fluxes and flux uncertainties in neighbouring pixels). The χ^2 values in the left and central regions are, therefore, expected in this case, while that in the right region may, instead, actually indicate a small degree of under-fitting.

3.6. Model for the central region

Figure 8 shows our fiducial profile fit to the absorption lines in the central region. The absorption is substantially weaker than in the right region for all ions, and the weak transitions used in the right region are not detected here. The fiducial velocity structure was, again, very difficult to construct, requiring many alternatives to be explored. In this case, the Fe II/Mg II column density ratio varies strongly across the feature at −237 km s^{-1}, making it more difficult to find a consistent velocity structure to fit to all ions. However, this was considerably simplified by the discovery that the two Mg II transition profiles were inconsistent

with each other due to additional absorption near −237 km s^{-1} in Mg II λ2803. It appears likely that this is associated with the additional absorption identified in the right region of the same transition but, similarly, we have not been able to identify its origin (see Sect. 3.5 above).

Four main features provide the constraints on $\Delta\alpha/\alpha$ in the central region: the strongest feature at −237 km s^{-1}, the double-peaked feature at −178 km s^{-1}, and the two moderately separated features at −135 and −123 km s^{-1}. For the former, the saturation in the Mg II lines means that the Mg I and Fe II transitions most strongly constrain $\Delta\alpha/\alpha$. For the other features, the Mg I absorption is too weak so the $\Delta\alpha/\alpha$ constraints come from the Mg II and Fe II transitions. The strongest absorption at −237 km s^{-1} also appears to be the narrowest, but we note that very narrow, 'cold' components were not required to fit this feature, in contrast to the 0 and 9 km s^{-1} features in the right region. While not formally detected, we do include the Mg I λ2026 transition (not shown on Fig. 8) in the fit because it helps to rule out the presence of any very narrow components in this feature; removing it from the fit has a small effect on the b parameter constraints for the velocity component fitted here, but a negligible effect on $\Delta\alpha/\alpha$, as expected.

Table 2 shows that the 40 components fitted to the central region provided a $\chi^2_\nu = 0.84$, indicating a statistically satisfactory model. There is no strong evidence in the normalised residual spectra, plotted above the individual transitions in Fig. 8, of unfitted velocity structure. The composite residual spectrum (top panel) also shows no evidence of any remaining structure. In contrast to the right region, Table 2 shows that our fit to the ESPRESSO data required substantially more components than Kotuš et al. (2017) used for the same region of their UVES spectrum (25 components). However, we note that their χ^2_ν was somewhat larger than ours, so it is possible that further exploration of alternative, more complex fits to the UVES spectrum would have yielded a statistically better fit. As in the right region, the fact that Milaković et al. (2021) fitted 17 components is likely due to the substantially lower S/N per km s^{-1} of their HARPS spectrum.

3.7. Model for the left region

Figure 9 shows our profile fit to the left region of the absorption system. Only the strongest transitions of Mg I/II and Fe II are detected and used to constrain $\Delta\alpha/\alpha$ here. While still challenging, this fit was considerably simpler to construct than those in the right and central regions. Nevertheless, the two features between −520 and −480 km s^{-1} presented some difficulties, with the five strongest components seen in Fe II not being adequate to fit the Mg I/II profiles. The latter required five additional, very weak components which appeared too weak in Fe II to be statistically required – VPFIT rejected these from the Fe II velocity structure. Given that this occurs over a small-but-contiguous portion of the profile, it is likely that this represents a real effect; that is, the Fe II/Mg II column density ratio may simply be lower in this part of the absorber. Indeed, visual comparison of the Fe II/Mg II absorption strength here with that in the strongest part of the region nearby, at −532 km s^{-1}, supports this conclusion.

In this left region of the absorber there are numerous distinct features which all contribute constraints on $\Delta\alpha/\alpha$, but the combination of Mg I and Fe II transitions in the strongest feature at −582 km s^{-1} are the most constraining. The Mg II lines are somewhat saturated in that feature, and the surrounding components are, therefore, deeper; this effectively broadens the feature in Mg II, reducing its impact on $\Delta\alpha/\alpha$ here. Nevertheless, Mg I is much weaker in the other features across the profile, so the
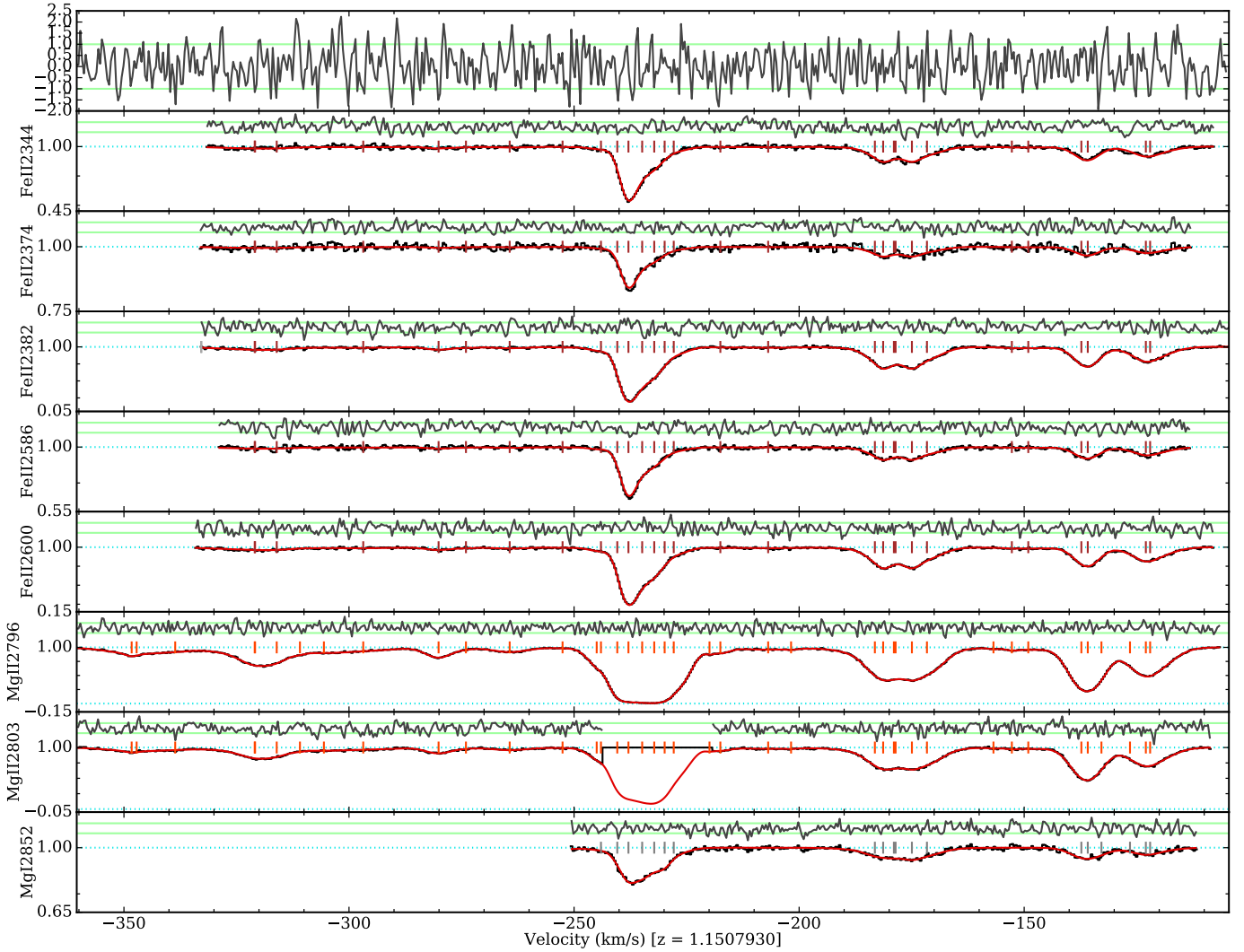
Fig. 8. Same as Fig. 4, but for the central region of the absorber and for all transitions fitted. For clarity, the very weak (formally undetected) Mg I λ2026 line is not plotted but it was included in the fit – see text for discussion.

dominant $\Delta\alpha/\alpha$ constraints arise from the combination of Mg II with Fe II there.

Our fit to the ESPRESSO spectrum required 41 components in the left region, with a final $\chi_\nu^2 = 0.80$, similar to the central region. As with the other two regions, there are no strong outliers revealed by the normalised residual spectra in Fig. 9 and the composite residual spectrum shows no evidence for remaining, unfitted velocity structure. Table 2 again shows that we fitted substantially more velocity components than Kotuš et al. (2017) and, particularly, Milaković et al. (2021), but we consider the reasons discussed for the same situation in the central region (Sect. 3.6 above) to apply here as well.

4. Results

4.1. Results with statistical uncertainties

Table 2 provides the best-fitting values of $\Delta\alpha/\alpha$ and 1σ statistical uncertainties for the three regions of the $z = 1.1508$ absorber. These are the results from VPFIT after minimising χ^2 between the fiducial models (described in Sects. 3.5–3.7) and the combined, unblinded ESPRESSO spectrum, as described in Sect. 3.3. Only the statistical uncertainties are provided here, as output by VPFIT, derived from the final parameter covariance matrix; that is, for a

given model, they are determined by the photon statistical noise in the fitted regions of the spectrum. The systematic error budget is assessed in detail below (Sect. 4.2).

The most important aspect of the results in Table 2 is that $\Delta\alpha/\alpha$ is within the 1σ statistical uncertainty of zero for all three regions. As expected, the right region provides the tightest constraint, with a statistical uncertainty of just 1.5 ppm. The weighted mean from our ESPRESSO spectrum is also within its 1σ uncertainty of zero (i.e. 1.3 ppm). This overall statistical precision is very similar to the ensemble precision obtained from the large samples of absorbers from Keck (1.1 ppm; Murphy et al. 2003, 2004) and VLT (1.2 ppm; Webb et al. 2011; King et al. 2012). As explained in Sect. 1, the high likelihood of large systematic effects in those studies undermines confidence in the results. The more recent 26 measurements from Keck, VLT and Subaru which are corrected for, or insensitive to, these systematic effects also have a similar ensemble precision to our new ESPRESSO measurement (Evans et al. 2014; Murphy et al. 2016; Murphy & Cooksey 2017).

Nevertheless, Table 2 shows that the 1σ uncertainty reported by Kotuš et al. (2017) from their UVES spectrum of HE 0515−4414 is approximately half of ours, despite the S/N per km s^{-1} being similar to that in our ESPRESSO spectrum. This is expected because the ESPRESSO spectrum does not cover
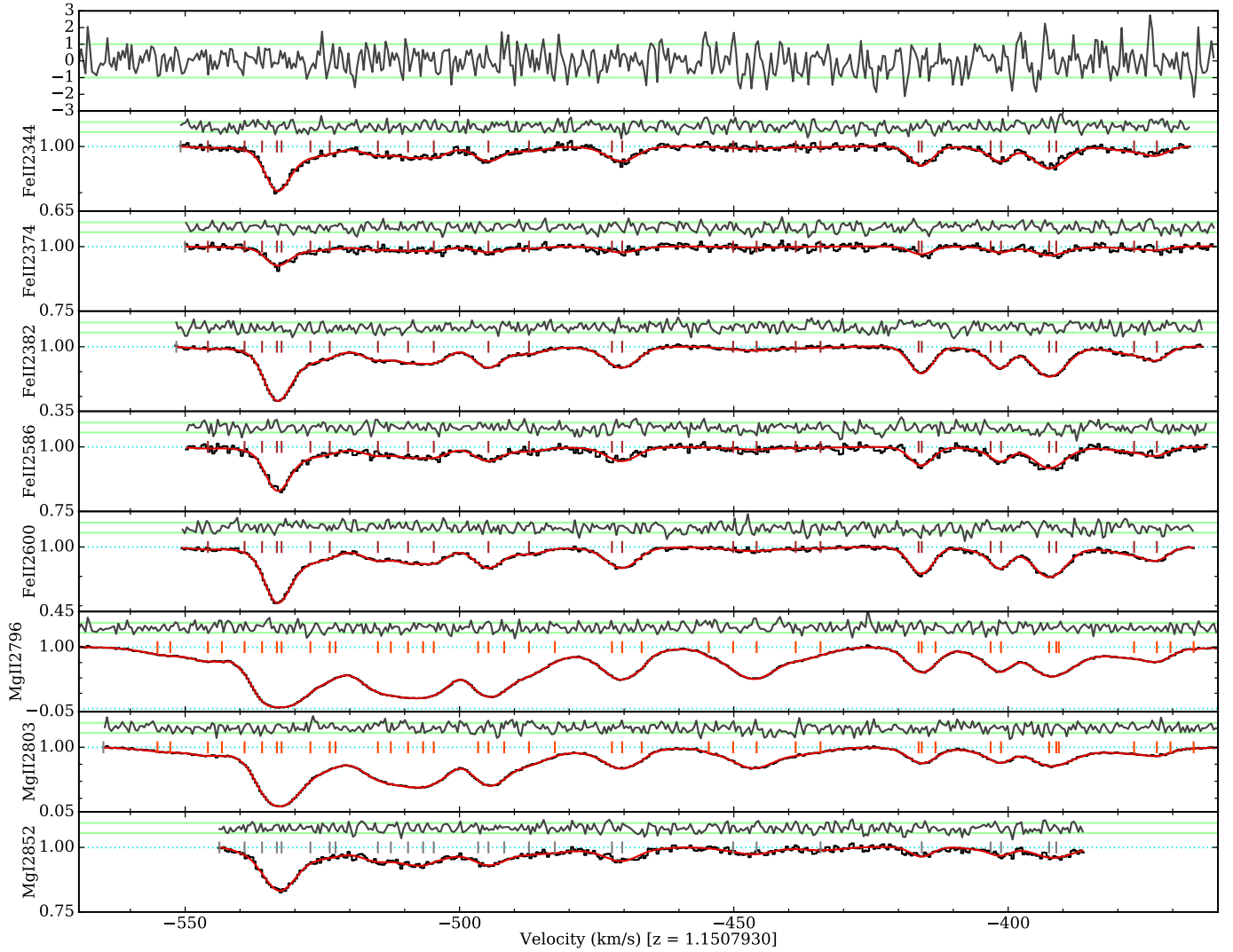
Fig. 9. Same as Fig. 4, but for the left region of the absorber and for all transitions fitted.

the Fe II $\lambda 1608$ transition which provides strong-but-unsaturated absorption in the right region and, most importantly, a large negative Q coefficient. This transition, combined with the other strong Fe II transitions which have large positive coefficients (see Fig. 3), provided the strongest constraint on $\Delta\alpha/\alpha$ in the UVES spectrum. Indeed, Kotuš et al. (2017) predicted that an ESPRESSO spectrum with similar S/N per km s^{-1} as the UVES spectrum would provide a statistical precision of ∼1.8 ppm.

4.2. Systematic errors

While the VPFIT χ^2 minimisation process provides a robust statistical uncertainty for a given model fit, it is clearly important to consider and quantify any systematic effects at the ≤1 ppm level. These have been considered in detail in many previous works (e.g. Murphy et al. 2001a, 2003; Fenner et al. 2005; Levshakov et al. 2005, 2007; Molaro et al. 2008a; Griest et al. 2010; Whitmore et al. 2010; Rahmani et al. 2013; Evans et al. 2014; Whitmore & Murphy 2015). In the slit-based echelle spectrographs on the Keck, VLT and Subaru telescopes, light-path differences between the quasar and ThAr calibration light are most likely responsible for the largest systematic effects: both long-range and intra-order distortions of the wavelength scale. These can produce spurious shifts in $\Delta\alpha/\alpha$ up to ∼10 ppm in

typical absorbers (Evans et al. 2014). When measuring $\Delta\alpha/\alpha$ from a single absorber, as we do here, it is vital that systematic errors are suppressed and any remaining ones assessed. In their UVES study of HE 0515−0515, Kotuš et al. (2017) used the existing HARPS spectrum of the same quasar to correct for the long-range distortions of UVES's wavelength scale. Nevertheless, uncertainties in this recalibration, and the intra-order distortions, meant that wavelength calibration errors remained the dominant factor in their 0.65 ppm systematic error budget. As discussed in Sect. 1, the LFC calibration of ESPRESSO effectively removes this dominant source of systematic uncertainties in $\Delta\alpha/\alpha$ measurements.

The remaining systematic effects in our measurements arise from three aspects of the data processing and analysis: spectral redispersion, profile modelling and convergence of the χ^2 minimisation process. These have been discussed in several previous works, including for HE 0515−4414 in Kotuš et al. (2017), and we discuss and quantify each of them below using the same approaches as that study. The results are summarised in Table 3.

4.2.1. Spectral redispersion effects

To combine the extracted spectra from multiple exposures, UVES_POPLER defines a final (log-linear) wavelength grid onto

Table 3. Summary of $\Delta\alpha/\alpha$ values from this work for the left, central, and right regions, in units of ppm.

Result type	Region			
	Left	Central	Right	Comb.
Fiducial fitting results				
$\Delta\alpha/\alpha$ value	2.17	1.57	1.14	**1.31**
1σ statistical error	3.31	5.59	1.45	**1.29**
Systematic errors				
Redispersion (4.2.1)	0.62	0.10	0.24	0.21
Profile modelling (4.2.2)	1.19	2.26	0.23	0.28
Convergence (4.2.3)	0.18	0.72	0.31	0.25
Combined	1.35	2.37	0.45	**0.43**
Systematic error checks				
LFC vs. ThFP calibration (4.3.1)			<0.02	
Combined LFC+ThFP calibration (4.3.2)			<0.01	
Arm shifts (4.3.3)			<0.14	
Trace shifts (4.3.4)			<0.16	
Isotopic abundance variations (4.3.6)			<0.7	

Notes. The 1σ statistical and systematic error contributions are described in Sects. 4.1 and 4.2, respectively. The combined value of $\Delta\alpha/\alpha$ is the weighted mean from the three regions, with 1σ statistical and systematic error values, as described in Sect. 4.4: $\Delta\alpha/\alpha = 1.31 \pm 1.29_{\text{stat}} \pm 0.43_{\text{sys}}$ ppm. We note that all quantities are quoted with extra significant figures only to allow reproducibility. The checks on additional systematic errors in Sect. 4.3 were conducted in the right region only, with no evidence found that they contribute to the systematic error budget. Section numbers are provided in parentheses for the systematic error checks.

which it redisperses the flux from each trace in each exposure. The phases of the initial and final bins are different, so the fluxes (and uncertainties) in neighbouring pixels are somewhat correlated in the final spectrum. This can lead to small errors in the best-fit centroid position of a spectral feature and, therefore, systematic errors in $\Delta\alpha/\alpha$. To quantify this effect, we created 10 alternative realisations of the combined ESPRESSO spectrum with UVES_POPLER with dispersions between 0.395 and 0.405 km s^{-1} per pixel (i.e. differing by 1 m s^{-1} per pixel) and re-computed $\Delta\alpha/\alpha$. For each realisation, the fiducial model was re-optimised with $\Delta\alpha/\alpha$ initially fixed to the fiducial result (from the 0.4 km s^{-1} per pixel spectrum), then run again with $\Delta\alpha/\alpha$ as a free parameter.

The resulting redispersion uncertainty for each region is provided in Table 3. The systematic uncertainty is taken as the standard deviation in the difference between $\Delta\alpha/\alpha$ for the realisations and the fiducial value for the region. As expected, Table 3 shows that the redispersion errors are small compared to the statistical uncertainties for all regions; for the right region, which is most important for determining $\Delta\alpha/\alpha$, the redispersion effect is ≲20% of the statistical uncertainty. It is possible that the redispersion error is underestimated in the central region (0.1 ppm) because the convergence error was larger for that model – see Sect. 4.2.3 below.

4.2.2. Profile modelling uncertainties

While our fiducial model provides a statistically acceptable fit, we cannot be sure that our fitted velocity structure is unique. This model uniqueness problem is discussed further in Sect. 5.4 after we compare our result with others in the literature (Sect. 5.2).

However, as described in Sect. 3.3 above, we explored an extensive range of alternative models while trying to determine which one provided the lowest χ^2 per degree of freedom for each region. Similar to the approach of Kotuš et al. (2017), we can use this set of alternative fits to estimate the systematic uncertainty in our fiducial $\Delta\alpha/\alpha$ measurements from model non-uniqueness.

For each region, there is an enormous number of possible alternative models, but the vast majority of these will provide much poorer fits to the data than our fiducial one. Therefore, we considered alternative models that contained only one more or fewer velocity component than our fiducial model. We also focussed most attention on the spectral features that most strongly constrained $\Delta\alpha/\alpha$ in each region (see Sect. 3.5). For example, the strong, narrow features near 0 km s^{-1} in the right region were thoroughly explored, while the weak, broad feature at 30−50 km s^{-1} is much less important. Appendix A provides the basic properties of the alternative models in Table A.1. Given our fitting approach (Sect. 3.3), many of these models were not acceptable: they produced large, correlated structures in the composite residual spectra, or were not 'stable'; that is, VPFIT rejected components in one species that were relatively strong in other species. We did not consider these alternative models any further. Finally, we considered the remaining alternative fits that had an Akaike information criterion, corrected for finite sample sizes, AICC (Akaike 1974; Sugiura 1978, as implemented in King et al. 2012) no more than 10 higher than the fiducial model's AICC. According to the Jeffreys' scale for interpreting the AICC, a difference of 10 corresponds to "very strong" evidence against the model with the higher AICC (Jeffreys 1961). This final selection criterion yielded a sample of 5, 6 and 6 alternative models for the right, central, and left regions, respectively.

The profile modelling error derived from the alternative models is shown for each region in Table 3. The systematic uncertainty is taken as the standard deviation in the difference between $\Delta\alpha/\alpha$ for the alternative models and the fiducial value for the region. Remarkably, we find the modelling uncertainty is smaller than the statistical uncertainty in all regions, especially in the right region where the fiducial fit was the most challenging to construct – there it is only ≲20% of the statistical uncertainty. Two additional alternative models for the right region had AICC differences of 11 and 12 – only slightly larger than our accepted threshold of 10 – but relatively small deviations in $\Delta\alpha/\alpha$ from the fiducial value; including them in the results would not significantly alter the systematic uncertainty estimate. We discuss this consistency between alternative models in light of recent literature in Sect. 5.4 below. It is important to emphasise here that all the alternative fits were constructed based on the blinded spectrum, with only the AICC criterion discussed above applied using the final, unblinded χ^2 and AICC values, thereby removing any possible (though highly unlikely) human biases in estimating this systematic uncertainty.

4.2.3. Convergence errors

As can be seen in Figs. 4−9, a very large number of velocity components were required to fit the high S/N ESPRESSO spectrum, many of which are within a resolution element of neighbouring components. This creates degeneracies between their parameters (i.e. their column densities, b parameters and redshifts) in the χ^2 minimisation process: it means χ^2 will have a weak dependence on these parameters, causing the minimisation algorithm to take smaller steps in all model parameters. This can mean the χ^2 difference between iterations can fall below the user-supplied stopping criterion before all parameters have completely reached

their true, best-fit values. We measure the uncertainty this convergence error causes by re-measuring $\Delta\alpha/\alpha$ after offsetting its initial guess from the fiducial value in Table 3 by $\approx 1\sigma_{\mathrm{stat}}$. More specifically, we first re-optimised the fit for a given region with $\Delta\alpha/\alpha$ fixed to its offset starting value, and then re-ran the χ^2 minimisation with $\Delta\alpha/\alpha$ as a free parameter.

Table 3 shows the convergence uncertainty for each region, which was taken as half the difference between the $\Delta\alpha/\alpha$ values measured after offsetting the starting point higher and lower than the fiducial value. For each region, the convergence uncertainty is small compared to the statistical uncertainty for that region. For the most important right region, $\Delta\alpha/\alpha$ in the offset models converged to the fiducial value within 25% of the statistical precision.

4.3. Checks for instrumental and astrophysical systematic effects

Further to the main contributors to the systematic error budget assessed above (Sect. 4.2), we considered several other possible effects below. We do not find evidence that these contribute to the systematic error budget, so we list them only as upper limits in Table 3. Nevertheless, they are important for demonstrating the robustness of the $\Delta\alpha/\alpha$ measurement from this single absorber.

4.3.1. LFC versus ThFP wavelength calibration

In principle, the laser frequency comb (LFC) should provide a much more reliable wavelength calibration for ESPRESSO than the combined ThAr+Fabry–Pérot (ThFP) approach. Each LFC mode's wavelength is known, in absolute terms, with $\ll 1\,\mathrm{cm\,s^{-1}}$ accuracy, while ThAr laboratory wavelengths typically have $\approx 2-20\,\mathrm{m\,s^{-1}}$ uncertainties (Redman et al. 2014) and, potentially, systematic errors that vary with wavelength. The number density of LFC modes also far exceeds that of ThAr lines, and adjacent modes typically have similar intensity, unlike the ThAr lines. However, instrumental effects that differ between the quasar and LFC exposure can still lead to systematic errors in the quasar wavelength scale. In particular, the quasar and calibration light will illuminate the optical fibre differently and, despite significant effort to ensure high optical scrambling efficiency (e.g. octagonal fibres and a double scrambler unit), their point-spread functions (PSFs) on the detector may differ. More importantly, any PSF difference may be a function of position on the detector, leading to intra-order and, possibly, longer-range distortions of the quasar relative wavelength scale and, therefore, systematic effects in $\Delta\alpha/\alpha$.

While we have no direct means to gauge this effect, some insight is gained from Schmidt et al. (2021)'s direct comparison of ESPRESSO's LFC and ThFP calibrations. Their Fig. 15 demonstrates that the two wavelength solutions differ by $0-20\,\mathrm{m\,s^{-1}}$, with $\sim 5\,\mathrm{m\,s^{-1}}$ modulation within the red arm's echelle orders (up to $\sim 15\,\mathrm{m\,s^{-1}}$ in the blue) and larger variations over longer wavelength scales. While the latter could plausibly be due to systematic errors in the ThAr laboratory wavelengths, the former must be due to instrumental effects that differ for the LFC and ThAr light, such as the differential PSF variations described above. It is therefore possible that similar systematics affect the quasar wavelength scale when calibrated with the LFC, as in our quasar exposures.

As a guide to the likely size of this effect on $\Delta\alpha/\alpha$, we re-measured it using ThFP-calibrated quasar exposures. Instead of the LFC calibration, we derived the wavelength solution for each quasar exposure from the standard ThFP calibration taken on the same day. The extracted quasar exposures were combined in exactly the same way as before, and the fiducial profile model for the right region was re-optimised on the new spectrum. The ThFP-calibrated $\Delta\alpha/\alpha$ differed from the LFC-calibrated value by only 0.02 ppm. This clearly demonstrates that our $\Delta\alpha/\alpha$ value is robust to systematic effects in the wavelength calibration of ESPRESSO. We emphasise that this result may not generalise to other ESPRESSO spectra or absorption systems (i.e. comparisons of different wavelength regions); future analyses of other absorbers should explicitly test ESPRESSO's wavelength scale for systematic errors.

4.3.2. Combined LFC+ThFP wavelength calibration

As described in Sect. 2.2, the LFC wavelength calibration only applies to the $\approx 4850-7000\,\text{\AA}$ range, while outside this range the wavelength scale is derived from the combined ThAr + Fabry–Pérot (ThFP) daytime calibration. The Mg I $\lambda 2026$, Zn II and Cr II transitions all fall bluewards of the LFC-calibrated range ($4357-4445\,\text{\AA}$). If the ThFP calibration differs substantially from the LFC one, it is possible that these transitions will be spuriously shifted relative to the other, redder transitions fitted in the right region. To test for this effect, we introduced a further free parameter – a velocity shift – for each fitting region associated with these six bluest transitions. These velocity shifts will be almost entirely degenerate with the constraints on $\Delta\alpha/\alpha$ from these transitions, while still allowing them to help constrain the velocity structure. We find a negligible change in $\Delta\alpha/\alpha$ when introducing these additional free parameters: $\Delta\alpha/\alpha$ increased by just 0.01 ppm relative to the fiducial value in the right region. This is expected because (i) these six weak transitions do not strongly constrain $\Delta\alpha/\alpha$ in this absorber; and (ii) the LFC–ThFP calibration difference is likely very small, $\sim 10\,\mathrm{m\,s^{-1}}$ (Schmidt et al. 2021).

4.3.3. Shifts between ESPRESSO arms

The blue and red arms of ESPRESSO are physically separated, with different collimators, cross-dispersers, cameras and detectors. A dichroic produces a small overlap in wavelength coverage between the two arms, $5191-5272\,\text{\AA}$. In the $z_{\mathrm{abs}} = 1.1508$ absorber, this means most transitions fall in the red arm, with Fe II $\lambda 2382$ and all bluer transitions in the blue arm. To test for any difference in calibration between the two arms, we again introduce a velocity shift parameter for each fitting region of the bluer transitions, as above for the LFC-versus-ThFP test. This means that only the red-arm transitions constrain $\Delta\alpha/\alpha$, while all transitions still constrain the velocity structure like in the fiducial fit. We find only a very small, 0.14 ppm change in $\Delta\alpha/\alpha$ when introducing these changes to the fit of the right region, with a slightly larger statistical uncertainty of $\sigma_{\mathrm{stat}} = 1.6$ ppm. That is, the change in $\Delta\alpha/\alpha$ is $\sim 0.1\sigma_{\mathrm{stat}}$, providing no indication for a systematic difference in wavelength calibration between the two ESPRESSO arms. The right region will be most sensitive to any such effect because it has the best statistical precision and includes more transitions in the blue arm, so we do not extend this test to the other regions.

4.3.4. Shifts between traces

The two traces of light from each fibre on the CCD are produced by an anamorphic pupil slicer, which is the first element

encountered by light emerging from the optical fibres within the spectrograph's vacuum enclosure (Pepe et al. 2021). In comparing the LFC and ThFP calibration solutions (see Sect. 4.3.1 above), Schmidt et al. (2021) found than one trace had considerably larger deviations between the solutions than the other. The effect was much larger in the blue arm, $0-20\,\mathrm{m\,s^{-1}}$, particularly towards the red side of echelle orders, while it remained below $\sim 3\,\mathrm{m\,s^{-1}}$ throughout the red arm. Thus, there appears to be illumination-dependent systematic errors that may have had an impact on our $\Delta\alpha/\alpha$ measurements. We tested for this effect by producing a 'sub-spectrum' of each trace with UVES_POPLER: after both traces were combined to form the main spectrum, we recombined spectra from each trace without modifying any clipped data regions, atmospheric line masking, or continuum fits. See Murphy et al. (2019) for details of this process. The fiducial model of the right region was fitted to the two trace sub-spectra simultaneously, but with free velocity shift parameters allocated to each fitting region of one trace; $\Delta\alpha/\alpha$ was therefore determined entirely by the other trace. Two separate $\Delta\alpha/\alpha$ values were measured in this way, one for each trace, but they differed by only 0.31 ppm with 1.59 ppm statistical uncertainties each (i.e. only a 0.14σ difference). The systematic error in $\Delta\alpha/\alpha$ for the right region is therefore $\lesssim 0.16$ ppm due to shifts between traces.

4.3.5. Velocity shifts between transitions

One of the main assumptions in the MM method for measuring $\Delta\alpha/\alpha$ in quasar absorbers is that the different ionic species share the same velocity structure – that is, that there are no intrinsic velocity shifts between transitions of different species. While this assumption has been questioned (e.g. Levshakov et al. 2005), it appears unlikely that different ions would really be physically separated in individual clouds and, therefore, have velocities differing by $\sim 20\,\mathrm{m\,s^{-1}}$, corresponding to ~ 1 ppm precision in $\Delta\alpha/\alpha$ – see discussion in Kotuš et al. (2017, their Sect. 4.3.2). The difficulty in modelling the velocity structure of the $z_{\mathrm{abs}} = 1.1508$ absorber is more likely to cause spurious shifts between transitions. For example, if too few velocity components have been fitted to a feature, this 'underfitting' will likely introduce shifts, especially between different ionic species but also between transitions of the same species. Small, intrinsic inconsistencies between the absorption profiles of different transitions may also arise if some of the absorbing gas only partially covers the background quasar, or if HE 0515−4414 is gravitationally lensed, with multiple, as-yet-unresolved sight-lines combining to produce the observed spectrum. Previous attempts to assess these effects in slit-based spectrographs (e.g. Kotuš et al. 2017) have been hampered by both the long-range and intra-order distortions of the wavelength scale (see Sect. 1). Clearly, these will both shift different transitions relative to each other depending on where they fall on the CCD mosaic. However, ESPRESSO's high-accuracy wavelength calibration avoids this. Therefore, ESPRESSO offers an important, alternative check on profile modelling errors to those derived above (Sect. 4.2.2).

Figure 10 shows the best-fitted velocity shifts between transitions for the right region. Here, $\Delta\alpha/\alpha$ was fixed to zero in the fiducial model and a velocity shift was introduced as a free parameter for all fitting regions except that for Mg I $\lambda2852$ – all velocity shifts are measured with respect to this transition. We note that, as illustrated in Fig. 5, the Mg I and Zn II $\lambda2026$ transitions share the same fitting region, as do Cr II and Zn II $\lambda2062$, so they also share a common velocity shift mea-
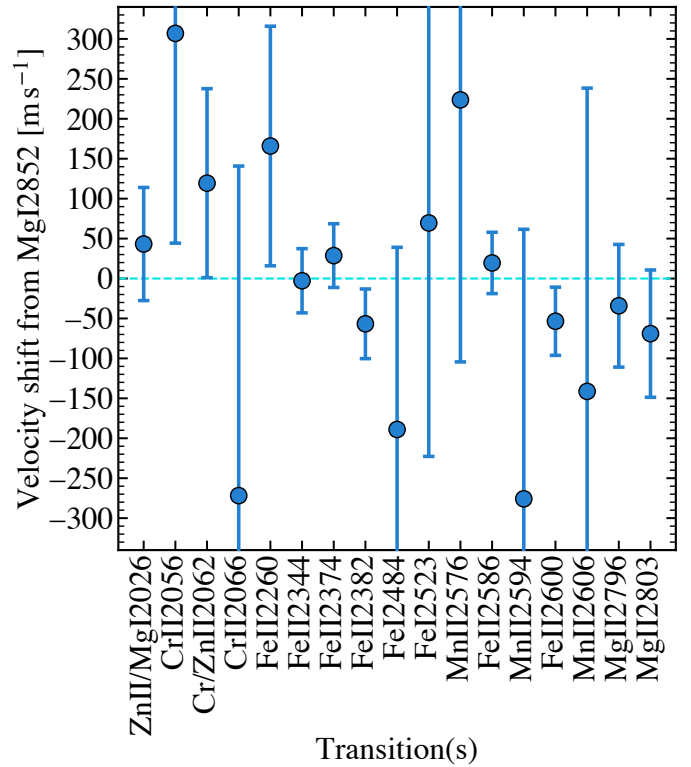
Fig. 10. Velocity shifts between transitions in the right region, measured relative to Mg I $\lambda2852$. $\Delta\alpha/\alpha$ was fixed to zero while a separate velocity shift parameter was introduced for each fitting region in the fiducial fit (Figs. 4 and 5). The best-fit values are shown together with their 1σ statistical uncertainties.

surement. Figure 10 clearly demonstrates that there are no systematic velocity shifts between transitions: all transitions are consistent with Mg I $\lambda2852$ and with each other. The stronger, unsaturated transitions in the right region also provide very precise velocity shift measurements: for example, the five strong Fe II lines only have $\approx 30\,\mathrm{m\,s^{-1}}$ uncertainties. Collectively, there is also no evidence for a shift between species: the weighed mean shift for the Fe II lines is $-6 \pm 16\,\mathrm{m\,s^{-1}}$ from the Mg I $\lambda2852$ line. Taken together, the results in Fig. 10 support the assumptions of the MM method and indicate that our profile fitting approach is robust.

4.3.6. Isotopic abundance variations

Most of the transitions fitted in the $z_{\mathrm{abs}} = 1.1508$ absorber have isotopic structures whose components span velocity ranges of $\approx 0.4\,\mathrm{km\,s^{-1}}$ for Fe II to $\approx 0.5-1\,\mathrm{km\,s^{-1}}$ for the Mg I/II transitions. As explained in Sect. 3.3, we assume the terrestrial isotopic abundances for each element in our analysis. However, if the abundances are different in the $z_{\mathrm{abs}} = 1.1508$ absorber, we will spuriously infer shifts between different transitions and, therefore, in $\Delta\alpha/\alpha$. This is a well-recognised problem (Murphy et al. 2001b; Ashenfelter et al. 2004a,b; Levshakov et al. 2005) but, unfortunately, there are currently no strong and reliable constraints on the isotopic abundances in quasar absorbers (e.g. Agafonova et al. 2011; Levshakov et al. 2009; Webb et al. 2014; Vangioni & Olive 2019; Noterdaeme et al. 2018). For the $z_{\mathrm{abs}} = 1.1508$ absorber, the $\approx 0.5\,\mathrm{km\,s^{-1}}$-wide isotopic structure of Mg I $\lambda2852$ is particularly important. If, for example, the dominant ^{24}Mg isotope (79% terrestrial abundance) was not present in

the absorber and the two heavier isotopes (^{25,26}Mg) were equally abundant, this would cause a $\approx$6 ppm spurious shift in $\Delta\alpha/\alpha$ (e.g. Fenner et al. 2005). Possible isotopic abundance variations are therefore a potentially important systematic effect for our $\Delta\alpha/\alpha$ measurements.

The velocity shift results in Fig. 10 demonstrate that large isotopic abundance variations in the $z_{abs} = 1.1508$ absorber are very unlikely. The Fe II lines are shifted by just -6 ± 16 m s^{-1} relative to the Mg I λ2852 line in the right region (see Sect. 4.3.5 above). For comparison, the velocity between the ^{24}Mg I λ2852 laboratory wavelength and the composite value with terrestrial isotopic abundances is 65 m s^{-1}. So, if the heavier isotopes (^{25,26}Mg) have zero abundance in the absorber, we should see a velocity shift of this order between the Fe II transitions and Mg I λ2852 (assuming $\Delta\alpha/\alpha$ is zero). Similarly, if only the heavier isotopes were present, with equal abundances in the absorber, we would expect to measure a 241 m s^{-1} shift between the Fe II and Mg I λ2852 lines. Clearly, these extreme isotopic abundance variations are not favoured by the results in Fig. 10. However, a robust estimate of the isotopic abundance of Mg in the $z_{abs} = 1.1508$ absorber is not possible with such a simplistic analysis, and we leave this for future work. Nevertheless, given the difference of $\approx$0.04 in Q between the Fe II and Mg I λ2852 transitions (see Fig. 3), the observed shift of -6 ± 16 m s^{-1} between them corresponds to a systematic error in $\Delta\alpha/\alpha$ of <0.7 ppm.

With their high S/N UVES spectrum, Kotuš et al. (2017) were able to strongly constrain the effect of isotopic abundance variations on their measurement of $\Delta\alpha/\alpha$ in the same $z_{abs} = 1.1508$ absorber. They removed the influence of Mg I and II transitions on $\Delta\alpha/\alpha$ by introducing velocity shift parameters (as we did for several tests in this Section), finding a negligible 0.24 ppm change in $\Delta\alpha/\alpha$, with only a marginal increase in the uncertainty. The latter point is due to the availability of the Fe II λ1608 transition in the UVES spectrum: its large negative Q coefficient offers a high contrast with the other Fe II transitions which all have large positive Q values. We cannot perform the same test with our ESPRESSO spectrum because it does not cover the Fe II λ1608 transition (which falls at $\approx$3460 Å, cf. ESPRESSO's blue wavelength limit of $\approx$3800 Å). Nevertheless, the consistency between our $\Delta\alpha/\alpha$ measurement, which depends on the Mg isotopic abundances, and Kotuš et al. (2017)'s which does not (see Sect. 5.2 below) indicates that the Mg isotopic abundance effect is likely smaller than $\approx$1 ppm in our measurement, consistent with the more direct estimate above.

We include our systematic error upper limit of <0.7 ppm in Table 3. As in previous studies (e.g. Kotuš et al. 2017; Milaković et al. 2021), we do not include this astrophysical effect in the formal systematic error budget for $\Delta\alpha/\alpha$. However, it is possible that the ESPRESSO data presented here, and the existing UVES spectrum, could be analysed together to formally, and more robustly, constrain (or perhaps measure) the ^{25,26}Mg/^{24}Mg abundance ratio directly. This would then provide a more stringent measurement or limit on the effect on $\Delta\alpha/\alpha$ in this specific absorber. We defer this to future work.

4.4. Combined result for the entire absorber

Table 3 summarises the fiducial fitting results with the 1σ statistical and systematic uncertainties for each of the three regions. The $\Delta\alpha/\alpha$ values and uncertainties for each region are independent, so straight-forward Gaussian error propagation can be used to compute the weighted mean result for the entire absorber:

$$\Delta\alpha/\alpha = 1.3 \pm 1.3_{stat} \pm 0.4_{sys} \text{ ppm}, \qquad (2)$$

with 1σ statistical and systematic error components. The right region dominates this combined result. In particular, the three main systematic effects (Sects. 4.2.1–4.2.3) cause similar uncertainties in $\Delta\alpha/\alpha$ for the right region, and so make similar contributions to the systematic error combined across the regions.

The result in Eq. (2) is consistent with no change in the fine-structure constant between the absorber at $z_{abs} = 1.1508$ and the current laboratory value. The systematic uncertainty is well below the statistical one, which is remarkable given that the latter is similar to the ensemble precision obtained from previous large samples of absorbers (see Sect. 4.1 for discussion). This is because ESPRESSO's accurate LFC wavelength calibration removes the systematic effects that undermine confidence in those large-sample results. Even the more recent results which were corrected for these effects using asteroid and/or solar twin spectra had residual wavelength calibration uncertainties 0.5−3.5 ppm, larger than our total systematic error budget (e.g. Evans et al. 2014; Murphy & Cooksey 2017). As explained in Sect. 4.2, Kotuš et al. (2017) corrected the wavelength scale of their UVES spectrum of HE 0515−4414 by using the HARPS spectrum of the same quasar. However, this procedure still left a ~0.6 ppm wavelength calibration uncertainty which dominated their systematic error budget. The recent $\Delta\alpha/\alpha$ measurement in the $z_{abs} = 1.1508$ absorber by Milaković et al. (2021) used an LFC-calibrated HARPS spectrum to measure $\Delta\alpha/\alpha$, thereby avoiding wavelength calibration uncertainties like our ESPRESSO spectrum. Of course, the lower S/N of the HARPS spectrum ($\approx$58 per km s^{-1}, cf. 170 for our ESPRESSO spectrum) limited the precision to 2.4 ppm (see further discussion in Sect. 5.2 below).

5. Discussion

5.1. No substantial wavelength calibration errors

As noted in Sect. 1, several of ESPRESSO's design features should strongly suppress systematic errors in $\Delta\alpha/\alpha$ measurements (i.e. relative velocity shifts between transitions at different wavelengths). However, even though the wavelength scale in this study was calibrated with the laser frequency comb (LFC), instrumental effects – particularly illumination-dependent PSF variations across the spectrum – may still affect measurements of $\Delta\alpha/\alpha$ in general (Schmidt et al. 2021). The four tests described in Sects. 4.3.1–4.3.4 all probe these possible wavelength calibration errors. The results are summarised in the lower section of Table 3. For the specific case of the $z_{abs} = 1.1508$ absorber studied here, the effect on any wavelength calibration errors on $\Delta\alpha/\alpha$ is negligible. However, we again emphasise that this may not be true for ESPRESSO studies of other absorbers, particularly those where important transitions fall towards the red edges of echelle orders in the blue arm. We recommend that quasar observations with ESPRESSO be complemented with asteroid or solar twin exposures to check the wavelength scale, especially if the LFC is not operational (e.g. Molaro et al. 2008b; Rahmani et al. 2013; Evans et al. 2014; Whitmore & Murphy 2015).

5.2. Comparison with previous measurements in the $z_{abs} = 1.1508$ absorber

The previous measurements of $\Delta\alpha/\alpha$ at $z_{abs} = 1.1508$ towards HE 0515−4414 are compared to our new value from ESPRESSO in Fig. 11 (blue points). All measurements are consistent with no variation in α, and are also consistent with our result. The early UVES measurements by Quast et al. (2004),
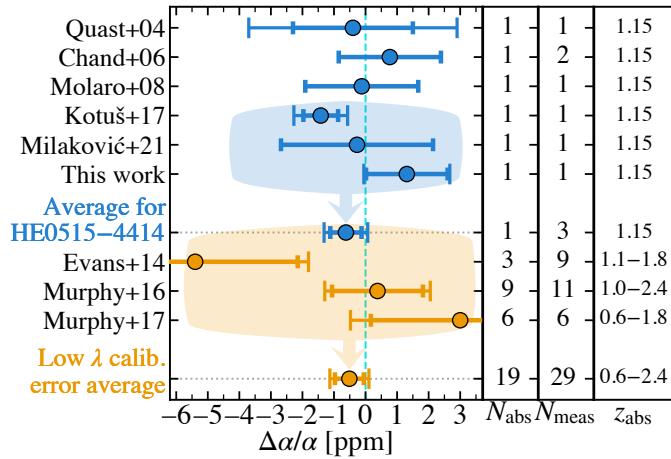
Fig. 11. Measurements of $\Delta\alpha/\alpha$ in the $z_{abs} = 1.1508$ absorber towards HE 0515−4414 (blue points), and recent measurements which were corrected for, or insensitive to, wavelength calibration errors (orange points). The shaded patches show the values included in the weighted mean values indicated by arrows and dotted lines. For each study, the number of absorption systems, N_{abs}, number of independent $\Delta\alpha/\alpha$ measurements (for the entire absorber), N_{meas}, and absorption redshift or range, z_{abs}, are shown in the *right-hand panels*. Shorter, thicker error bars indicate 1σ statistical uncertainties while the longer error bars with taller terminators show the quadrature sum of statistical and systematic uncertainties.

Chand et al. (2006), and Molaro et al. (2008a, who revised their earlier measurement in Levshakov et al. 2006) will all be subject to additional systematic errors, to differing degrees, from the long-range wavelength calibration distortions discussed in Sect. 1 (e.g. Rahmani et al. 2013; Whitmore & Murphy 2015). Kotuš et al. (2017) demonstrated this explicitly, finding that these earlier results most likely require corrections by ∼+2 ppm, similar to their statistical uncertainties. Kotuš et al. (2017) corrected for these effects in their $\Delta\alpha/\alpha$ measurement by directly comparing their UVES spectra to a (ThAr-calibrated) HARPS spectrum of the same quasar. The HARPS ThAr calibration has no evident long-range distortions (Whitmore & Murphy 2015). Milaković et al. (2021) performed the first LFC-calibrated measurement of $\Delta\alpha/\alpha$ with new HARPS observations, with a separate study of wavelength calibration errors demonstrating that they are negligible for that work (Milaković et al. 2020).

Focussing on these three recent, independent results with relatively low wavelength calibration errors, Fig. 11 indicates the formal weighted mean value for the $z_{abs} = 1.1508$ absorber towards HE 0515−4414:

$$\Delta\alpha/\alpha = -0.62 \pm 0.50_{stat} \pm 0.48_{sys}\ \text{ppm}. \tag{3}$$

Again, this is consistent with no variation in α. It is also consistent, at 1.6σ, with the expected value from the α dipole of King et al. (2012), 2.2 ± 1.6 ppm. This value derives from the simplest dipole model where $\Delta\alpha/\alpha = 10.2^{+2.2}_{-1.9}\cos(\Theta)$ ppm for Θ the angle between HE 0515−4414 and the model's pole (RA = 17.4 ± 0.9 h, Dec = $-58 \pm 9°$). The line of sight towards HE 0515−4414's is almost orthogonal to the dipole direction ($\Theta = 78°$), so the model expectation is near zero; that is, HE 0515−4414 does not provide a strong test of the dipole model.

We note that the weighted mean in Eq. (3) is dominated by the Kotuš et al. (2017) measurement from wavelength-corrected UVES spectra. Despite their considerable efforts to estimate

the error budget correctly, it remains plausible that their statistical and systematic uncertainties may be underestimated. For example, the model non-uniqueness problem discussed below (Sect. 5.4) could affect both the statistical uncertainty and contribute a larger systematic error than previously estimated. These problems potentially extend to all measurements, but it is prudent to be cautious when combining results, especially when one dominates. If we exclude that measurement, the weighted mean of Milaković et al. (2021)'s HARPS result and our ESPRESSO measurement is $\Delta\alpha/\alpha = 0.93 \pm 1.14_{stat} \pm 0.33_{sys}$ ppm, where the statistical uncertainty is, of course, larger than in Eq. (3) but the systematic uncertainty is smaller, reflecting the suppression of systematic effects in the two LFC-calibrated measurements.

5.3. Comparison with recent measurements in other absorbers

Figure 11 also compares our result, and the average for the $z_{abs} = 1.1508$ absorber (Eq. (3)), to other recent measurements which had relatively small wavelength calibration uncertainties (orange points). Evans et al. (2014) used asteroid and iodine-cell exposures to correct the wavelength scales of UVES, HIRES, and Subaru/HDS spectra of the same quasar. Similarly, Murphy & Cooksey (2017) corrected HDS spectra of two quasars using exposures of solar twins. Using a different approach, Murphy et al. (2016) measured $\Delta\alpha/\alpha$ using only the (normally weak) Zn/Cr II transitions in "metal-strong" absorbers towards 9 quasars. These ions' transitions have large sensitivity coefficients (Q) with opposite sign, making them sensitive probes of α but, most importantly, the transitions have very similar wavelengths, rendering the $\Delta\alpha/\alpha$ measurements insensitive to long-range calibration distortions. Combining the 28 results of these studies with the weighted mean for $z_{abs} = 1.1508$ towards HE 0515−4414 (Eq. (3)), we obtain

$$\Delta\alpha/\alpha = -0.51 \pm 0.48_{stat} \pm 0.42_{sys}\ \text{ppm}. \tag{4}$$

This weighted mean result represents the most reliable measurement using quasar absorption lines, and is relatively free of the long-range wavelength calibration errors that affected all previous spectra, including the large samples from HIRES (Webb et al. 2001; Murphy et al. 2003, 2004) and UVES (Webb et al. 2011; King et al. 2012). As with Eq. (3), the Kotuš et al. (2017) measurement in the $z_{abs} = 1.1508$ absorber also dominates Eq. (4). Excluding that single result yields a weighted mean of $\Delta\alpha/\alpha = 0.51 \pm 0.82_{stat} \pm 0.37_{sys}$ ppm, which still represents ppm-level precision. Again, as in Sect. 5.2, we urge some caution when combining many results together: if uncorrected systematic errors remain, or some measurements have underestimated total uncertainties, these will significantly affect the weighted mean result.

For 27 of the 29 measurements with low wavelength calibration errors in Fig. 11 – those prior to 2021 – Martins & Pinho (2017) and Murphy & Cooksey (2017) have already noted that they do not support the α dipole model. Ignoring uncertainties in the model itself, those 27 results are inconsistent with it at the $\approx 4\sigma$ level. However, if the dominant Kotuš et al. (2017) measurement is removed, this falls to $\approx 2.5\sigma$ (Murphy & Cooksey 2017). Our new ESPRESSO measurement, and Milaković et al. (2021)'s HARPS measurement, do not appreciably alter this conclusion, given the near orthogonal directions to HE 0515−4414 and the dipole.
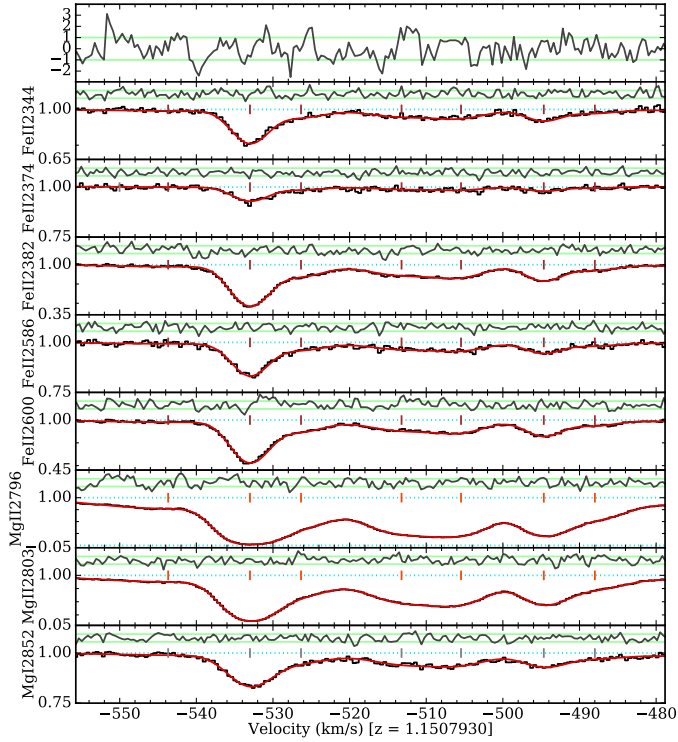
Fig. 12. Same as Fig. 9, but using the velocity structure of Milaković et al. (2021) for their region I (corresponding to part of our left region) which used a compound broadening model (i.e. turbulent and thermal broadening mechanisms). This model was used to fit our ESPRESSO spectrum with VPFIT. We note the significant, correlated structures in the composite residual spectrum (*top panel*) and in individual transitions, especially Mg II λ2796 and Fe II λ2600 near the strongest absorption feature.

5.4. Velocity structure

As shown in Table 2, our fitting approach requires a large number of velocity components to fit all the statistical structure of the ESPRESSO spectrum: 48 in the right region of the $z_{abs} = 1.1508$ absorber, compared to 49 in Kotuš et al. (2017)'s fit to their UVES spectrum, and 26 for Milaković et al. (2021)'s HARPS spectrum. We mentioned in Sect. 3.5 that the former two are likely similar because the ESPRESSO and UVES spectra have similar S/N per km s^{-1} and because the fitting approach was very similar. The HARPS spectrum's S/N per km s^{-1} is much lower than ESPRESSO's ($\approx$58 vs. $\approx$170), so we should expect fewer components to be fitted by Milaković et al. (2021). However, they also used a very different fitting approach: an automated algorithm for determining the velocity structure with VPFIT, called AI-VPFIT (Lee et al. 2021a), and both turbulent and thermal broadening for each component's b parameter ("compound" broadening). As a simple illustration, Fig. 12 shows the compound model of Milaković et al. (2021) for part of their region I (corresponding to part of our left region), with χ^2 minimised using VPFIT on our ESPRESSO spectrum. It is immediately clear that this fit does not adequately account for the statistical structure in the data: for example, the composite residual spectrum shows several systematic excursions, correlated over many pixels, particularly around the strongest absorption at −540 to −525 km s^{-1}, but at other velocities as well. This clearly demonstrates that the larger S/N per km s^{-1} of the ESPRESSO spectrum drives the need for a more complex fit; differences in the fitting approach cannot fully explain this.

Lee et al. (2021b) explored how the AI-VPFIT approach produced different velocity structures from different starting points of the algorithm and different broadening model assumptions, and how $\Delta\alpha/\alpha$ varied between these models. A similar "non-uniqueness" problem arises for the more interactive, human-guided approach we have employed: different people will choose to introduce new components at different velocities, with different initial parameters, while building up a complex fit, and so produce somewhat different velocity structures for the same dataset. We have attempted to ensure our measurement is robust to this effect by thoroughly exploring alternative velocity structures and incorporating their variance in $\Delta\alpha/\alpha$ into our error budget – see Sect. 4.2.2. However, we limited that search to adding or subtracting one component from our fiducial model, as the sheer number of other possibilities is not possible to explore in this way. Lee et al. (2021b) find that turbulent-only models within AI-VPFIT produce a larger variety of fits, with a larger range of $\Delta\alpha/\alpha$ values than the compound broadening approach. While this is certainly a concern, it is important to recognise that this is unlikely to apply simplistically outside AI-VPFIT itself. Our turbulent-only fitting approach maximises the number of components using χ_ν^2 as the information criterion, while the addition of new components – where in the profile they are added, and with what initial parameters – is judged by the human user. Importantly, that judgement seeks to avoid leaving inconsistencies in the model between different ions and/or transitions. Currently, AI-VPFIT appears to solve such potential inconsistencies, at least sometimes, by adding blending components (i.e. from an unidentified species in an unidentified absorber) in some fitting regions; these random blends differ when the algorithm is started with different random seeds – see Figs. A.1–A.3 of Lee et al. (2021b) which cover the same fitting range as Fig. 12. A full AI-VPFIT analysis of our ESPRESSO spectrum would be very welcome: we make our ESPRESSO spectrum publicly available in Murphy et al. (2021) online for others to analyse, or we will undertake this if AI-VPFIT becomes available.

6. Conclusions

In this work we have made the first precise measurement of $\Delta\alpha/\alpha$ with the new ESPRESSO spectrograph on the VLT. HE 0515−4414 was observed for 16.1 h to achieve a very high S/N ($\approx$105 per 0.4 km s^{-1} pixel at 6000 Å) at resolving power $R \approx 145\,000$, with the wavelength scale of all 17 exposures calibrated using the laser frequency comb (LFC). The entire analysis procedure was developed using a blinded version of the spectrum to avoid human biases; once the blinding was removed, the analysis was re-run without alteration or human interaction. The high R and S/N revealed that the $z_{abs} = 1.1508$ absorption system was more complex than previous studies found: strong constraints on the relative optical depths of two different Mg I lines confirmed the presence of very narrow velocity components ($b < 0.5$ km s^{-1}) in the strongest absorption feature (Sect. 3.5). A total of 129 velocity components were required to fit the $\approx$720 km s^{-1}-wide absorption profile, with the strongest Mg I and Fe II transitions providing the main constraints on $\Delta\alpha/\alpha$.

Our new measurement at $z_{abs} = 1.1508$ from the ESPRESSO spectrum is

$$\Delta\alpha/\alpha = 1.3 \pm 1.3_{stat} \pm 0.4_{sys} \text{ ppm}, \tag{5}$$

consistent with no cosmological variation in the fine-structure constant. The total error is similar to the ensemble precision of the previous, large samples of absorbers from Keck/

HIRES and VLT/UVES that indicated variations at the ≈ 5 ppm level (Webb et al. 2001, 2011; Murphy et al. 2003, 2004; King et al. 2012), which likely arose from long-range distortions in the wavelength scale (e.g. Rahmani et al. 2013; Whitmore & Murphy 2015). Our measurement is free from those systematic errors due to the specific design features of ESPRESSO to suppress them (e.g. octagonal fibre feed, stable vacuum environment, and LFC calibration). However, we emphasise that wavelength calibration errors are still possible with ESPRESSO, and future studies of $\Delta\alpha/\alpha$ should specifically test for them, especially when important transitions fall in the blue arm (Sect. 5.1). The main systematic uncertainties in our measurement arise from ambiguities in fitting the absorption profile, effects from redispersion of the spectra, and convergence of the fitting procedure (Sect. 4.2).

HE 0515−4414 is very bright (*Gaia G* = 14.9 mag), which has enabled many previous measurements of $\Delta\alpha/\alpha$ in the z_{abs} = 1.1508 absorber. Our new $\Delta\alpha/\alpha$ measurement in Eq. (5) is consistent with the two most recent of those from two different spectrographs: the first LFC-calibrated measurement, from HARPS (Milaković et al. 2021, $\Delta\alpha/\alpha$ = −0.3 ± 2.4 ppm), and the most precise measurement in a single absorber, from a very high S/N UVES spectrum (Kotuš et al. 2017, $\Delta\alpha/\alpha$ = −1.4 ± 0.6$_{stat}$ ± 0.6$_{sys}$ ppm). Both are relatively free from wavelength calibration uncertainties (see Sect. 5.2) and can be combined with our ESPRESSO measurement – see (Eq. (3)):

$$\Delta\alpha/\alpha = -0.6 \pm 0.5_{stat} \pm 0.5_{sys} \text{ ppm.}$$

A sample of 26 other measurements with HIRES, UVES, and Subaru/HDS which have been corrected for, or are insensitive to, long-range wavelength distortions are also consistent with our new measurement and have a similar overall precision: $\Delta\alpha/\alpha$ = 0.4 ± 1.3$_{stat}$ ± 0.7$_{sys}$ ppm (Evans et al. 2014; Murphy et al. 2016; Murphy & Cooksey 2017). Combining the 29 independent, low-wavelength calibration error measurements of $\Delta\alpha/\alpha$ at z_{abs} = 0.6−2.4 above yields a weighted mean (Eq. (4)):

$$\Delta\alpha/\alpha = -0.5 \pm 0.5_{stat} \pm 0.4_{sys} \text{ ppm.}$$

Figure 11 summarises these comparisons with previous measurements.

ESPRESSO promises to provide a larger sample of well-calibrated, high quality quasar absorption spectra for measuring $\Delta\alpha/\alpha$ through the instrument consortium's Guaranteed Time Observations and open, competitive observing time. One important, outstanding problem in improving quasar absorption measurements of $\Delta\alpha/\alpha$ is the lack of observational constraints on how the isotopic abundances in the absorbers differ from the terrestrial values. The overall velocity shifts between Mg I $\lambda 2852$ and the strong Fe II lines in our ESPRESSO spectrum limit the effect on our measurement of $\Delta\alpha/\alpha$ to <0.7 ppm (Sect. 4.3.6). It may be possible to directly measure the Mg isotopic abundance ratio in this absorber with our spectrum, given ESPRESSO's excellent wavelength calibration accuracy and, in principle at least, detailed map of the instrumental line-shape function provided by the LFC and Fabry–Pérot calibration exposures. However, given its complex absorption profile, a robust isotopic ratio estimate in the z_{abs} = 1.1508 absorber will likely require an even more extensive exploration of alternative velocity structures than we have performed here.

References

Agafonova, I. I., Molaro, P., Levshakov, S. A., & Hou, J. L. 2011, A&A, 529, A28

Akaike, H. 1974, IEEE Trans. Autom. Control, 19, 716

Aldenius, M. 2009, Phys. Scr. Vol. T, 134, 014008

Ashenfelter, T., Mathews, G. J., & Olive, K. A. 2004a, Phys. Rev. Lett., 92, 041102

Ashenfelter, T. P., Mathews, G. J., & Olive, K. A. 2004b, ApJ, 615, 82

Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33

Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123

Bainbridge, M. B., & Webb, J. K. 2017, MNRAS, 468, 1639

Barlow, M. J., Crawford, I. A., Diego, F., et al. 1995, MNRAS, 272, 333

Blackwell-Whitehead, R. J., Toner, A., Hibbert, A., Webb, J., & Ivarsson, S. 2005, MNRAS, 364, 705

Carswell, R. F., & Webb, J. K. 2014, Astrophysics Source Code Library [record ascl:1408.015]

Cersullo, F., Coffinet, A., Chazelas, B., Lovis, C., & Pepe, F. 2019, A&A, 624, A122

Chand, H., Srianand, R., Petitjean, P., & Aracil, B. 2004, A&A, 417, 853

Chand, H., Srianand, R., Petitjean, P., et al. 2006, A&A, 451, 45

Dumont, V., & Webb, J. K. 2017, MNRAS, 468, 1568

Dzuba, V. A., Flambaum, V. V., & Webb, J. K. 1999a, Phys. Rev. A, 59, 230

Dzuba, V. A., Flambaum, V. V., & Webb, J. K. 1999b, Phys. Rev. Lett., 82, 888

Dzuba, V. A., Flambaum, V. V., Kozlov, M. G., & Marchenko, M. 2002, Phys. Rev. A, 66, 022501

Evans, T. M., Murphy, M. T., Whitmore, J. B., et al. 2014, MNRAS, 445, 128

Fenner, Y., Murphy, M. T., & Gibson, B. K. 2005, MNRAS, 358, 468

Griest, K., Whitmore, J. B., Wolfe, A. M., et al. 2010, ApJ, 708, 158

Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90

Jeffreys, H. 1961, The Theory of Probability, 3rd edn. (Oxford: Oxford University Press)

Jones, A., Noll, S., Kausch, W., Szyszka, C., & Kimeswenger, S. 2013, A&A, 560, A91

King, J. A., Webb, J. K., Murphy, M. T., et al. 2012, MNRAS, 422, 3370

Kotuš, S. M., Murphy, M. T., & Carswell, R. F. 2017, MNRAS, 464, 3679

Lange, R., Huntemann, N., Rahm, J. M., et al. 2021, Phys. Rev. Lett., 126, 011102

Lee, C.-C., Webb, J. K., Carswell, R. F., & Milaković, D. 2021a, MNRAS, 504, 1787

Lee, C.-C., Webb, J. K., Milaković, D., & Carswell, R. F. 2021b, MNRAS, 507, 27

Levshakov, S. A., Centurión, M., Molaro, P., & D'Odorico, S. 2005, A&A, 434, 827

Levshakov, S. A., Centurión, M., Molaro, P., et al. 2006, A&A, 449, 879
Levshakov, S. A., Molaro, P., Lopez, S., et al. 2007, A&A, 466, 1077
Levshakov, S. A., Agafonova, I. I., Molaro, P., Reimers, D., & Hou, J. L. 2009, A&A, 507, 209
Malec, A. L., Buning, R., Murphy, M. T., et al. 2010, MNRAS, 403, 1541
Martins, C. J. A. P. 2017, Rep. Progr. Phys., 80, 126902
Martins, C. J. A. P., & Pinho, A. M. M. 2017, Phys. Rev. D, 95, 023008
Milaković, D., Pasquini, L., Webb, J. K., & Lo Curto, G. 2020, MNRAS, 493, 3997
Milaković, D., Lee, C.-C., Carswell, R. F., et al. 2021, MNRAS, 500, 1
Molaro, P. 2009, Astrophys. Space Sci. Proc., 9, 389
Molaro, P., Murphy, M. T., & Levshakov, S. A. 2006, in The Scientific Requirements for Extremely Large Telescopes, eds. P. Whitelock, M. Dennefeld, & B. Leibundgut, 232, 198
Molaro, P., Reimers, D., Agafonova, I. I., & Levshakov, S. A. 2008a, Eur. Phys. J. Spec. Top., 163, 173
Molaro, P., Levshakov, S. A., Monai, S., et al. 2008b, A&A, 481, 559
Molaro, P., Centurión, M., Whitmore, J. B., et al. 2013a, A&A, 555, A68
Molaro, P., Esposito, M., Monai, S., et al. 2013b, A&A, 560, A61
Murphy, M. T. 2016, https://doi.org/10.5281/zenodo.44765
Murphy, M. T., & Berengut, J. C. 2014, MNRAS, 438, 388
Murphy, M. T., & Cooksey, K. L. 2017, MNRAS, 471, 4930
Murphy, M. T., Webb, J. K., Flambaum, V. V., Churchill, C. W., & Prochaska, J. X. 2001a, MNRAS, 327, 1223
Murphy, M. T., Webb, J. K., Flambaum, V. V., et al. 2001b, MNRAS, 327, 1208
Murphy, M. T., Webb, J. K., & Flambaum, V. V. 2003, MNRAS, 345, 609
Murphy, M. T., Flambaum, V. V., Webb, J. K., et al. 2004, in Constraining Variations in the Fine-Structure Constant, Quark Masses and the Strong Interaction, eds. S. G. Karshenboim, & E. Peik, 648, 131
Murphy, M. T., Webb, J. K., & Flambaum, V. V. 2007a, Phys. Rev. Lett., 99, 239001
Murphy, M. T., Udem, T., Holzwarth, R., et al. 2007b, MNRAS, 380, 839
Murphy, M. T., Webb, J. K., & Flambaum, V. V. 2008, MNRAS, 384, 1053
Murphy, M. T., Malec, A. L., & Prochaska, J. X. 2016, MNRAS, 461, 2461
Murphy, M. T., Kacprzak, G. G., Savorgnan, G. A. D., & Carswell, R. F. 2019, MNRAS, 482, 3458
Murphy, M. T., Molaro, P., Leite, A. C. O., & the ESPRESSO Science Team 2021, https://doi.org/10.5281/zenodo.5512490
Nave, G., & Sansonetti, C. J. 2011, J. Opt. Soc. Am. B Opt. Phys., 28, 737
Nave, G., Johansson, S., Learner, R. C. M., Thorne, A. P., & Brault, J. W. 1994, ApJS, 94, 221
Noll, S., Kausch, W., Barden, M., et al. 2012, A&A, 543, A92
Noterdaeme, P., Ledoux, C., Zou, S., et al. 2018, A&A, 612, A58
Olive, K. A., Peloso, M., & Uzan, J.-P. 2011, Phys. Rev. D, 83, 043509
Pepe, F., Cristiani, S., Rebolo, R., et al. 2021, A&A, 645, A96
Quast, R., Baade, R., & Reimers, D. 2002, A&A, 386, 796
Quast, R., Reimers, D., & Levshakov, S. A. 2004, A&A, 415, L7
Rahmani, H., Wendt, M., Srianand, R., et al. 2013, MNRAS, 435, 861
Redman, S. L., Nave, G., & Sansonetti, C. J. 2014, ApJS, 211, 4
Reimers, D., Hagen, H. J., Rodriguez-Pascual, P., & Wisotzki, L. 1998, A&A, 334, 96
Reimers, D., Baade, R., Quast, R., & Levshakov, S. A. 2003, A&A, 410, 785
Rosenband, T., Hume, D. B., Schmidt, P. O., et al. 2008, Science, 319, 1808
Schmidt, T. M., Molaro, P., Murphy, M. T., et al. 2021, A&A, 646, A144
Sugiura, N. 1978, Commun. Stat. – Theor. Meth., 7, 13
Vangioni, E., & Olive, K. A. 2019, MNRAS, 484, 3561
Webb, J. K., Flambaum, V. V., Churchill, C. W., Drinkwater, M. J., & Barrow, J. D. 1999, Phys. Rev. Lett., 82, 884
Webb, J. K., Murphy, M. T., Flambaum, V. V., et al. 2001, Phys. Rev. Lett., 87, 091301
Webb, J. K., King, J. A., Murphy, M. T., et al. 2011, Phys. Rev. Lett., 107, 191101
Webb, J. K., Wright, A., Koch, F. E., & Murphy, M. T. 2014, Mem. Soc. Astron. It., 85, 57
Whitmore, J. B., & Murphy, M. T. 2015, MNRAS, 447, 446
Whitmore, J. B., Murphy, M. T., & Griest, K. 2010, ApJ, 723, 89
Wilczynska, M. R., Webb, J. K., King, J. A., et al. 2015, MNRAS, 454, 3082
Wilczynska, M. R., Webb, J. K., Bainbridge, M., et al. 2020, Sci. Adv., 6, eaay9672

Appendix A: Alternative profile models

Here we detail the alternative models used to derive an estimate of the systematic errors from profile model non-uniqueness in Section 4.2.2. Table A.1 identifies the velocity at which components were removed from, or added to the fiducial models for the right, central, and left regions plotted in Figs. 4, 8, and 9 respectively. In general, we focused on components in the spectral features that most strongly influence $\Delta\alpha/\alpha$. We also investigated other features where two components were fitted close to each other – we removed one of them to assess whether it was really statistically required. The stability, or otherwise, of each model is noted in the table. We note that Table A.1 does not show all models attempted: many models were initially investigated which, upon preliminary analysis, were either highly unstable or produced substantially larger χ^2 per degree of freedom (χ^2_ν) than the fiducial model and/or large deviations in the composite residual spectra; these were discarded without further consideration.

For the remaining models, $\Delta\alpha/\alpha$ was measured in the same way as for the fiducial model (Section 3.3), first using the blinded spectrum and again after unblinding (see Section 3.4). Initially, $\Delta\alpha/\alpha$ was fixed at the fiducial value while the alternative model was constructed and refined by minimising χ^2 in VPFIT. Then VPFIT was re-run with $\Delta\alpha/\alpha$ as a free parameter. Table A.1 shows how $\Delta\alpha/\alpha$ deviates from the fiducial value for the region, and the increase in χ^2 and the Akaike information criterion, corrected for finite sample sizes, AICC, as a result. We note that some stable alternative models have a lower AICC than the fiducial model. At first, it may seem that one of these models was, therefore, more appropriate to adopt as fiducial. However, this is not correct for two reasons: (i) the fiducial model was selected while the analysis was blinded, so the χ^2_ν and AICC values were based on the blinded spectrum, while those in Table A.1 are from the final, unblinded analysis; and (ii) we used χ^2_ν as a simpler information criterion for deciding which model should be the fiducial one, while AICC is perhaps more justified for comparing very similar models as we do in this systematic error analysis. The most important aspects of the alternative models are discussed below for each region.

A.1. Right region

The spectral features that most strongly influence $\Delta\alpha/\alpha$ are those at 0 and 9 km s^{-1}, and many components were already required for our fiducial model in these features, particularly the former (five components). We already discussed in Section 3.5 that no reasonable fit could be constructed without the very narrow 'cold' components in the 0 km s^{-1} feature. However, removing a single broader component in the 0 km s^{-1} feature, like that at -3 km s^{-1}, results in a much higher AICC. The only velocity at which a new component could be added to these features was at 14 km s^{-1} which produced a somewhat higher AICC (12.4) but very little change in $\Delta\alpha/\alpha$ (0.3 ppm). In the rest of the profile, a 0.4 ppm change in $\Delta\alpha/\alpha$ resulted from removing one of the two components in the sharp right-hand edge of the 75 km s^{-1} feature. This feature is saturated in Mg II and the two components in the fiducial fit were found to provide a marginally smaller AICC (by 0.4).

A.2. Central region

The structure of the most important spectral features for measuring $\Delta\alpha/\alpha$, at -237 and -140 to -120 km s^{-1}, were thoroughly explored in this region. In the former, removing one of the

Table A.1. Alternative profile models attempted in each region.

Velocity (km s^{-1})	Added or removed?	Stable?	$\Delta(\Delta\alpha/\alpha)$ (ppm)	$\Delta\chi^2$	Δ(AICC)
Right region					
-38	$+$	Yes	-0.24	-8.8	-0.1
-24	$+^a$	Yes	0.02	4.3	6.5
-3	$-$	Yes	-2.28	53.4	31.5
6	$-$	No			
14	$+$	Yes	0.26	-0.7	12.4
16	$-$	No			
18	$+$	Yes	0.31	-0.8	7.9
29	$+$	Yes	-0.15	0.0	11.0
52	$-$	No			
55	$-$	No			
60	$-$	Yes	0.04	15.7	4.7
76	$-$	Yes	0.42	15.7	0.4
Central region					
-321	$-$	Yes	0.70	11.9	5.5
-244	$-$	Yes	0.07	2.6	2.6
-238	$-$	No			
-232	$-$	Yes	5.06	15.9	5.1
-219	$+$	Yes	-1.26	-9.3	5.8
-218	$-$	Yes	-0.11	10.7	8.7
-179	$-$	No			
-127	$-$	Yes	-2.00	2.1	-6.6
-123	$-$	No			
Left region					
-543	$-$	Yes	0.02	10.4	3.8
-536	$-$	Yes	-0.37	23.6	12.5
-523	$-$	Yes	-0.20	9.6	-1.4
-513	$-$	Yes	0.10	20.0	6.8
-507	$-$	Yes	-0.07	12.8	1.8
-492	$-$	Yes	1.80	5.2	-10.2
-467	$-$	No			
-416	$-$	Yes	-2.99	87.1	76.1
-391	$-$	Yes	1.58	10.2	3.6

Notes. The velocity at which a component was added ($+$) or removed ($-$) is measured relative to the absorption redshift, $z_{\rm abs} = 1.150793$. A model is deemed unstable if VPFIT rejected important components in one species that were relatively strong in other species. For stable fits, the final two columns provide the change in $\Delta\alpha/\alpha$ and the Akaike information criterion, corrected for finite sample sizes, AICC, relative to the fiducial model in Figs. 4, 8, and 9. [a] The component at -24 km s^{-1} was only added to Fe II as it already exists in Mg I/II in the fiducial model (see Fig. 4). While this model is listed as stable here, this additional Fe II component's column density slowly reduces over repeated runs of VPFIT, with χ^2 barely changing, and it is rejected eventually.

reddest 3 components leaves a stable, statistically acceptable fit with an AICC 'only' 5 more than the fiducial fit. This fit leads to quite a large change in $\Delta\alpha/\alpha$, 5 ppm, approximately the same as the statistical uncertainty in this region. While the AICC change of 5 is within our accepted AICC change threshold of 10, it is still regarded as "strong" evidence against this model (Jeffreys 1961). Nevertheless, we include it in our estimate for profile modelling systematic errors in this region. For the double feature between -140 and -120 km s^{-1}, removing the component at -127 km s^{-1} in Mg I/II results in a lower AICC, but $\Delta\alpha/\alpha$ only changes by 2 ppm. Although the feature at -179 km s^{-1} is relatively broad and does not constrain $\Delta\alpha/\alpha$ strongly, it is interesting that removing one of the two components at its centre

resulted in an unstable fit, causing other surrounding components to be rejected by VPFIT in some species.

A.3. Left region

$\Delta\alpha/\alpha$ is mainly constrained by the features at −533, −495 and −470 km s^{-1}, and the three between −420 and −380 km s^{-1}. Table A.1 shows that the velocity structure fitted to the strongest of these, at −533 km s^{-1}, seems to have little effect on $\Delta\alpha/\alpha$, while that fitted to the −495 and −390 km s^{-1} features is less well determined – shifts in $\Delta\alpha/\alpha$ of $\approx$1.5–2 ppm are found for alternative models which have similar AICC values as the fiducial model. We found no alternative stable fits with more components than the fiducial model.