Fundamental physics and cosmology using astronomical laser frequency combs

Dinko Milaković

München 2020

Fundamental physics and cosmology using astronomical laser frequency combs

Dinko Milaković

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Dinko Milaković
aus Zagreb, Kroatien

München, den 9 September 2020

Contents

List of Figures

List of Tables

Zusammenfassung

Diese Dissertation präsentiert neue Erkenntnisse in Methoden zur Kalibration der Wellenlängen astronomischer Spektrographen. Der Kontext dieser Forschung ist die Suche nach neuen physikalischen Zusammenhängen anhand astronomischer Beobachtungen von Quasaren. Ziel ist es, die bestmögliche Genauigkeit bei Messungen der Wellenlänge mittels eines Laserfrequenzkamms ("laser frequency comb" oder LFC) zu erreichen. Sehr genaue Messungen der Wellenlänge sind von Bedeutung für verschiedene wissenschaftliche Projekte, welche den Bau großer optischer und astronomischer Observatorien, wie dem Extremely Large Telescope, erfordern.

In vielen Verallgemeinerungen der allgemeinen Relativitätstheorie ändern sich die Werte fundamentaler Konstanten wie der Feinstrukturkonstante (α). Dieser Effekt ist messbar, da eine Änderung von α die Energielevel atomarer und molekularer Übergänge beeinflusst. Kosmologische Änderungen im Wert von α können durch die Messung kleiner Verschiebungen der Wellenlänge von Übergängen in Spektra von Quasaren überprüft werden. Bisherige Suchen nach einem solchen Effekt fanden Hinweise für eine Dipolschwankung von α über den Himmel, mit einer Signifikanz von $\approx 4\sigma$. Falls sie sich bewahrheitet, könnte dieses Ergebnis den Weg zu einer großen vereinheitlichten Theorie ebnen. Es ist daher unumgänglich, alle systematischen Fehler, welche Messungen von α beeinträchtigen, zu beheben. Ein solcher Fehler betrifft die Kalibration der Wellenlängen astronomischer Spektrographen.

Im Folgenden präsentiere ich Forschungen, die mit der ersten Generation astronomischer LFC an einem der präzisesten astronomischen Spektrographen, dem High Accuracy Radial-velocity Planet Searcher (HARPS), durchgeführt wurden. Die Resultate sind sehr vielversprechend: LFCs liefern beispiellose Genauigkeit bei der Kalibration der Wellenlänge. Die zahlreichen und dicht gestaffelten Frequenzlinien der LFCs erlauben es, praktisch alle Kalibrationsfehler zu beheben und läuten eine neue Ära für Messungen von α bei hoher Rotverschiebung ein. Die Frequenzlinien ermöglichen zusätzlich die Messung von Detektoreigenschaften auf eine neue, zuvor nicht realisierbare Art und Weise.

Die Anwendung dieser verbesserten Methoden zur Kalibration der Wellenlänge auf HARPS-Beobachtungen des Quasars HE0515$-$4414 erlaubt es, die Änderung von α in einem Absorptionssystem bei $z_{abs} = 1.15$, welches zwischen Beobachter und Quasar gelegen ist, einzuschränken. Unterstützt durch künstliche Intelligenz erreichen wir 40 Messungen der relativen Änderung von α in diesem System. Der Mittelwert der Messungen ist $\Delta\alpha/\alpha = -0.27 \pm 2.41 \times 10^{-6}$, konsistent mit der Vorhersage des Dipols für die Position des Quasars am Himmel, aber ebenfalls konsistent mit keiner Änderung. Die große Anzahl an Messungen in diesem System ermöglicht es uns, Änderungen von α auf kleinen Skalen zum ersten Mal einzuschränken: $\Delta\alpha/\alpha \lesssim 9 \times 10^{-5}$ über Entfernungen von $\approx 25\,\mathrm{kpc}$.

Abstract

This thesis presents advances in procedures for wavelength calibration of astronomical spectrographs. The context of this research is a search for new physics using astronomical observations of quasars. The aim is to reach the highest possible accuracy of wavelength measurements using laser frequency comb technology. Very accurate wavelength measurements are important for several scientific projects motivating the construction of large optical astronomical facilities, such as the Extremely Large Telescope.

In many extensions of General Relativity, the values of fundamental constants, such as the fine structure constant (α), change. This effect is measurable, as a change in α perturbs the energy levels of atomic and molecular transitions. Cosmological changes in the value of α can be probed by measuring small wavelength shifts of transitions in quasar spectra. Previous searches for such an effect found evidence for a dipole variation of α across the sky with $\approx 4\sigma$ significance. If proven to be correct, this result could possibly pave the way to a Grand Unified Theory. It is therefore imperative to remove all systematic effects which may spoil α measurements. One such effect pertains the wavelength calibration of astronomical spectrographs.

Here, I present research performed using the first generation of astronomical laser frequency combs (LFC) on one of the most stable astronomical spectrographs: the High Accuracy Radial-velocity Planetary Searcher (HARPS). The results are highly encouraging; the LFCs provide unprecedented wavelength calibration accuracy. The numerous and densely spaced LFCs lines allow practically all wavelength calibration effects to be removed, starting a new era for α measurements at high redshift. The LFC lines also allowed detector characteristics to be measured in a way that was not possible before.

Applying these advanced wavelength calibration methods to HARPS observations of the quasar HE0515$-$4414 allowed us to constrain variations in α in an absorption system at $z_{abs} = 1.15$ seen towards this quasar. We obtained 40 measurements of the fractional change in α in this system, aided by artificial intelligence methods. The average of the measurements is $\Delta\alpha/\alpha = -0.27 \pm 2.41 \times 10^{-6}$, consistent with the prediction of the dipole for the sky position of this quasar but also consistent with zero change. The large number of measurements in this system allowed us to constrain small-scale variations of α for the first time: $\Delta\alpha/\alpha \lesssim 9\times10^{-5}$ across $\approx 25\,\mathrm{kpc}$ scales.

Chapter 1

Introduction

This thesis is concerned with achieving the highest possible accuracy for wavelength measurements. The context of the research is the measurement of wavelength shifts in absorption lines in the spectra of distant quasars. Such shifts may occur if the laws of physics are different at early times or in different locations in the universe. Increasingly accurate measurements would help us determine whether the physics of the early universe is the same as the physics we observe in our laboratory.

1.1 The Laws of Physics

Scientific inquiry relies on mathematical constructs to describe and predict the behaviour of various types of phenomena or entire systems. In this sense, it is useful to identify a set of the most basic mathematical constructs which are able to describe the broadest range of observations. When talking about physical systems such as our universe, we refer to these as *the laws of physics*.

1.1.1 The Standard Model

Our current understanding of the universe is best described by the Standard Model of particle physics (SM) and General Theory of Relativity (GR). Everything we can directly observe, either using our senses or advanced equipment, is composed of fundamental particles. The known particles are the six quarks, six leptons, twelve force carriers, and the Higgs boson. Particles have fundamental properties such as mass and various quantum numbers (e.g. spin, lepton and baryon numbers) and charges (e.g. electric charge, colour charge, weak charge, etc.). Charges determine the type and strength of interactions with other particles. There are four known types of interactions, also called forces: the weak and the strong nuclear forces, the electromagnetic force, and the gravitational force.

Quantum Field Theory provides the mathematical formalism for all but the gravitational force. The electromagnetic and the weak nuclear force used to be indistinguishable before the electroweak symmetry was broken (Salam & Ward 1964; Weinberg 1967). This symmetry can be restored, and the two forces reunited, at sufficiently high energies. It is generally believed that all forces were once united

and that a Grand Unified Theory (GUT), one able to provide the mathematical basis for all four forces, will emerge.

Whilst a quantum description of gravity is currently lacking, GR (Einstein 1917) provides the mathematical tools for understanding gravitational interactions between macroscopic objects and the universe as a whole. It is expected that a quantum theory of gravity will supersede GR, but no specific candidate theory stands out above the rest. Furthermore, numerous observational tests in the strong gravitational field regime, where quantum effects are expected to be important, do not deviate from GR predictions. Notable examples of this are the indirect (Hulse & Taylor 1975) and the direct (Abbott et al. 2016) detection of gravitational waves from inspiraling massive compact objects, constraints on monopole and dipole gravitational wave radiation[1] from pulsar timing (Antoniadis et al. 2013), observations of the gravitational redshift of light during a periastron passage of the star S2 around what is most likely a supermassive black hole in the centre of the Galaxy (Gravity Collaboration et al. 2018), the measurement of Schwarzschild precession of S2's orbit (Gravity Collaboration et al. 2020), and the observation of the black hole shadow by the Event Horizon Telescope Collaboration et al. (2019). All of these are in excellent agreement with GR predictions, demonstrating its power in studying the universe even in the absence of a quantum theory of gravity. Berti et al. (2015) give a comprehensive overview of possible tests in the strong field regime for several classes of quantum gravity theories.

A significant contribution of GR to our understanding of the universe as a whole pertains to the interpretation of the observed correlation between the recessional velocity (parametrized by redshift, z) and distances of galaxies outside of ours (Hubble 1929; Hubble & Humason 1931). Astronomers therefore commonly use redshift to indicate the passage of time because light travels longer from distant (high redshift) galaxies than from nearby (low redshift) ones. Interpreted in the context of GR, the distance-redshift relation is evidence of the expansion of the universe. This lead to a paradigm shift: the universe must have evolved from an initial dense state (possibly a singularity) after a sudden change of state which we call the Big Bang.

It is unknown what happened before or during the Big Bang [numerous theoretical models are proposed, e.g. Guth (1981); Linde (1982), see also Lyth & Riotto (1999); Armendáriz-Picón et al. (1999) for reviews]. It is hypothesised that the four known forces were unified into a single force at this time (Georgi & Glashow 1974; Pati & Salam 1974). At approximately the same time (or possibly simultaneously) the universe adiabatically expanded by ≈ 30 orders of magnitude in what we call inflation (Guth 1981), forming the fabric of spacetime and perhaps causing symmetry breaking due to which the unified force split into the four known forces (Linde 1982).

---

[1]In GR, the lowest order multipole term that can generate gravitational radiation is the quadrupole term.

1.1.2   Beyond the Standard Model

Ultimately, our understanding of the laws of physics is either incomplete or a low-energy limit of a more fundamental theory. This is true on the basis of theoretical considerations (GR is not a quantum theory of gravity and is thus inconsistent with the SM) and experimental considerations. Some attempts to unify gravity with the SM introduce additional spacetime dimensions, e.g. Kaluza-Klein theories (Klein 1926; Kaluza 2018, the latter being a republication of the original paper from 1921), superstring theories (Green 1985), and braneworld theories (Maartens 2004), all of which lack detailed experimental confirmation. A comprehensive recent review of different quantum gravity theories is given by Clifton et al. (2012).

Experimental evidence for the incompleteness in our knowledge of physics comes from particle physics experiments [e.g. neutrino flavour oscillations reported by Ahmad et al. (2002) and Fukuda et al. (1998)] and astronomical observations accrued throughout the 20th century. Astronomical observations are described below.

Dark matter

Observed accelerations of stars in galaxies (Rubin et al. 1980) and galaxies in galaxy clusters (Zwicky 1933) are inconsistent with GR (in its Newtonian limit) if stars, gas, or dust are the dominant source of gravitating mass in these systems. An extended gravitational potential comprising of material that does not create nor significantly obscure light (particles, massive compact objects, or both) or a modification to the law of gravity (e.g. Bekenstein 2004; Moffat 2006) is necessary to explain the observations. Simultaneous detection of gravitational and electromagnetic radiation emitted in a kilonova GW170817 (Abbott et al. 2017) falsified many modified gravity theories and left only a few viable modifications to GR (Ezquiaga & Zumalacárregui 2017). Massive compact objects in galaxy haloes which emit little or no electromagnetic radiation, e.g. brown dwarfs and black holes with masses $M \leq 100 M_\odot$, have been excluded by microlensing and gravitational wave observations (Zumalacárregui & Seljak 2018).

Current consensus in the physics community therefore gives more credibility to the existence of weakly interacting massive particles (WIMPs), even more so as several problems in astronomy [e.g. the positions of peaks in the Cosmic Microwave Background (CMB) power spectrum[2], measurements of the total baryonic matter fraction in the universe, structure formation in the early universe, galaxy dynamics] and in particle physics (e.g. the strong CP and the hierarchy problems), could be simultaneously solved by postulating additional gauge symmetries of the universe (also known as supersymmetry). Non-baryonic fundamental particles with limited (or no) interaction with the currently known ones come out naturally from these theories if supersymmetry is spontaneously broken at low energies. An overview of supersymmetry is given in Olive & Particle Data Group (2014, pp.1555-1568) whilst Bertone et al. (2005) give an overview of arguments for the existence of non-baryonic particles.

---

[2]Skordis & Złosnik (2020) recently produced the first satisfactory fit of the CMB in relativistic modified gravity framework, possibly paving the way for a viable extension of GR.

A new type of non-baryonic matter in the universe (the first being neutrinos) was thus introduced into physics. This new type of matter, called "dark matter" (DM), interacts with the baryonic matter through gravity and (possibly) through weak nuclear force. Astronomical observations and simulations of structure formation constrain the motion of DM particles to be non-relativistic (i.e. the particles should have a low dynamical temperature), adding a prefix "cold" to the name (CDM). Springel et al. (2006) describes the use of simulations in structure formation studies.

In order to explain the observations, CDM must compose 25% of the matter-energy content of the universe (Planck Collaboration et al. 2016). However, none of the many direct and indirect searches for a new particle have yielded detections (for a recent overview see, e.g., Marrodán Undagoitia & Rauch 2016). The search for DM particle(s) is still ongoing, but it is not clear whether detection will ever be possible either due to the large mass of DM particles, their limited interaction with baryonic matter or because they do not exist.

Dark energy

The second major evidence revealing our lack of knowledge about the laws of physics was first discovered from the observations of supernovae type Ia. This type of supernovae can be standardised to serve as distance indicators with high precision (see, e.g., Leibundgut 2017). The distances to these supernovae were measured to be larger than expected, demonstrating that the universal expansion rate is increasing for the last billion years (Perlmutter et al. 1999; Riess et al. 1998; Schmidt et al. 1998). Within the GR framework, this requires that the universe is permeated by a non-relativistic fluid with negative pressure, acting as a repulsive force on the largest scales and pushing galaxies away from each other at an ever increasing rate. Measurements using the baryonic acoustic oscillations method (e.g. Alam et al. 2017) and from the CMB (e.g. Planck Collaboration et al. 2016) lead to the same conclusion.

The nature of this fluid, termed "dark energy" (DE), is one of the greatest mysteries of contemporary physics. It seems to comprise 70% of the total matter-energy content of the universe (Planck Collaboration et al. 2016), yet almost nothing is known about it beyond its influence on the cosmic dynamics. There are numerous theoretical models for DE, summarised in a review by Copeland et al. (2006). The simplest model supported by the data is a cosmological constant (or Λ) associated to the vacuum energy in particle physics.

More complicated DE models, however, can emulate Λ. For example, an additional scalar field ϕ minimally coupled to matter can emulate Λ in the current epoch, but would have behaved differently in the early universe (Ratra & Peebles 1988; Armendariz-Picon et al. 2000). The DE phenomenon can also arise from a scalar field which couples to dark matter in more complicated models reviewed by Copeland et al. (2006).

It is clear that DM and DE point out the weaknesses in our understanding of physics. The best currently available model of the universe and its evolution, both in terms of its simplicity and agreement with observations, assumes GR is correct

and sufficient to explain gravitational interactions, and postulates the existence of an exotic form of non-baryonic matter (CDM) and the existence of an exotic form of energy arising from the quantum vacuum (Λ). We call this model ΛCDM. To fit with observations, CDM and Λ must comprise 95% of the matter-energy content of the universe with the remaining 5% being attributed to the detectable, baryonic matter.

Revealing the true nature of DM and DE phenomena will necessarily transform our knowledge of the universe and possibly pave the way to a Grand Unified Theory. This is why, over the last several decades, significant resources have been devoted to DM searches and to detailed studies of DE properties. Astronomical experiments that determine the structures occupied by baryonic matter in our universe and their evolution with time, aim to constrain the properties of the dark sector (see, e.g., Albrecht et al. 2006). Most of these use optical data collected by photometric surveys covering large areas of the sky up to high redshift, with ancillary spectroscopic observations used to constrain systematic errors. However, a different type of dedicated spectroscopic studies may probe a different region of the parameter space, as I go on to explain in Section 1.1.3. Combined with photometric surveys, these spectroscopic studies provide much more stringent constraints on DE properties and allow us to more easily discard unrealistic models. Chapter 4 presents results of a study to explore the nature of DE, constraining models in which the cause for the DE phenomenon also causes a change in the strength of electromagnetic force.

1.1.3 Fundamental constants and when they are expected to vary

Any particular definition of the laws of physics require a set of free parameters, or fundamental constants, to provide a satisfactory description of reality. Their values are not given *a priori* and cannot be calculated from first principles of the theory. Instead, values of fundamental constants have to be measured experimentally. The choice of what is a fundamental constant is somewhat arbitrary and depends on the mathematical framework used. Whilst a discussion presented in Duff et al. (2002) argues that only three (or possibly even two) constants are necessary within a string theory framework, we currently use 22 constants (Uzan 2011). Additional seven constants would be needed to incorporate massive neutrinos. The 22 fundamental constants used in the SM and GR, together with their best measured values, are given in Table 1. See Uzan (2011) and Martins (2017) for recent reviews of the role that fundamental constants have in our physical theory.

Which set of parameters are considered fundamental is defined by the theoretical framework used and, as such, changes with as our understanding of the universe improves. Lévy-Leblond (1979) classifies constants into three categories depending on their universality: those characteristic of a particular system, those characteristic of a class of physical phenomena, and truly universal constants. An example in each category, from less to more universal, would be the electrical resistance of a material, the electric permittivity of the vacuum, and the speed of light. A constant can move between categories (as was the case with, e.g., the speed of

Table 1.1: A list of fundamental constants used in our formulation of laws of physics. The list was produced from Uzan (2011). Values were updated by values from Tanabashi et al. (2018). The figures in parentheses after the values give the 1-standard-deviation uncertainties in the last digits.

Constant	Symbol	Value
Speed of light in vacuum	c	$299\,792\,458\,\mathrm{m\,s^{-1}}$
Planck constant (reduced)	$\hbar$	$1.054\,571\,726(47) \times 10^{-34}\,\mathrm{J\,s}$
Newton constant	G	$6.673\,84(80) \times 10^{-11}\,\mathrm{m^2\,kg^{-1}\,s^{-2}}$
Weak coupling constant (at m_Z)	$g_2(m_Z)$	0.6520 ± 0.0001
Strong coupling constant (at m_Z)	$g_3(m_Z)$	1.221 ± 0.022
Weinberg angle		$0.2229(3)$
Electron Yukawa coupling	h_e	2.935×10^{-6}
Muon Yukawa coupling	h_μ	0.000607
Tauon Yukawa coupling	h_τ	0.0102057
Up Yukawa coupling	h_u	0.0000126 ± 0.000003
Down Yukawa coupling	h_d	0.000027 ± 0.000002
Charmed Yukawa coupling	h_c	0.00732 ± 0.00002
Strange Yukawa coupling	h_s	0.00055 ± 0.00003
Top Yukawa coupling	h_t	0.994 ± 0.002
Bottom Yukawa coupling	h_b	0.0240 ± 0.0002
Quark CKM matrix angle	$\sin\theta_{12}$	0.2244 ± 0.0014
	$\sin\theta_{23}$	0.04221 ± 0.0014
	$\sin\theta_{13}$	0.0249 ± 0.0005
Quark CKM matrix phase	δ_{CKM}	1.05 ± 0.24
Higgs potential quadratic coefficient	$\hat{\mu}^2$	$7835.02 \pm 20.03\,\mathrm{GeV^2}$
Higgs potential quartic coefficient	λ	0.1295 ± 0.021
QCD vacuum phase	θ_{QCD}	$< 10^{-9}$

light) or removed altogether if a more fundamental constant is revealed. We thus implicitly assume that true fundamental constants are indeed constant and independent from each other. A corollary of this is that we can test our fundamental physical theories by looking for variations in their fundamental constants.

The fine structure constant

This thesis focuses on the fine-structure constant, α, a dimensionless quantity which characterizes the strength of the electromagnetic force. It is defined (using SI units) as

$$\alpha \equiv \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c},$$
(1.1)

where ϵ_0 is the permittivity of the vacuum, e is the electron charge, $\hbar$ is the reduced Planck constant, and c the speed of light.

Throughout the thesis, variation in the value of α is defined as

$$\frac{\Delta\alpha}{\alpha} = \frac{\alpha_{obs} - \alpha_{lab}}{\alpha_{lab}},$$
(1.2)

with α_{obs} and α_{lab} referring to the values of α in the observed system and in a laboratory on Earth.

Variation of α is, in principle, almost expected. Its variation is predicted by many GUTs and generalisations of the SM and GR aiming to provide a physical explanation for DM/DE. In these theories, α becomes a dynamical quantity and its value is determined by its coupling to some new degrees of freedom. In higher-dimensional theories, e.g. Kaluza-Klein and string theories, α can vary due to contraction or expansion of the compactified higher dimensions. α also varies in theories in which DE is attributed to an additional scalar field, ϕ, if ϕ couples to the baryonic matter (Bekenstein 1982; Sandvik et al. 2002; Copeland et al. 2004; Shaw & Barrow 2005; Barrow & Lip 2012) or if the vacuum expectation value of ϕ depends on matter density (Silva et al. 2014). Furthermore, α may vary with the gravitational potential (Dicke 1959; Mota & Barrow 2004a,b), and in supersymmetric theories in which dark matter particles interact with baryons (Olive & Pospelov 2002; Stadnik & Flambaum 2015).

The exact mechanism behind the variation depends on the specifics of the model, but they may be roughly grouped into models in which α varies in time, in space, or in a strong gravitational field. Martins (2015) classifies "varying constant theories" into those which also provide all DE required by observations (Class I, e.g. Carroll 1998; Dvali & Zaldarriaga 2002) and those which contribute only partially or do not contribute at all to DE (Class II, e.g. Bekenstein 1982; Sandvik et al. 2002). Extensive reviews of theoretical models are given by Uzan (2011) and Martins (2017).

Searching for variations of α is therefore a powerful method to explore the nature of DE, complementary and independent from the methods mentioned in Section Albrecht et al. (2006).

1.1.4 Literature constraints on the fine structure constant variation

Variations in α have been explored through terrestrial laboratory measurements, the geological record, and from astronomical observations of white dwarfs, molecular clouds in the Galaxy, the CMB, galaxies, quasars, and others. Constraints from quasar spectra specifically are discussed in Section 1.2.4.

Laboratory measurements

Laboratory measurements mainly focus on measuring the relative frequency drifts of transitions of two atomic clocks, where each clock measures frequencies of a different atomic species. If α varies in time, the frequency ratio of the two atomic clocks should change as well. The best current atomic clock measurements, using Hg^+ and Al^+ optical clocks (Rosenband et al. 2008), constrain

$$\dot{\alpha}/\alpha = (-1.6 \pm 2.4) \times 10^{-17} \, \mathrm{yr}^{-1},$$

where $\dot{\alpha}$ denotes the time derivative of α. A completely independent method is based around dysprosium isotopes. The two isotopes (^{163}Dy and ^{162}Dy) have transitions which are almost degenerate in energy, but are of opposite parity. As a consequence of the opposite parity, the two transitions have opposite sensitivity coefficients for variation in α. Using this method, Cingöz et al. (2007) obtain

$$\dot{\alpha}/\alpha = (2.7 \pm 2.6) \times 10^{-15} \, \mathrm{yr}^{-1}.$$

The field of laboratory searches for varying constants is broad and therefore beyond the scope of this thesis. A review of various laboratory methods and recent results is given by Safronova et al. (2018).

The Oklo phenomenon

Constraints of similar magnitude were obtained by radioactive isotope analysis of the Oklo phenomenon, a naturally occurring nuclear fission reactor active in Gabon approximately 1.8 billion years ago (Naudet & Ronson 1974; Maurette 1976; Petrov 1977). Neutrons produced in the fission of ^{235}U were captured by ^{149}Sm to produce ^{150}Sm through resonance capture at ≈ 0.1 eV. Had the value of α been different at the time the reactor was active compared to its value today, the resonance would been at a slightly different energy, and the reaction rates would be slightly different. Current relative abundance of ^{150}Sm/^{149}Sm in the extinguished reactor therefore probes the neutron capture process at $z = 0.2$. This type of measurement constraints variations in the strong nuclear force (e.g. Shlyakhter 1976), and variations in α. Damour & Dyson (1996) constrain α variation from Oklo data to

$$-6.7 \times 10^{-17} \, \mathrm{yr}^{-1} < \dot{\alpha}/\alpha < 5.0 \times 10^{-17} \, \mathrm{yr}^{-1},$$

where the numbers above represent 95% confidence intervals. The Oklo data were subsequently analysed by Fujii et al. (2000) and Davis & Hamdan (2015).

Specifically, Davis & Hamdan (2015) improve the calculations of Damour & Dyson (1996) and find, with 95% confidence,

$$|\dot{\alpha}/\alpha| < 0.61 \times 10^{-17} \, \mathrm{yr}^{-1}.$$

An important caveat of these measurements is that they strongly depend on nuclear physics models and on assumptions of the physical conditions in the reactor.

Meteorite dating

Products of long-lived α- and β-decays probe the time evolution of $\Delta\alpha/\alpha$, as small variations in α will manifest as different ratios of the decay products (Wilkinson 1958; Dyson 1972). Meteorites provide constraints of $\Delta\alpha/\alpha$ at the time of formation of the Solar System ($z \approx 0.43$). The tightest constraints from measurements of this type, using ^{187}Re (Olive et al. 2004), yield

$$|\Delta\alpha/\alpha| \lesssim 10^{-7}.$$

White dwarfs

In models in which an additional scalar field couples to matter, variation of α can occur in strong gravitational fields (see, e.g. Dicke 1959; Magueijo et al. 2002). This was explored through high-resolution observations of metal-polluted white dwarfs. Berengut et al. (2013) and, more recently, Hu et al. (2020) find that

$$|\Delta\alpha/\alpha| \lesssim 10^{-5}$$

in presence of gravitational field $\sim 10^5$ times stronger than on Earth. The largest uncertainty on $\Delta\alpha/\alpha$ from white dwarf measurements comes from the uncertainties on laboratory wavelengths.

A measurement of $\Delta\alpha/\alpha$ near the Galactic centre

If α changes in the presence of a strong gravitational field, this could be made apparent in the region around the supermassive black hole in the Galactic centre, Sagittarius A*. A recent spectroscopic study (Hees et al. 2020) of five evolved stars in direct proximity of Sagittarius A* constrained possible variations of α in the presence of gravitational field ≈ 500 times stronger than on Earth to

$$\Delta\alpha/\alpha = 1.0 \pm 5.8 \times 10^{-6}.$$

Cosmic Microwave Background

A different value of α at the time of matter-radiation decoupling ($z \approx 1100$) would affect the Thompson cross-section, the binding energy of hydrogen, and the recombination rates. These would change the time of the decoupling and alter the baryon-to-photon ratio at that time, leaving an imprint on the CMB power spectrum at sub-1° angular scales (Hannestad 1999; Kaplinghat et al. 1999). These methods were applied to BOOMERanG and MAXIMA observations, which prefer

a smaller value of α (compared to today) at the time of last scattering, with at least 63% certainty (see table 1 in Battye et al. 2001). The latest and most precise constraints were obtained from the *Planck* satellite (Planck Collaboration et al. 2015), finding

$$\Delta\alpha/\alpha = (3.6 \pm 3.7) \times 10^{-3}$$

with 68% confidence. Spatial variations using the CMB are constrained down to $(\Delta\alpha/\alpha)_{rms} < 3.4 \times 10^{-3}$ (O'Bryan et al. 2015).

Big Bang nucleosynthesis

Physical conditions just after the Big Bang led to the production of light elements: hydrogen, deuterium, helium, lithium (Gamow 1948a,b) in what is called the Big Bang nucleosynthesis (BBN). Their relative abundances depend, besides the temperature and particle density, on the values of coupling constants for various forces: α for electromagnetism, α_s for the strong nuclear force, α_w for the weak nuclear force, and Newton's constant for gravity. Kolb et al. (1986) first used BBN abundances to measure $\Delta\alpha/\alpha$ and constrain higher-dimensional theories, considering also degeneracies when other couplings change simultaneously. This work was later expanded by Barrow (1987) and Campbell & Olive (1995). Landau et al. (2008) combined BBN constraints with CMB data to constrain variations in α and in the Higgs vacuum expectation value at the time of BBN, but they note that ^{7}Li should be measured with better accuracy to yield better constraints. Analysing the measured abundances of nuclei up to lithium in primordial gas, Clara & Martins (2020) find a preference in the data for more positive α at BBN, of the order of several parts per million. This result, however, depends on the priors used in the analysis.

Comment on the variation of the strong nuclear force

Mechanisms responsible for variation in the electromagnetic force can also cause variations in the strong nuclear force. This would change the binding energy in nucleons, effectively changing their masses. Because electron mass would not be affected in this scenario, variation of the strong nuclear force can be probed through measurements of the proton-to-electron mass ratio, $\mu \equiv \frac{m_p}{m_e}$.

Thompson (1975) was the first to point out that molecular transitions can be used to measure variations in constants such as α, μ, and various gyromagnetic factors. The first reliable astronomical measurement of μ was derived from optical transitions of molecular hydrogen at $z = 2.811$ (Varshalovich & Levshakov 1993), yielding:

$$\Delta\mu/\mu = 3.4 \times 10^{-3}$$

Recently, using optical CO transitions, Daprà et al. (2016) obtained:

$$\Delta\mu/\mu = (2.3 \pm 1.7) \times 10^{-5}$$

Frequencies in the radio bands, such as transitions of polyatomic molecules, can generally be measured very accurately. Truppe et al. (2013) improved laboratory

frequencies of CH molecular transitions, reaching $3\,\mathrm{Hz}$ accuracy, and obtained

$$\Delta\alpha/\alpha = (0.3 \pm 1.1) \times 10^{-7}$$

and

$$\Delta\mu/\mu = (-0.7 \pm 2.2) \times 10^{-7}$$

in a molecular cloud within the Galaxy. A comprehensive review of recent μ measurements is given by Ubachs (2018), and a review of measurements of various combinations of fundamental constants by Martins et al. (2018).

21-cm line measurements

Recently, Lopez-Honorez et al. (2020) calculated the observable effects of varying α on the 21-cm hydrogen line during the reionization period ($6 < z < 30$). They predict that the future Square Kilometre Array will be able to measure α with precision of 0.1%. Whilst not as precise as lower-redshift quasar measurements, these results would be a unique and independent check on varying constants in this redshift range.

1.2 Fundamental constant measurements from high resolution quasar spectroscopy

Quasars are a class of extremely luminous extragalactic objects visible at cosmological distances: the lowest redshift quasars appear at redshift $z \gtrsim 0.1$. The first quasars were discovered as strong radio sources associated with star-like optical counterpart (Schmidt 1963), and hence named Quasi-Stellar Radio sources. For a historical review of events leading to the discovery of the first quasar, see Hazard et al. (2018). The ubiquity of quasars on the sky and at almost all redshifts make them an excellent probe of varying α models, particularly those in which α varies with time or with direction on the sky.

Quasars produce intense electromagnetic radiation by converting the gravitational potential energy of material in direct proximity of a central supermassive black hole into radiation (Salpeter 1964; Zel'dovich 1964; Lynden-Bell 1969). The radiation covers the entire electromagnetic spectral range and shows features of different origin.

1.2.1 Quasars and high-resolution spectroscopy

We distinguish two types of features in quasar spectra, illustrated in Figure 1.1. The first type is directly connected to the surroundings of the central black hole, e.g. synchrotron emission from relativistic electrons, X-rays from the hot viscous material in direct proximity of the black hole, broad and narrow ultra-violet and optical emission lines originating from the material located at short distances from the black hole (< 1 parsec for the narrow and < 100 parsec for the broad lines), etc. Whilst these features are important for understanding high-energy physics of accretion disks and the role of the central black hole for host galaxy evolution
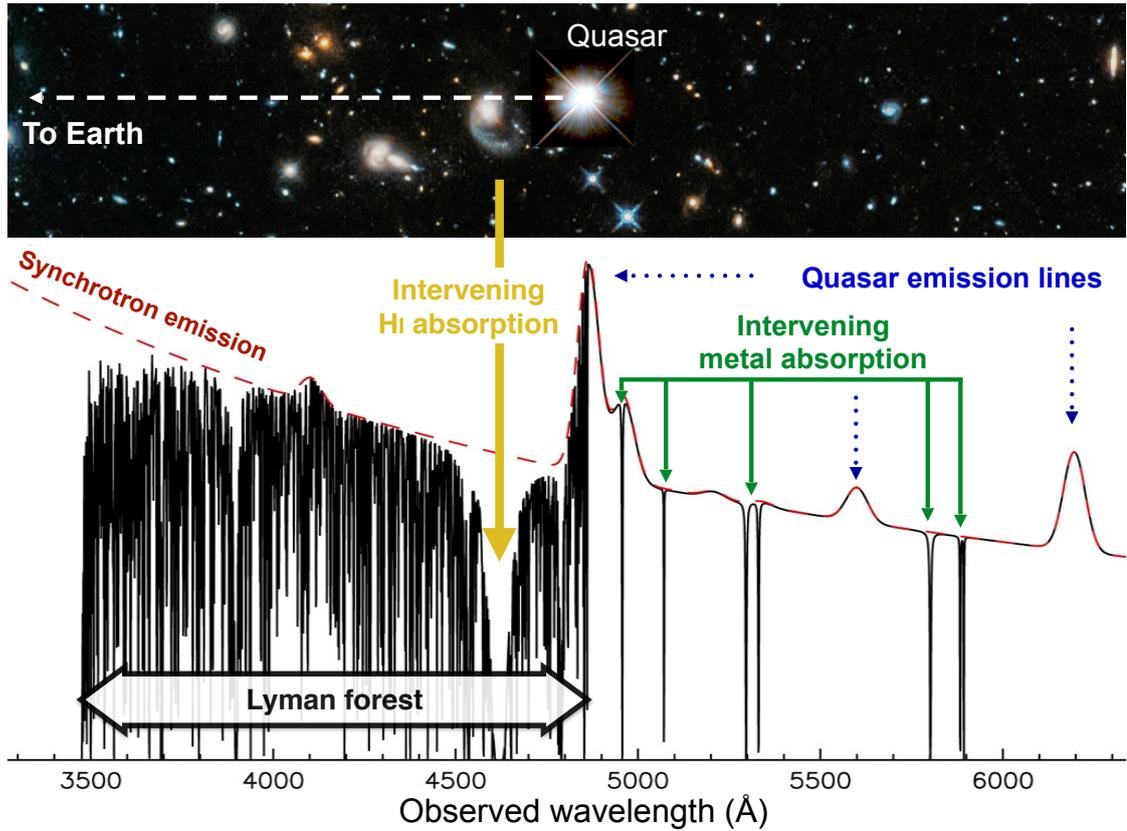
Figure 1.1: A simulated optical quasar spectrum, showing the most significant features. The synchrotron continuum emission (red dashed line) as well as broad and narrow emission lines (indicated by blue dotted arrows) result from high-energy processes occurring close to the supermassive black hole powering the quasar. These processes are extremely energetic, making quasars visible at distances comparable to the size of the observable universe. The light travelling towards Earth is absorbed by the intergalactic gas clouds along the line of sight. The clouds leave an imprint in the spectrum, mostly in the form of a forest of Lyman lines (see, e.g. Rauch 1998). More massive clouds, sometimes associated with massive galaxies, leave strong Lyman-α (thick yellow arrow) and metal (thin green arrows) absorption lines. These absorption lines can be used to search for variations in fundamental constants. Modified from an image provided by John K. Webb.

(Kormendy & Ho 2013, and references therein), they are unimportant for the purpose of this thesis and therefore a more detailed description is not provided.

The second type of features result from the interaction of the light emitted by the quasar with the material located between the quasar and Earth. As the quasar light travels towards us, it is absorbed and re-emitted by intervening gas clouds, producing a series of absorption lines in the spectrum. The most common is the Lyman-α absorption line associated with the neutral hydrogen gas in the intergalactic medium (Lynds 1971; Sargent et al. 1980) or, in the case of strong absorption, outer parts of massive galaxies. The numerous Lyman-α lines (also called the Lyman-α forest) trace the large-scale structure of the universe and its

evolution with time. The Lyman-α lines have been used for structure formation studies, both in observations and in simulations (e.g. Cen et al. 1994; Miralda-Escudé et al. 1996), but also to constrain cosmological parameters via baryonic acoustic oscillations and redshift space distortions (e.g. eBOSS Collaboration et al. 2020). Combined with the CMB and supernovae samples, the forest also constrains DM properties, such as the mass of (hypothesized) sterile neutrinos, and therefore cosmological scenarios of structure formation (Viel et al. 2006; Seljak et al. 2006; BOSS Collaboration et al. 2015). An extensive description of absorption lines in quasar spectra is given in Rauch (1998) and Bechtold (2001).

1.2.2 Requirements for measuring the fine structure constant in quasar spectra

The quasar spectrum may exhibit groups of absorption lines associated with clouds arising in a single system. The complexity of the absorption system seen in a spectrum depends on the number of clouds, their physical size, particle density, the clustering scale, and the way quasar light intersects each cloud (i.e. the impact parameter). The relative abundances of heavy elements in each cloud depend on the star-formation history and interactions it previously had with its environment. Therefore, in addition to hydrogen absorption, these gas clouds often also show absorption from transitions of spatially coincident metal atoms (Burbidge et al. 1966; Stockton & Lynds 1966). In this context, metals are all elements heavier than hydrogen. The most commonly observed metals are iron (Fe), magnesium (Mg), manganese (Mn), aluminium (Al), chromium (Cr), zinc (Zn), and others. These transitions provide a direct probe of star and structure formation and of chemical evolution of the universe. Most importantly for the work in this thesis, they provide a way to measure the value of α at the time light was absorbed, several billion years ago.

The following is required to measure α from quasar absorption systems:

1. The absorption system must contain transitions with a range of sensitivities to α variation (see Section 1.2.3). Several hundred suitable systems are known today with large compilations of α measurements in, e.g., Murphy et al. (2004); King et al. (2012); Pinho & Martins (2016) and Wilczynska et al. (2020).

2. Observations must have sufficiently high spectral resolution ($R = \lambda/\Delta\lambda$) and signal-to-noise ratio (S/N) to resolve individual absorbing components. This allows accurate modelling of the absorption system and therefore accurate measurements of the physical properties of the gas (see Section 1.3.1). Observations of bright targets (or long integration times for fainter ones) using high-resolution astronomical spectrographs are needed to fulfil this condition. The spectrum should generally have $R > 40000$ and S/N > 10 per pixel in the continuum. The most commonly used are UVES, HIRES, HARPS, and HDS spectrographs. New high-quality observations are expected from the newly constructed ESPRESSO spectrograph.

3. Spectrograph wavelength calibration should be significantly more accurate than the uncertainty on absorption line centres, and comparable to (or better than) the uncertainty on laboratory wavelengths (see Section 1.3.2).

   This is currently the most demanding constraint. Whilst the uncertainty on absorption line centres depends on data quality, some laboratory wavelengths are known with $\Delta\lambda/\lambda = 10^{-8}$ accuracy, an order of magnitude better than what is currently achievable with astronomical spectrographs. Chapters 2 and 3 of this thesis deal with novel wavelength calibration techniques to bridge this gap.

1.2.3   Measuring the fine structure constant

Variation in the value of α manifests as a perturbation of energy levels of atoms and results in small frequency shifts of atomic transitions with respect to their measured laboratory values. These can equivalently be expressed as shifts in the observed wavelength $\lambda = c/f$ or wavenumber $\omega = 1/\lambda$, where c and f are the speed of light and the frequency.

First astronomical constraints on α variation were derived by measuring the separation between fine-structure doublet transitions, which scales as α^2. This method, called the *Alkali Doublet* method, was first applied to galaxy emission lines with an uncertainty of $\sim 10^{-2}$ (Savedoff 1956). Bahcall et al. (1967) used narrow quasar absorption lines to obtain $\Delta\alpha/\alpha = (-2 \pm 5) \times 10^{-2}$ at $z = 1.95$. Later, Varshalovich et al. (1996) used Si II doublets and obtained $|\Delta\alpha/\alpha| < 1.6 \times 10^{-4}$ (95% confidence limit) for $z \approx 2.8-3.1$. The tightest constraints from applying the Alkali Doublet method to quasar spectra were obtained by Murphy et al. (2001b),

$$\Delta\alpha/\alpha = (-0.5 \pm 1.3) \times 10^{-5}.$$

The Many Multiplet method

An order of magnitude improvement in precision was achieved by the *Many Multiplet* (MM) method (Dzuba et al. 1999b; Webb et al. 1999). The MM method parametrises the shift of the wavenumber of a particular transition in terms of α at redshift z and on Earth (α_z and α_0), and sensitivity coefficients q and q_2:

$$\omega_z = \omega_0 + q\left[\left(\frac{\alpha_z}{\alpha_0}\right)^2 - 1\right] + q_2\left[\left(\frac{\alpha_z}{\alpha_0}\right)^4 - 1\right], \tag{1.3}$$

Here ω_z and ω_0 are the wavenumbers in the restframe at redshift z and in a laboratory on Earth, respectively. The values of coefficients q and q_2 are obtained by solving relativistic many-body equations in Hartree-Fock potential (Dzuba et al. 1999a). The typical value of q_2 is an order of magnitude smaller than typical value of q and can therefore be ignored. Converting into the observed frame using $\omega_z = \omega'(1+z)$ and rearranging Eq. (1.3) to obtain the expression for the observed frame laboratory wavelength when $\Delta\alpha \neq 0$, one gets:

$$\lambda' \approx \lambda_0(1+z)\left[1 - \frac{2q}{\omega_0}\frac{\Delta\alpha}{\alpha}\right]. \tag{1.4}$$
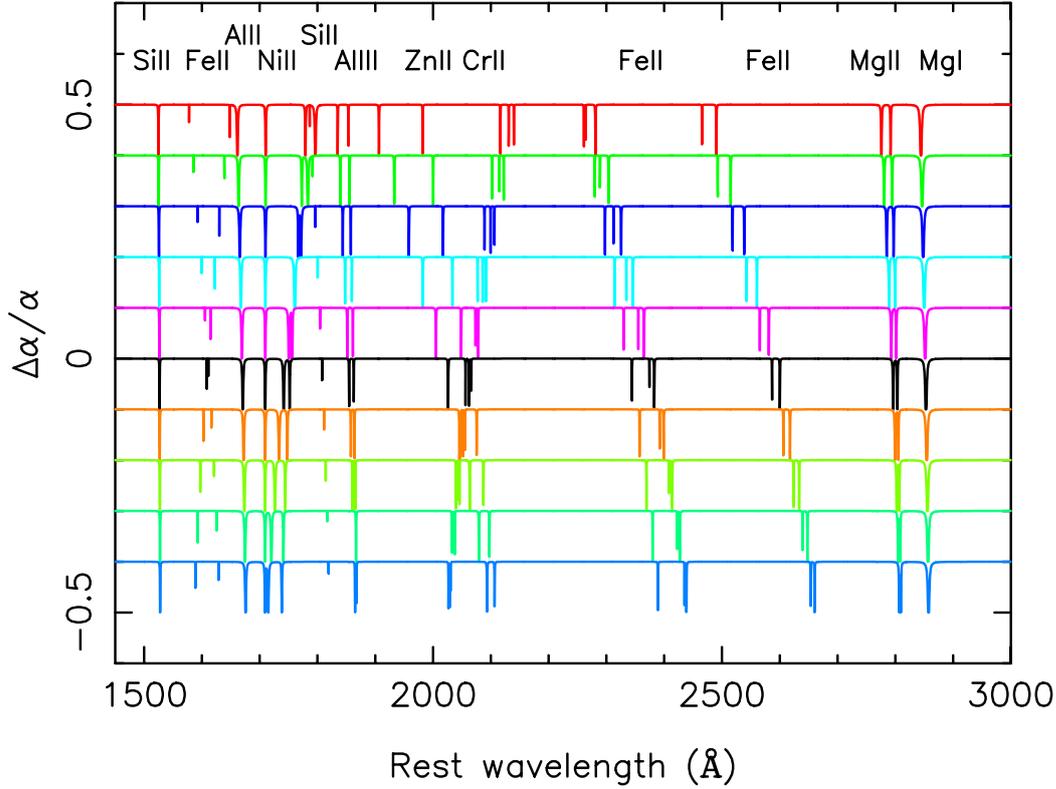
Figure 1.2: Illustration of the effect that varying α has on the wavelengths of transitions commonly observed in quasar spectra. The black line shows a spectrum for which α is at its terrestrial value. The coloured spectra above and below illustrate how particular transitions shift when $\Delta\alpha/\alpha$ is changed in steps of 0.1 and -0.1, respectively. For positive $\Delta\alpha/\alpha$, wavelengths of transitions with negative (positive) q-coefficients shift towards the red (blue). The opposite is true for negative $\Delta\alpha/\alpha$. Transitions with negative q in this Figure include Fe II $\lambda 1608$ and Cr II $\lambda 2066$. Positive q transitions include Zn II $\lambda\lambda 2026, 2062$ and Fe II $\lambda\lambda 2344, 2374$. Transitions with small q-coefficients of either sign experience only small wavelength shifts and serve as anchors, e.g. Si II $\lambda 1526$ and Mg II $\lambda\lambda 2796, 2803$. Variation in α therefore produces a unique and retrievable pattern of velocity shifts in the astronomical spectra. Simultaneous measurement of shifts for transitions with different q-coefficients provides a stringent constraint on $\Delta\alpha/\alpha$ in the Many Multiplet method, used in this thesis. Image credit: John K. Webb.

Here, ω' and λ' are observed frame quantities, λ_0 is the laboratory wavelength, and only the first two terms of the Taylor series expansion were used to approximate the second term in Eq (1.3).

It is clear from Eq. (1.4) that, should $\Delta\alpha \neq 0$, the second multiplicative term acts as an additional wavelength shift independent from redshift. However, one must use transitions with different q-coefficients (and, in fact, as different as possible) in order to distinguish shifts caused by $\Delta\alpha \neq 0$ from z. For convenience and convention, I will express wavelength shifts $\Delta\lambda$ in velocity units throughout the

text,

$$\frac{\Delta v}{c} = \frac{\Delta \lambda}{\lambda}, \tag{1.5}$$

where c is the speed of light.

Transitions are categorised into the following three categories according to their q-coefficients: (i) transitions with large and negative q coefficients; (ii) transitions with large and positive q coefficients; and (iii) transitions with small q coefficients of either sign. For a given $\Delta \alpha / \alpha$, the velocity shift for each transition is different. This is illustrated using commonly observed transitions in Figure 1.2. Together, they leave a unique pattern of velocity shifts which can be retrieved through modelling (see Section 1.3.1). Having transitions of all three categories present in the absorption system is very desirable, as their different sensitivities to α increase the measurement precision. The most up-to-date collection of q coefficients for transitions commonly observed in quasar spectra are provided by Dzuba et al. (1999a, 2002); Dzuba & Flambaum (2009); Murphy & Berengut (2014).

1.2.4   A possible spatial dipole in α

The most exciting and controversial results are certainly those published in a series of papers by Webb et al. (1999), Webb et al. (2001), Murphy et al. (2003a), Webb et al. (2011), and King et al. (2012). These authors measured $\Delta \alpha / \alpha$, using the MM method, in several hundred quasar absorption systems observed by the Ultra-Violet Echelle Spectrograph (UVES, Dekker et al. 2000) installed on the Unit Telescope 2 of the Very Large Telescope operated by the European Southern Observatory and the High Resolution Echelle Spectrometer (HIRES, Vogt et al. 1994) installed on the Keck 10m telescope of the Keck Observatory.

Analysing 30 systems observed by the HIRES instrument, Webb et al. (1999) found that α could have been smaller in the past. Murphy et al. (2003a) increased the sample to 128 systems and came to the same conclusion. Later, Webb et al. (2011) and King et al. (2012) analysed 153 systems observed by the UVES instrument and found a contrary result, i.e. evidence for α being larger in the past. Individual measurements for both samples are plotted in Figure 1.3. Systematic errors due to, e.g., laboratory wavelength errors, heliocentric velocity variation, hyperfine structure effects, magnetic fields, kinematic effects, airvacuum wavelength conversion, were found to be negligible (Murphy et al. 2001a, 2003b).

King et al. (2012) expanded the HIRES sample (to 140 systems) and combined it with the UVES sample (for a total of 293 systems) interpreting the results as evidence for a spatial dipole of α across the universe. The 1σ confidence level for the "north pole" of the dipole, derived independently on the HIRES and UVES are shown as green and blue shaded regions in Figure 1.3. The combined dipole, shown as a red shaded region on the same Figure, has a significance of $> 4.1\sigma$ and has generated significant discussion in the astronomical community and in the general public.

Several other papers have appeared from groups that work independent of the group headed by John K. Webb at UNSW Sydney. For example, a series of papers reporting a small number of measurements using ESO VLT data do not find spatial
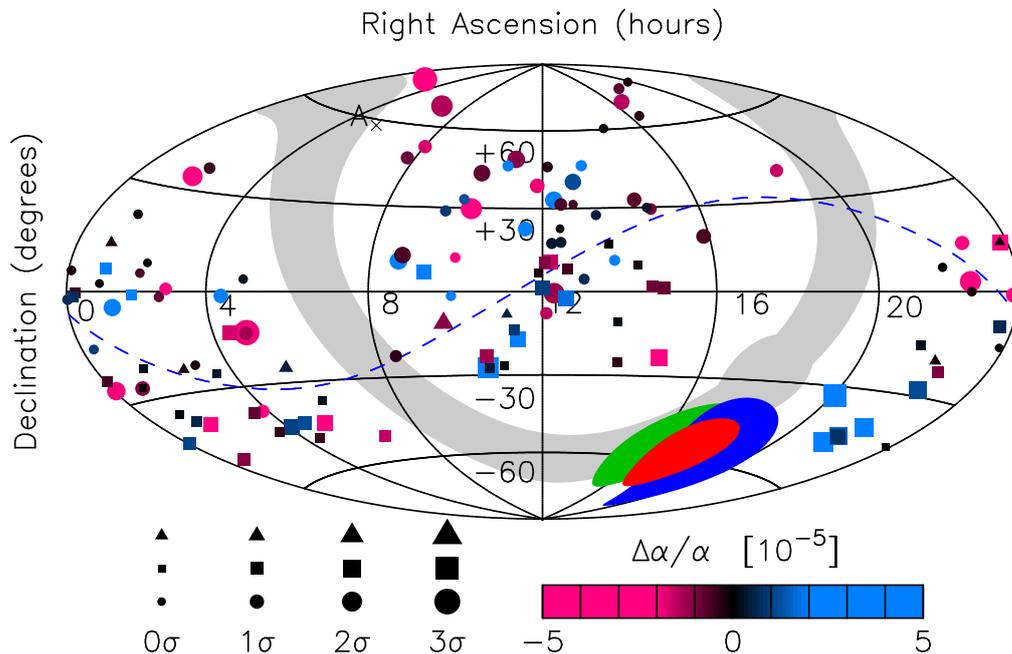
Figure 1.3: Sky map of $\Delta\alpha/\alpha$ measurements collected and analysed in King et al. (2012). The symbols represent the normalised residuals, $r = (\Delta\alpha/\alpha)/\sigma_{\Delta\alpha/\alpha}$, where larger symbols represent larger r values. Measurements made from UVES observations are marked by squares, those from HIRES observations by circles, and those from both instruments by triangles. The pink-blue colour code indicates the difference between the measured $\Delta\alpha/\alpha$ and zero. The red shaded region represents the 1σ confidence limits on the dipole of the form $\Delta\alpha/\alpha = A\cos\theta$ fitted to all measurements. The blue and the green regions show the same for dipole models fitted to the UVES and HIRES observations. The antipole is marked with an "A". The gray shaded region indicates the Galactic plane and the bulge indicates the Galactic Centre. Figure taken from King et al. (2012).

variation, although the sample size and overall precision are really insufficient to provide a stringent test (Molaro et al. 2013a; Rahmani et al. 2013; Evans et al. 2014). The most complete compilation of all results to date are given in Wilczynska et al. (2020).

1.3   Methodology

1.3.1   Absorption system modelling

Information on the physical properties of the intervening cloud are contained in the absorption line it imprints onto the quasar spectrum. This information can be retrieved through absorption line modelling and fitting. The goal of modelling is therefore to obtain a statistically acceptable, physical description of the absorption system from which interesting quantities can be measured. The most commonly used tool for this, when also trying to measure $\Delta\alpha/\alpha$, is VPFIT. VPFIT and AI-VPFIT are both extensively used in Chapter 4. This Section provides a brief description of their methodology with the intention of clarifying how models in Chapter 4 were obtained.

General considerations

Absorption profiles seen in quasar spectra generally consist of several blended absorption lines (also termed "absorption components" or "velocity components"). Individual lines can be modelled using a Voigt profile, which represents an atomic line for which line broadening contains independent contributions of a Gaussian and Lorentzian shapes (van de Hulst & Reesinck 1947). The Gaussian shape arises from the turbulent and thermal motions of the gas, and Lorentzian from the finite lifetime of excited states of absorbing atoms. Voigt profiles are generally evaluated numerically, using look-up tables [e.g. Harris (1948); Armstrong (1967), but see also Carswell & Webb (2014)].

The Voigt profile of a transition is described by a number of fixed and free parameters. The fixed parameters are determined empirically in a laboratory, such as the rest-frame wavelength λ, the oscillator strength f, and the inverse of the transition lifetime Γ. Free parameters are determined by fitting models to the data, such as the integrated surface density (column density) of gas in the line-of-sight (N, in units cm^{-2}), the redshift (z), and the total velocity width (b-parameter, in units km s^{-1}). We include gas temperature (T) and $\Delta\alpha/\alpha$ as additional free parameters. Parameters such as redshift, temperature, and $\Delta\alpha/\alpha$ are the same for all transitions associated with the same absorbing could.

VPFIT

VPFIT is a tool designed to fit multiple Voigt profiles to spectral data (Carswell & Webb 2014). The data and the model are provided by the user at input. The model is fully specified by the parameters of the constituent Voigt profiles and, if required by the data, additional parameters such as the continuum and zero-level corrections.

Initial values (i.e. first guesses) for the free parameters of the model, denoted together by the vector $\mathbf{x}$, are chosen by the user after a careful examination by eye and method of trial-and-error. Starting from the initial values, VPFIT uses

non-linear least-squares to find $\mathbf{x}$ which minimizes

$$\chi^2 = \sum_{i=1}^{N} \frac{[F_i^{model}(\mathbf{x}) - F_i^{data}]}{\sigma_i^2}. \qquad (1.6)$$

Here, i denotes the values associated to the ith pixel, $F^{model}(\mathbf{x})$ is the model prediction given $\mathbf{x}$, and F^{data} and σ are the observed flux and the corresponding 1σ uncertainty. The sum runs over N pixels.

VPFIT iteratively descents through the multi-dimensional χ^2-space until it reaches a minimum. This generally occurs when the decrease in χ^2 between consecutive iterations is much smaller than unity (Press et al. 1992). To ensure that $\mathbf{x}$ is as close as possible to its lowest χ^2 values, we demand that the fractional decrease in χ^2 is $\Delta\chi^2/\chi^2 < 1 \times 10^{-6}$ for more than 3 consecutive iterations. VPFIT has been extensively tested over several decades, and has been shown to produce reliable results on simulated data (Murphy et al. 2003a; King et al. 2010).

A statistically acceptable fit to the data is achieved when the normalized χ^2 is around unity, where the normalized χ^2 is defined as $\chi^2_\nu \equiv \chi^2/\nu$ and ν is the number of degrees of freedom. For such models, the dispersion of the data around the predictions of the model is consistent with the uncertainty on the data. Conversely, models for which $\chi^2_\nu \gg 1$ probably require additional free parameters (i.e. are "under-fitted"), and those for which $\chi^2_\nu \ll 1$ contain too many free parameters (i.e. are "over-fitted"). Decisions on which parameters should be added and/or removed from the model are made by a human.

It is possible to find more than one model with $\chi^2_\nu \lesssim 1$ for the same data. This can happen, for example, if two different people model the same data. Due to differences in the number of free parameters or their initial values, each model may yield a different value of $\Delta\alpha/\alpha$ (Wilczynska et al. 2015). At this point, one must decide which model is more physical or more likely. There are several statistical methods to compare models, including the F-test, the Bayesian Information Criterion, Akaike Information Criterion, and others (e.g. Bainbridge & Webb 2017; Bainbridge et al. 2017). We use the Akaike Information Criterion (AIC, Akaike 1974) to decide which model should be used.

The AIC $= \chi^2 + 2p$, where p is the number of free parameters in the model. The second term penalises the introduction of parameters which do not sufficiently lower χ^2. An additional corrective term to AIC is necessary when the sample size is small, as is the case for our data. This is called AICc (Sugiura 1978) and is calculated as

$$\text{AICc} = \chi^2 + 2p + \frac{2p(p+1)}{n-p-1}, \qquad (1.7)$$

where n is the number of data points fitted. Both AIC and AICc are related to Kullback-Leiber entropy (Kullback & Leibler 1951) and measure the information lost from representing some true distribution with a specific model. When provided with several models, all of which provide a statistically acceptable fit to the data (i.e. have $\chi^2_\nu \lesssim 1$), choosing the lowest AICc model minimizes information loss.

AI-VPFIT

`AI-VPFIT` (Artificial Intelligence `VPFIT`) is a new development in spectral modelling of high-resolution spectra. It was designed to automate the model building process outlined above, address the issue of human bias, and improve the objectivity and reproducibility of measurements. It combines a genetic approach introduced in `GVPFIT` (Bainbridge & Webb 2017) with Monte Carlo methods and advanced model-building algorithms to produce a model with the most information (i.e. the lowest AICc). The more comprehensive description of `AI-VPFIT` is provided in Lee et al. (2020). A short description is provided here in order to understand the models derived in Chapter 4.

Human involvement in `AI-VPFIT` is limited to specifying how the modelling should proceed. The user declares which type of parameters should be considered, but model-building is performed automatically. For example, the user specifies which atomic species are present in the data, how their lines are broadened (turbulently, thermally, or through a combination of both mechanisms), should $\Delta\alpha/\alpha$ be a free parameter, and which parts of the data (if any) require additional free parameters for continuum or zero level adjustments. The user must also specify which atomic species should be used to produce a preliminary model of the velocity structure, or a "primary species", as described below. Following this, `AI-VPFIT` follows a set of well defined six stages, based on a generational approach, to produce the final model. All stages rely on `VPFIT` for parameter estimation.

In Stage 1, a preliminary model of the entire dataset is constructed using the primary species' transitions. The primary species should have transitions which are likely to provide a reliable initial model for later stages. These should be strong yet unsaturated, such that the velocity structure can be determined with reasonable certainty. Initially, a single trial absorption line, with pre-defined[3] values of N and b, is placed at a random redshift within the absorption complex of the primary species and its best-fit parameters are estimated. If the primary species contains transitions with sufficiently large q-coefficients, $\Delta\alpha/\alpha$ may be left as a free parameter (using a single value for the entire model) not to bias the final result towards terrestrial α value. Otherwise, the $\Delta\alpha/\alpha$ parameter is included in Stage 2.

The AICc of this model (Generation 1) is computed and saved internally to `AI-VPFIT`. The complexity of the model is increased by randomly placing a second trial line and optimizing model parameters. The model is accepted as the parent for the next generation if the AICc decreased with respect to that of the parent. Otherwise, the model is rejected and `AI-VPFIT` places the trial line at a different redshift, repeating the process. The process continues until AICc does not decrease for N_{lines} consecutive trial line placements, the latter being a parameter defined by the user at input. Components can be removed if they fall below pre-defined threshold values for N or b during optimization.

Having found a preliminary model for the primary species data, in Stage 2, `AI-VPFIT` copies the velocity structure of the lowest AICc model from Stage 1 onto all other (secondary) transitions and ties the redshifts of individual velocity

---

[3]Trial lines always have $N = 1 \times 10^{12} \, \mathrm{cm}^{-2}$ and $b = 5 \, \mathrm{km \, s}^{-1}$ in our case.

components. Depending on user choice, the b-parameters of primary and secondary species can be related either by the turbulent (equal b for all species) or thermal (b is scaled by the atomic mass ratio of the two species) broadening mechanism, or allow for contributions from both. All free parameters are optimized simultaneously to fit the entire dataset, and the complexity of the model is again increased until AICc fails to decrease for 100 consecutive trials.

Stage 3 considers blends with lines of unknown species, called "interlopers". In this Stage, free parameters from the end of Stage 2 are temporarily fixed. Contrary to previous stages, trial lines are now allowed to be placed anywhere in the entire dataset (e.g., an interloper appears in Fe II $\lambda2344$, but not in any other transition). The interloper is kept if it decreases the AICc for the entire dataset. The number and placement of interlopers is revisited in Stage 5.

Stage 4 is a two-step approach to allow continuum and zero level corrections to be made. In the first step, all parameters fixed in Stage 3 are allowed to vary again and new velocity components are added until no AICc descent occurs for 100 trials. In the second step, parameters related to continuum and zero level correction are included (for transitions specified by user at input). At the end of this Stage, the model is already very good, but some procedures need repeating to adjust for the introduction of these new parameters.

Stage 5 addresses potential issues of the modelling process so far. For example, because continuum-level corrections were not yet considered, the modelling process might have previously compensated for the incorrect continuum estimate by placing additional lines at specific locations. The program therefore temporarily fixes all continuum-level and zero-level parameters and removes all previously found interlopers. The algorithm is returned to Stage 2, but model parameters are initialized to values at the end of Stage 4. Stages 2 through 4 are repeated in their entirety before proceeding to Stage 6.

Stage 6 deals with the problem of accidentally lost components and over-fitting. Some heavy element lines might have been incorrectly dropped if, e.g. the column density falls below a previously defined dropping criterion. To recover from this, `AI-VPFIT` reproduces the velocity structure of the primary species into all secondary species, removes all dropping criteria, and limits the step size for all relevant parameters. The latter stabilises the calculations and prevents irrecoverably moving the parameters away from their optimal values. Furthermore, the continuum and zero level parameters are fixed to their values at the end of stage 4, and the redshifts of all heavy elements and interlopers are fixed. A more detailed explanation is provided in Lee et al. (2020). Finally, all fixed parameters are allowed to vary once more and parameters are refined. It is possible that the number of interlopers and heavy element components included in the model is too large, i.e. that the data have been over-fitted. This is tested for by removing interlopers and components, one at a time, and keeping it only if AICc does not increase. The final `AI-VPFIT` model is the one with the lowest AICc at the end of this Stage.

1.3.2 Wavelength calibration of astronomical spectrographs

The value of α retrieved from the model depends on the accuracy with which one can measure wavelengths in the spectrum. Besides the spectral resolution and S/N (affecting the precision with which line centres can be measured in the data during modelling), the accuracy on α depends on the accuracy of the laboratory and astronomical instruments' wavelength calibration. Transitions useful for α measurements have laboratory wavelength measurements with a typical uncertainty of $20\,\mathrm{m\,s^{-1}}$, with the most precise measurements (e.g. for Mg) reaching $0.1\,\mathrm{m\,s^{-1}}$. To fully exploit this high accuracy of laboratory wavelengths, wavelength calibration of astronomical instrumentation should ideally be of comparable or better accuracy.

ThAr calibration methods

The vast majority of data used for α measurements in the literature were obtained using echelle spectrographs and calibrated using ThAr methods. This method relies on recording a high-resolution, high S/N spectrum of a ThAr hollow cathode lamp on the detector, against which the astronomical spectrum is compared (Breckinridge et al. 1975). Numerous sharp Th and Ar emission lines recorded on the detector are identified from atlases (e.g., that of Palmer & Engleman 1983), which catalogue wavelengths for several thousands lines in the optical and ultraviolet ranges. Although it has many advantages to other calibration techniques, the accuracy of the ThAr method is limited by lamp purity, the number of sharp lines available for calibration, large centroid errors for low flux lines, non-linear detector effects and saturation for high flux lines, ThAr line mis-identifications, uncertainties in line wavelengths, and unavoidable relative line strength variations with pressure inside the lamp. The best accuracy one can hope to achieve from the ThAr method, using the highest resolution instruments and best available data, is $\approx 30\,\mathrm{m\,s^{-1}}$ (see Section 4.2.2).

Instrument design and calibration procedures complicate matters further, as the ThAr exposure taken at the beginning of the night (common practice on, e.g. UVES) may be inappropriate for calibration of exposures taken later in the night. This can be due to changes in the instrument settings (e.g. exchange of the dispersion element), telescope slewing, uneven slit illumination, changes to the rotation angle of the atmospheric dispersion correction unit, or environmental effects such as temperature and pressure changes during the night. For this reason, some varying constant studies used additional calibration information, e.g. asteroid or Solar-twin observations taken close in time and using the same instrument settings as used for scientific observations (Molaro et al. 2008; Rahmani et al. 2013; Whitmore & Murphy 2015) or inserting iodine absorption cells into the light path (Griest et al. 2010; Whitmore et al. 2010).

The aforementioned studies revealed the presence of wavelength distortions, i.e. differences in the wavelengths inferred from the calibration with respect to the true wavelengths of spectral features, in the calibration scales of UVES and HIRES, both extensively used for fundamental constant measurements. In a study by Whitmore & Murphy (2015) [but see also Molaro et al. (2008) for the first study
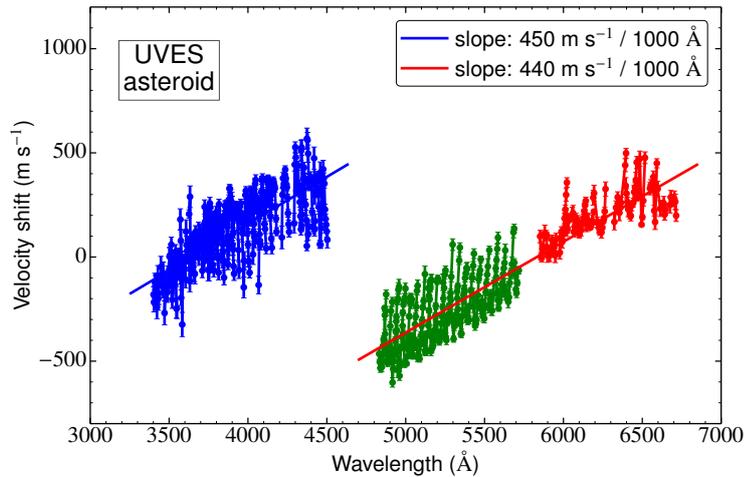
Figure 1.4: Asteroid observations revealed the presence of long range wavelength scale distortions (coloured dots and lines) in the UVES wavelength calibration. These distortions, i.e. differences in measured wavelengths of spectral features compared to their true values, systematically shift the wavelengths of measured quasar absorption lines in a way which may go some way towards mimicking variations in fundamental constants. Similar distortions exist for the HIRES instrument (see, e.g., figure 7 in Whitmore & Murphy 2015). Their influence can be minimised by appropriate modelling (Dumont & Webb 2017). However, in this thesis, the aim is to show how to avoid the distortions altogether using Laser Frequency Combs. Figure taken from Whitmore & Murphy (2015).

of the sort], a large number of solar lines (reflected from an asteroid) had their wavelengths measured from the spectrum wavelength calibrated standard ThAr procedures and compared to significantly more accurate wavelength measurements obtained using the Fourier Transform Spectrometer (Chance & Kurucz 2010). The comparison revealed the distortion pattern, illustrated on Figure 1.4, with an approximately linear long-range trend and an amplitude of $450\,\mathrm{m\,s^{-1}}$ across $1000\,\text{Å}$ (Whitmore & Murphy 2015). Additional short-range distortions are seen in individual echelle orders. Whilst these distortions can be minimized by using this additional calibration information (e.g. Molaro et al. 2008; Rahmani et al. 2013; Evans et al. 2014) or included into the spectral modelling process (e.g. Dumont & Webb 2017), they are limited to the accuracy of the ThAr calibration, $\approx 30\,\mathrm{m\,s^{-1}}$. Therefore, reaching wavelength accuracy comparable to that of laboratory measurements requires new ways of calibrating astronomical spectrographs.

Laser frequency comb for astronomy

Laser frequency comb (LFC) systems promise to provide wavelength calibration accuracy limited only by the number of incident photons. Originally developed for precise frequency metrology in laboratories (Udem et al. 2002; Hänsch 2006), LFCs were recognised as ideal wavelength calibrators for astronomical spectrographs. Preliminary considerations of astronomical LFCs are presented in Oster-

man et al. (2007) and Murphy et al. (2007), and the technology was soon adapted into astronomical LFCs, also known as *astrocombs*. Several such systems were prototyped (Steinmetz et al. 2008; Wilken et al. 2010; Ycas et al. 2012; Phillips et al. 2012; Probst et al. 2015a; McCracken et al. 2017b) and later installed on spectrographs such as HARPS (Probst et al. 2015a), FOCES (Brucalassi et al. 2016), ESPRESSO (Frank et al. 2018), HARPS-N (Ravi et al. 2019b), EXPRES (Blackman et al. 2020), and others.

The LFC produces thousands of emission lines equidistant in frequency space and stabilised to an atomic clock reference. The frequency (and by extension, the wavelength) of each line is known with extremely high accuracy, $\Delta f/f = 10^{-11}$ or several $\mathrm{mm\,s^{-1}}$. The most important astronomical applications of the LFC technology pertains to searches for Earth-mass exoplanets around Solar-mass stars, testing the immutability of fundamental constants, and measuring the acceleration of universal expansion in real time. These projects are major science drivers for the construction of the new generation of 30-m class telescopes, such as the ESO's Extremely Large Telescope (Hook 2009; ESO ELT team 2010, 2011).

1.4   Structure of the thesis

The first Chapter of this thesis introduced the concept of the *laws of physics* and presented some of the outstanding problems in modern fundamental physics and cosmology – the requirement for unknown particles (cold dark matter) and a type of exotic energy (dark energy) in the universe – for which we currently have no fundamental theory. Significant effort has gone into producing theories which should provide a physical mechanism behind the dark matter and dark energy phenomena, thus expanding (or reformulating) the laws of physics. References to this work are given in the same Chapter.

Many extensions to the current formulation of the laws of physics predict that the values of fundamental constant should vary due to their coupling to the new degree(s) of freedom. One such constant is the fine structure constant, α. A series of analyses of large samples of quasar absorption systems have found evidence for a spatial variation in α across the sky Webb et al. (1999, 2001); Murphy et al. (2003a); King et al. (2012); Wilczynska et al. (2020). This variation, best modelled as a dipole, is $\approx 4\sigma$ significant over the model in which α is at its terrestrial value everywhere in the universe. That analysis comprised approximately 300 measurements, half from the Keck telescope and half from the VLT (Section 1.2.4). Later, independent studies identified a systematic effect associated with wavelength calibration distortions of astronomical spectrographs which might have biased these measurements.

This thesis introduces the tools necessary to remove this systematic uncertainty using laser frequency combs for wavelength calibration, and then presents the first measurement of α derived from observations void of wavelength scale distortions; thus making new ground in the field of varying constants, fundamental physics, and cosmology in general.

The following three chapters present the bulk of the research performed during the writing of this thesis. Two Chapters are published as research articles in

international journals, whist the third Chapter is submitted for review. Chapter 2 presents an analysis of data collected during the commissioning of an LFC on the High Accuracy Radial velocity Planet Searcher (HARPS) instrument in 2015. The same Chapter focuses on describing the hardware of the LFC and presents the results of a careful analysis of its performance at the end of the LFC development. At this time, a second astrocomb was also installed on HARPS and the performance of the two systems was compared. Chapter 3 presents another analysis of the same dataset, this time focused on developing optimal wavelength calibration methods for high-resolution astronomical spectrographs using LFCs. Chapter 4 presents new α measurements from the first quasar spectrum calibrated using an LFC system. The spectrum was produced from HARPS observations of the quasar HE0515$-$4414 and was subsequently analysed using new, advanced algorithms, producing unbiased and robust results.

Thesis conclusions, together with an overview of the most significant results that came out of my research (and some other projects I was involved in) are given in Chapter 5. The results presented here should be relevant for other projects, most notably the ESPRESSO spectrograph and the future Extremely Large Telescope. Finally, I place my research into the context of these new facilities.

Chapter 2

A crucial test for astronomical spectrograph calibration with frequency combs

2.1 Overview

Laser frequency combs (LFCs) are well on their way to becoming the next-generation calibration sources for precision astronomical spectroscopy (Wilken et al. 2012; Molaro et al. 2013b; Glenday et al. 2015; Yi et al. 2016; Löhner-Böttcher et al. 2017; Obrzud et al. 2019). This development is considered key in the hunt for low-mass rocky exoplanets around solar-type stars whose discovery with the radial-velocity method requires $\mathrm{cm\,s^{-1}}$ Doppler precision (Fischer et al. 2016). In order to prove such precise calibration with an LFC, it must be compared to another calibrator of at least the same precision. Being the best available spectrograph calibrator, this means comparing it to a second fully independent LFC. This test had long been pending, but our installation of two LFCs at the ultra-stable spectrograph HARPS presented the so far unique opportunity for simultaneous calibrations with two separate LFCs. Although limited in time, the test results confirm the $1\,\mathrm{cm\,s^{-1}}$ stability that have long been anticipated by the astronomical community. First developed for laboratory-based spectroscopy, where they triggered a spectacular gain in accuracy, LFCs have become widely used tools for precision metrology (Udem

et al. 2002). An LFC is generated by a mode-locked laser that is phase-stabilized to an accurate radio-frequency (RF) reference such as an atomic clock. Its optical spectrum consists of a series of equally spaced, narrow spectral lines (modes), whose frequencies are known to the accuracy of the RF reference. The frequency of the nth mode is $f_n = f_0 + n f_r$, with f_0 the offset frequency and f_r the mode spacing. For astronomical echelle spectrographs, this regular pattern of lines, whose frequencies can directly be traced back to the SI second, comes close to an ideal calibrator (Murphy et al. 2007). LFCs for astronomical applications have thus been developed (Steinmetz et al. 2008; Wilken et al. 2012; Ycas et al. 2012; Probst et al. 2014; Glenday et al. 2015; Yi et al. 2016; McCracken et al. 2017a,b; Obrzud et al. 2019; Suh et al. 2019), which stand out through their particularly large mode spacing of $> 10\,\mathrm{GHz}$, allowing spectrographs to resolve the mode structure. When monitored on a second spectrograph channel during observations, LFCs permit tracking spectrograph drifts more precisely than ever before (Wilken et al. 2012; Glenday et al. 2015). LFCs are about to become the future standard calibration sources, replacing thorium-argon arc lamps. The lamps currently limit the attainable precision as their spectrum suffers from strongly irregular line intensity and spacing, blended lines, saturation effects, and from line drifts while the lamp ages.

The resulting gain in precision would greatly benefit exoplanet searches with the radial-velocity method, which looks for periodic Doppler shifts in stellar spectra caused by the gravitational interaction with orbiting planets. This method has enabled the discovery of the first exoplanet around a Sun-like star (Mayor & Queloz 1995) and continues to yield numerous important discoveries including Earth-mass planets (Anglada-Escudé et al. 2016; Astudillo-Defru et al. 2017) and super-Earths (Pepe et al. 2011). However, the technique has so far been insensitive to Earth-mass planets in the habitable zone of Sun-like stars. An Earth-Sun analogue would manifest in a radial-velocity variation of only $\pm 9\,\mathrm{cm\,s^{-1}}$ when viewed edge-on, with a 1-year period. On such time scales, thorium-argon lamps are limited in stability to the $\mathrm{m\,s^{-1}}$ level, whereas LFCs are expected to provide a $1\,\mathrm{cm\,s^{-1}}$ calibration precision over arbitrary time horizons (Murphy et al. 2007). LFCs are also anticipated to enable two astronomical tests of fundamental physics: (1) direct measurement of the cosmic expansion rate history, requiring a $1\,\mathrm{cm\,s^{-1}}$ precision over about two decades (Liske et al. 2008); and (2) measuring the potential cosmological variation in the value of fundamental constants (Webb et al. 2011).

The HARPS spectrograph (Mayor et al. 2003) is one of several cutting-edge spectrographs that have recently been equipped with an LFC (Glenday et al. 2015; Brucalassi et al. 2016; Probst et al. 2016; Löhner-Böttcher et al. 2017; McCracken et al. 2017b; Hao et al. 2018). After a number of test campaigns to demonstrate the technique (Wilken et al. 2010, 2012; Molaro et al. 2013b), HARPS has been given a permanently installed LFC (LFC1) as its future routine calibrator in May 2015. The installation was accompanied by the temporary deployment of a second LFC (LFC2), which was thereafter installed at the Wendelstein Observatory for operation with FOCES (Pfeiffer et al. 1998; Brucalassi et al. 2016). The scientific goal of the campaign was to characterize the relative performance of two LFCs in a series of repeated simultaneous calibrations in the two HARPS input channels.
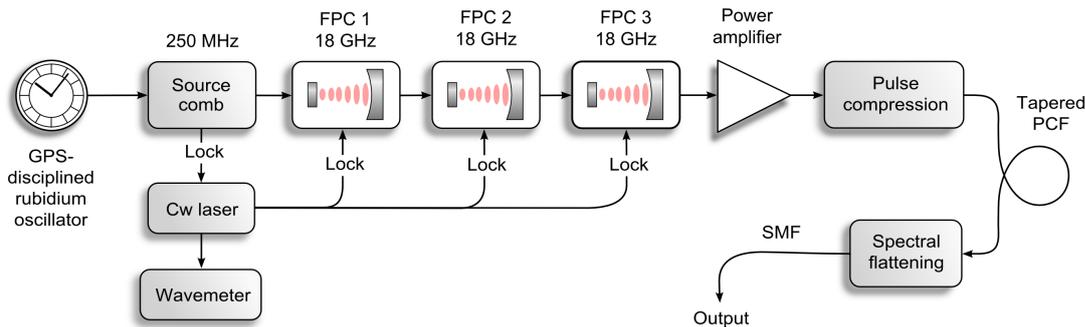
Figure 2.1: Setup of the Laser Frequency Comb (LFC) system. An Yb-fibre-based LFC with 250 MHz mode spacing is used as a source comb and is stabilized to GPS. A series of three identical Fabry-Prot cavities (FPCs) increases the mode spacing to 18 GHz (or 25 GHz in the case of LFC2). The FPCs are stabilized in length by a continuous-wave (cw) laser, that itself is stabilized on a transmitted comb mode. Monitoring of the wavelength of this laser with a wavemeter reveals which subset of modes is transmitted through the FPCs. After amplification and pulse compression, the filtered comb spectrum is broadened in a tapered photonic crystal fibre (PCF). Finally, the spectrum is reshaped into a flat-top by the spectral flattening unit. SMF: single-mode fibre.

Similar studies had previously been made with a single LFC for both channels, which indicated a stability of around $2\,\mathrm{cm\,s^{-1}}$ (Wilken et al. 2012; Glenday et al. 2015). Such tests are well suited to disclose potential uncertainties from the spectrograph light injection, imaging system, and image read-out mechanism. Strictly speaking, however, they are incomplete, since they do not rule out any unidentified uncertainties from the LFC itself, which should be common mode in the two channels. This is solved with our relative measurement of two independent LFCs. We therefore report on the most rigorous and precise test conducted so far for proving LFCs as precision calibrators for astronomy.

The setup of the LFCs is shown in Figure 2.1. As a light source, each LFC uses a mode-locked laser with 250 MHz mode spacing and 1040 nm centre wavelength, phase-stabilized to a GPS-disciplined quartz oscillator as an RF reference. The mode spacing is multiplied to 18 GHz (LFC1) through 3 identical Fabry-Prot cavities (FPCs) with a finesse of 2600, suppressing all but 1 out of 72 modes. The series of 3 FPCs ensures sufficient suppression for a $1\,\mathrm{cm\,s^{-1}}$ calibration accuracy (see Section 2.2.1). The FPCs are stabilized in length by a continuous-wave laser in a Pound-Drever-Hall scheme (Probst et al. 2014). The filtered comb light is then amplified to 12 W of average power, and compressed to a train of ultrashort pulses of 130 fs duration in a grating-prism compressor. This generates sufficient peak power to drive spectral broadening in a tapered photonic crystal fibre (Probst et al. 2015b), which extends the initially infrared spectrum into the visible range. Finally, the broad but structured spectrum is reshaped into a flat-top in the spectral flattening unit (Probst et al. 2013, 2015b) allowing all comb lines to be of roughly equal signal level on HARPS. LFC2 has a larger mode spacing of 25 GHz, which is matched to the lower resolution of FOCES (Brucalassi et al. 2016).
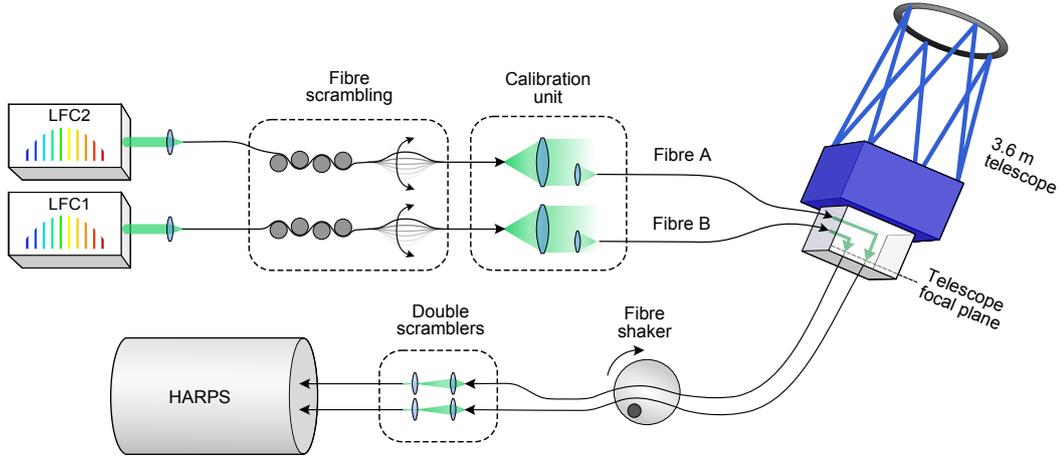
Figure 2.2: Coupling of instruments through optical fibres. The output of each LFC is injected into a multimode fibre subject to static bends and agitation. In the calibration unit, the fibre output is projected on another set of fibres A and B having motorized fibre inputs and couplers. This allows the operator to select different calibration sources by moving them to different slots. The fibres guide the light to the telescope, where it is projected through the image plane into the fibres leading to HARPS. The last set of fibres is again equipped with a combination of static and dynamic scramblers.

Both LFCs are coupled to HARPS through a sequence of multimode fibres (Figure 2.2). While astronomical applications generally favour multimode fibres over single-mode fibres, as they facilitate efficient throughput of light from astronomical sources, they also come with the issue of modal noise (Mahadevan et al. 2014): the beam profile at the output of a multimode fibre depends on light injection conditions at its input and on fibre bend. This is particularly true for coherent light as emitted by an LFC, which acquires laser speckles through modal interference (Wilken et al. 2010). As a countermeasure, we agitate the first and last set of fibres with electric motors. This makes the speckle pattern change quickly, while the spectrograph averages over it with its much longer exposures. In addition, we have a static bend structure in place on the first fibre pair in the sequence, as well as a double scrambler (Hunter & Ramsey 1992) on the last pair. These static scramblers globally homogenize the beam by coupling different spatial fibre modes to one another.

Figure 2.3a shows a part of the echellogram recorded with HARPS using the two LFCs. The full echellogram is shown in the Figure 2.8. From the two-dimensional image, we extract one-dimensional spectra by projecting each echelle order of each channel on the spectral direction (Figure 2.3b). For spectrograph calibration, we determine the line centres by fitting each line with a Gaussian function (see Figure 2.3c). A coarse thorium-argon calibration is used for mode identification to unambiguously assign frequencies f_n. Figure 2.3d shows the line amplitude spectrum, adjusted by the grating blaze function. The flat-top region ranges from 455 to 691 nm (76% of the HARPS spectral range), and is flat within 13% (root-
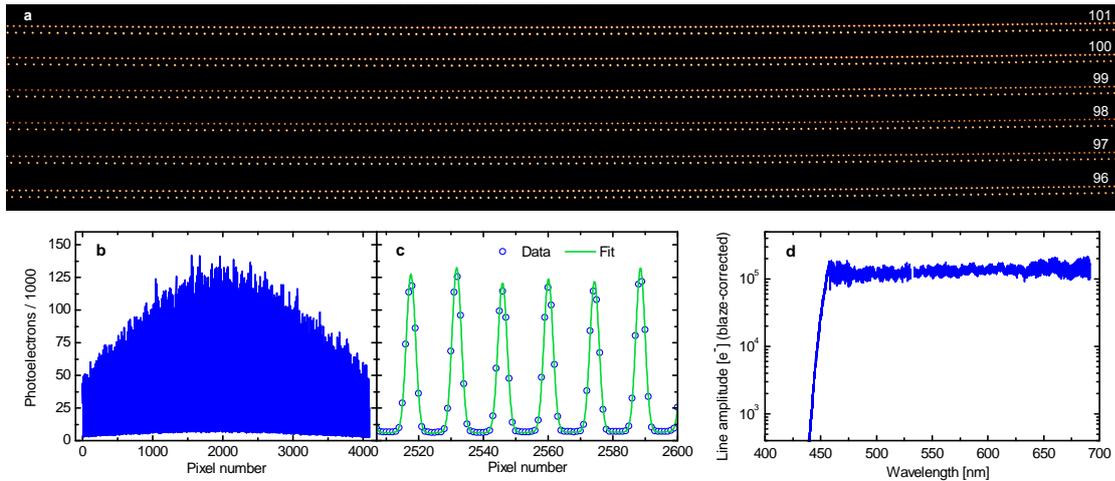
Figure 2.3: Two laser frequency combs (LFCs) on the HARPS spectrograph. Panel (a): Part of the echellogram with the echelle orders labelled with their physical diffraction order. The upper part of each order is channel A (here using LFC1, 18 GHz mode spacing), and the lower part is channel B (LFC2, 25 GHz mode spacing). Panel (b): Channel A, order 99 (centre wavelength: 618 nm) after data extraction. Panel (c): Section of the data in part b with a Gaussian function fitted to each line. Panel (d): Fitted peak values of the LFC lines, corrected by the blaze function of each echelle order.
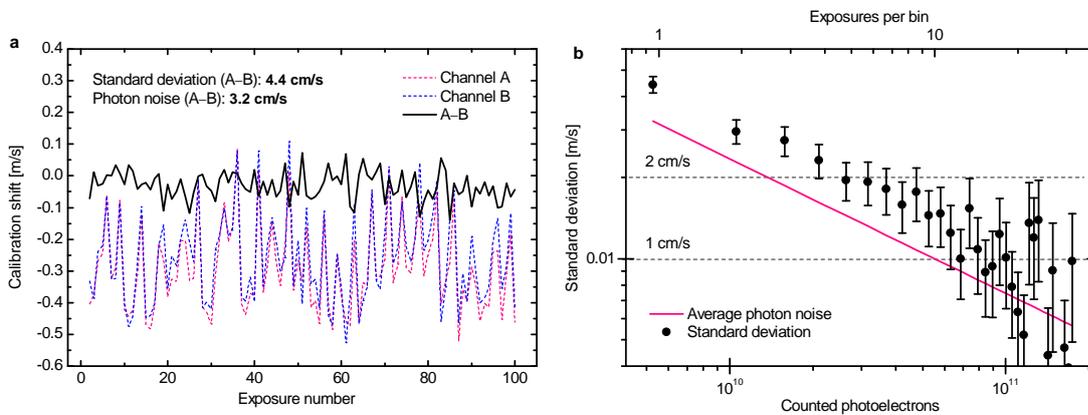
Figure 2.4: Relative stability measurement of two laser frequency combs (LFCs). LFC1 (18 GHz mode spacing) is in channel A and LFC2 (25 GHz mode spacing) in channel B. Panel (a): Series of 100 spectrograph calibrations with one exposure every 61 s (integration time: 30 s, readout time: 22.6 s), total duration of the series: 102 min. Panel (b): Results obtained with binned exposures of increasing size. The filled circles represent the standard deviation in A−B. The error bars quantify the uncertainty of the standard deviation estimated from the size of the statistical sample.

mean-square).

To assess the relative stability, we repeatedly calibrate the spectrograph and measure by how much each exposure is shifted relative to a reference exposure at
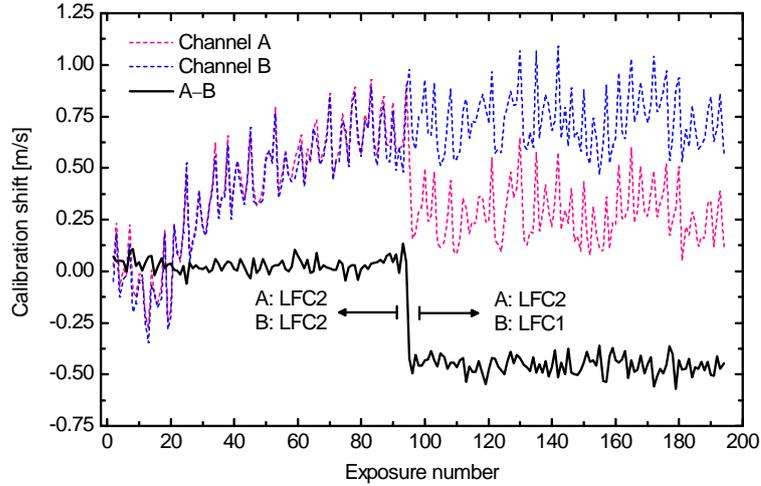
Figure 2.5: Calibration reproducibility with two different LFCs: mean shift in line positions. After exposure number 94 of 194, the source in channel B is changed from LFC2 to LFC1. LFC2 is continually kept on channel A to track spectrograph drifts. One acquisition was taken every $61\,\mathrm{s}$ (integration time: $30\,\mathrm{s}$. Readout time: $22.6\,\mathrm{s}$), with a 100 min pause before the change to the other LFC. The analysis was performed by comparing mean shifts of individual lines. Changing the source in channel B leads to an average shift of $49\,\mathrm{cm\,s^{-1}}$, after taking into account the different line structure of LFC1 (see Section 2.2.3). We ascribe this shift to differing spectrographs illuminations (see text), not to the LFCs themselves.

the beginning of the series (Figure 2.4a). This tracks spectrograph drifts, which are generally assumed to be equal in both channels, within the limits given by photon noise. With LFC1 in channel A and LFC2 in channel B, the standard deviation in the differential shifts is of $4.4\,\mathrm{cm\,s^{-1}}$ with a photon noise of $3.2\,\mathrm{cm\,s^{-1}}$. We attribute the deviation from photon noise to residual modal noise, as the fiber scrambling configuration had to be well optimized to reduce excess noise to this level (see Section 2.3.3). Interestingly, the remaining uncertainty appears purely statistical in nature: if we bin subsequent exposures and repeat the analysis for the series of binned exposures, we see the stability improving along with the photon noise (Figure 2.4b). This continues down to a standard deviation of about $1\,\mathrm{cm\,s^{-1}}$, where the limited size of our statistical sample inhibits us from demonstrating a further improvement.

Besides their relative stability, we also compared the two LFCs for their absolute consistency. For this we changed the calibration source from LFC2 to LFC1 on one channel, while continuously keeping LFC2 on the other channel to safely track spectrograph drifts. This reveals a $49\,\mathrm{cm\,s^{-1}}$ systematic shift between the calibrations from different LFCs (Figure 2.5). This was validated using a second, independent analysis (see Section 2.3.4). The discrepancy can be explained by the differing illumination of the fibres A and B in the calibration unit and/or variation in the intensity of the LFC light. When altering the alignment at this point, we observed comparable shifts despite the use of mode scramblers on the subsequent
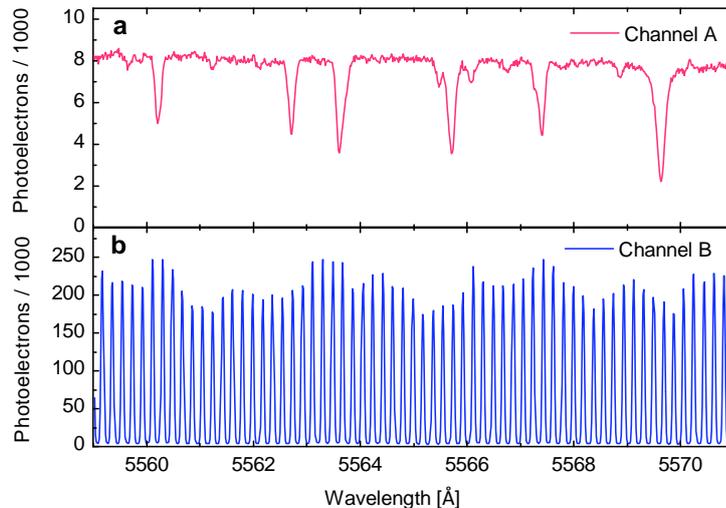
Figure 2.6: Observation of Ceres. Panel (a): Ceres spectrum observed in channel A, order 110, integrated over 900 s. The wavelength scale has been calibrated with the LFC at the start of the night and was continually adjusted for spectrograph drifts using the LFC that was kept on Channel B. Panel (b): 25 GHz LFC spectrum that was observed in channel B at the time of the Ceres observation.

fibres (see Section 2.3.3).

The LFC as a calibrator is not only extremely stable and reproducible, it also features unparalleled accuracy. We demonstrate how to make use of this property by verifying the absolute velocity of a solar system body. For this we select the dwarf planet Ceres, whose light features the reflected, Doppler shifted solar spectrum. The spectrograph was calibrated with LFC1 (channel A) and LFC2 (channel B) before the observation. While Ceres was observed on channel A, LFC2 remained on channel B to adjust the calibration for spectrograph drifts (Figure 2.6). By cross-correlating the observed, calibrated spectrum with a mask modelling solar spectrum, we measure a Doppler shift of $-21\,800.6\,\mathrm{m\,s^{-1}}$. From JPL Horizons[1] we expect this value to be $-21\,797.6\,\mathrm{m\,s^{-1}}$, thus showing our ability to measure absolute Doppler velocities with about $3\,\mathrm{m\,s^{-1}}$ accuracy. At this level, our measurement is limited by solar activity and by inhomogeneity of the albedo of Ceres in combination with its rotation (Lanza et al. 2016). Observing sunlight reflected from solar system bodies offers the benefit of evenly averaging the solar spectrum across the entire solar disc. In combination with the accuracy provided by the LFC, this opens up new possibilities such as an accurate measurement of the solar gravitational redshift by comparing various spectral lines formed in different altitudes of the Sun to laboratory wavelengths.

In this Chapter we have shown two LFCs with a relative stability of $1\,\mathrm{cm\,s^{-1}}$ over 102 minutes. Since the two LFCs are fully independent, each LFC by itself must be stable to this level. Common to both LFCs is merely the GPS reference, which features proven stability (see Section 2.2). The most demanding astrophys-

---

[1]Available at: https://ssd.jpl.nasa.gov/horizons.cgi

ical applications require the level of stability demonstrated in this Chapter, but maintained over years or decades. This calls for a comparison of two LFCs with a very long time base. Our present work is an important step into this direction by probing the limiting statistical and systematic effects. Our results show the critical influence of reproducible fibre illumination conditions and mitigation of modal noise even more clearly than earlier investigations (Wilken et al. 2012; Hao et al. 2018). For finding Earth-Sun analogues, it will also be essential to develop more advanced methods for separating weak orbital signals from the noise created by stellar activity phenomena. Precise, LFC-calibrated observations are invaluable in this respect (Dumusque et al. 2015a).

2.2 Methods

2.2.1 Operation and characterization of the LFCs

During the measurements presented above, the LFCs were operated at a repetition rate (offset frequency) of 18 GHz (5.27 GHz) and 25 GHz (9.70 GHz), respectively. Both values can be altered within certain limits, which we did for some measurements in Section 2.3. This is indicated where being the case. For spectral broadening, the LFCs employed a tapered photonic crystal fibre (PCF) with a design very similar to the one described in Löhner-Böttcher et al. (2017); Wu et al. (2020). The main difference is the higher air-filling fraction (80%), which allows the PCF to generate a somewhat wider spectrum. The broad but structured spectrum was flattened out using a liquid-crystal-on-silicon spatial light modulator (SLM) as an adaptive spectral filter (Probst et al. 2013, 2014, 2015b). The filter was set to truncate the spectrum at 20 dB below its peak within the visible range. This is accomplished with the help of a small CCD spectrometer that is integrated in the spectral flattening unit. However, when the spectrum appears flat on the systems internal spectrometer, it is not measured to be flat on HARPS, due to differences in spectral sensitivity and wavelength-dependent losses in the feed-through. Therefore, we first recorded a flattened but uncorrected spectrum with HARPS. From this we derived a spectral correction factor, which could then be taken into account by the software controlling the flattening unit. This yielded the flat-top spectrum shown in Figure 2.3d. The procedure was carried out separately with each LFC.

The accuracy provided by the spectral lines of the LFC can be influenced by the weak presence of unwanted modes (side-modes) that are not fully suppressed by the Fabry-Prot cavities (FPCs) in the setup. With the side-modes being unresolved by HARPS, they can shift the centroid of the calibration lines, in case their intensities are asymmetric around the observed line centre. Using the method described in Probst et al. (2014), we measure the finesse of our FPCs to be 2588 (geometric average for the FPCs in LFC2). From this measurement, we calculate the suppression of the strongest side-mode after the series of three FPCs to be 111.4 dB for LFC1 and 102.8 dB for LFC2. The subsequent spectral broadening is known to re-amplify side-modes. An upper limit on the side-mode amplification can be gained from the width of the broadened spectrum (Probst et al. 2013),

yielding 62 dB with our PCFs. This method has been shown to be reliable in several measurements (Probst et al. 2013) including a PCF that is very similar to ours (Wu et al. 2020). The remaining side-mode suppression should ensure a calibration accurate to $0.2\,\mathrm{cm\,s^{-1}}$ for LFC1 and $1\,\mathrm{cm\,s^{-1}}$ for LFC2. The centroid of a mode can also be shifted through a distortion of its shape while being transmitted through an FPC, in case the filter function is not accurately centred on the mode. For a series of three FPCs with the measured finesse, we calculate the worst-case line shift to be $0.5\,\mathrm{cm\,s^{-1}}$ for LFC1 and $0.3\,\mathrm{cm\,s^{-1}}$ for LFC2. The dominant error is thus $0.5\,\mathrm{cm\,s^{-1}}$ for LFC1 and of $1\,\mathrm{cm\,s^{-1}}$ for LFC2. Both LFCs were connected to the same RF reference, which was a Datum Model 9390 atomic clock with a 10 MHz GPS-disciplined Rb oscillator, which was part of the existing infrastructure of the observatory. The unit had previously been characterized relative to another 10 MHz Rb oscillator (Standford Research Systems PRS10, not locked to GPS) over 24 hours using a frequency counter. Employing a 30 s integration time in this comparison (identical to the integration time we used in our HARPS exposures), the two RF signals proved to be stable within 5.6×10^{-12}. This corresponds to a radial velocity uncertainty of $< 0.2\,\mathrm{cm\,s^{-1}}$.

2.2.2 Light delivery to HARPS

Both LFCs are each coupled into a 1 mm thick acrylic plastic fibre with a 980 µm core and a numerical aperture (NA) of 0.51. We impose a static bend structure on this fibre, which couples the fibre spatial modes to one another. This globally homogenizes the far-field beam profile at the fibre output, thereby making it less dependent on the illumination conditions at the input, which reduces the related systematic calibration uncertainties. However, for a highly coherent light source such as an LFC, such static scramblers are not sufficient. Interference between different spatial fibre modes creates a speckle pattern at the fibre output, which changes with the slightest motion of the fibre, inducing calibration errors on the $\mathrm{m\,s^{-1}}$ scale (Wilken et al. 2010). Therefore, we attach a small electric motor to the fibre. The motor spins an eccentric weight to make the fibre vibrate. This causes the speckle pattern to change quickly, making the light behave similar to spatially incoherent light on the much longer time scales of the spectrograph exposures. The large NA and core diameter of the plastic fibre makes this process very efficient. This is because it supports a large number of spatial modes, and thus creates a large number of very small speckles, that are highly sensitive to the motions of the fibre. We have also tested a silica fibre with a 200 µm octagonal core in the place of the plastic fibre. The octagonal core by itself is an excellent static mode scrambler, as it is very effective in mixing spatial fibre modes. Yet, the calibration results with this fibre were consistently a factor of 45 above the photon noise limit. Most likely, this is due to its smaller core and lower NA, causing the fibre to support a lower number of spatial modes, which reduces the effect of the fibre agitation.

The two scrambled plastic fibres, that each carry the light of an LFC, lead to the calibration unit (see Figure 2.2), where their output produces a large spot of light, into which the motorized inputs of the fibres A and/or B can be moved to choose a calibration source for each channel. The fibres A and B are a set of

silica fibres with a round core of 300 μm in diameter. This pair of fibres leads to the telescope, where the light is projected through the image plane onto the entrance facets of the next set of fibres that feed the spectrograph. From this point, the calibration light takes the same path that star light from the telescope takes when used for spectroscopy. The last set of fibres has a 70 μm circular core. They incorporate a double scrambler (Hunter & Ramsey 1992) to redistribute the spatial modes within the fibre, which is an excellent static mode scrambler. The purpose of the double scrambler is to mitigate the effect of telescope guiding errors on spectroscopy of astronomical objects. This is needed because guiding errors entail varying illumination of the fibre entrance. Although our calibration tests also profit from this static type of scrambler, we additionally installed a dynamic scrambler that agitates the last set of fibres. It consists of a rotating wheel with an off-centred support moving the fibres up and down. While testing the relative stability of the two LFCs, we could only come close to the photon-noise limit with this fibre shaker added to the last set of fibres (see Section 2.3.3).

2.2.3 Data acquisition and processing

The HARPS data were typically recorded in sequences of spectrograph exposures with one exposure per minute. For the data shown in the main article we used a 30 s integration time and a 22.6 s readout time. The images are first processed with the HARPS pipeline which automatically subtracts a dark image and the detector bias, performs spectral localization, flat-fielding, cosmic ray removal and spectral extraction (optimal extraction after Horne (Horne 1986)). The extracted spectrum contains 4096 data points per echelle order in every channel (see Figure 2.3b for an example of a single extracted echelle order). The centre positions of the comb lines are determined by fitting them with a model function. Experimenting with several line models, we found that a simple Gaussian fit leads to unreliable results. The reason for this is the relatively strong continuum background in the LFC spectrum, that keeps the signal from dropping to zero between the calibration lines (see Figure 2.3b, c, and Section 2.3.2). The background can have a non-zero slope that if not properly taken into account by the data analysis can shift the detected line positions. Remarkably, the background level consistently follows the structure of the envelope of the comb lines. Tests that we conducted with strongly structured spectra revealed global calibration errors of up to $1\,\mathrm{m\,s^{-1}}$ using a simple Gaussian fit. Hence, we add a linear polynomial to the fit, modelling each line as a Gaussian function and the background around each line as a first-order polynomial:

$$f(x) = a_0 + a_1 x + a_2 \exp\left[-\frac{(x - a_3)^2}{a_4^2}\right] \qquad (2.1)$$

Here, x is the position on the sensor in pixels. We define a_3 as the centre position of the line. This function is fitted to the data using chi-square minimization with the Levenberg-Marquardt method. The uncertainties of the data points in the fit are given by $\sqrt{|N| + R^2}$, where N is the number of photons detected in each pixel, and R is the readout noise. Besides the fitted parameters, the fit routine also returns the uncertainties of the parameters computed through Gaussian error propagation

from the uncertainties of the data points, which are usually dominated by photon noise. The primary interest of our data analysis lies in a precise assessment of the relative shifts in the calibration. To this end, we determine the overall shift in each channel for each exposure relative to a reference exposure. The reference exposure is consistently chosen to be the first exposure (exposure number 1) of each analysed sequence of exposures, unless stated otherwise. Our standard way of measuring this shift is to average the individual shifts of all lines, weighting each line by its inverse variance in a_3 as returned by the fit routine. This yields traces for channel A (red dashed line) and channel B (blue dashed line) as seen in Figure 2.4 and Figure 2.5. We remove spectrograph drifts by taking the difference between the two traces (AB, black solid line), thus revealing the relative stability of calibrations in the two channels. We use the standard deviation of AB as a measure of stability and compare it to the associated photon noise (compound uncertainty in a_3 over all lines). Note that switching channel B to carry light from LFC1 after exposure 94 in Figure 2.5 requires interpolation between the comb lines in order to make the calculation of line shifts meaningful. We do this by linearly interpolating between the lines of LFC2 in the reference exposure to derive the expected line positions of LFC1 and their uncertainties.

2.2.4 Ceres observation

Ceres was observed over 66 minutes on April 18, 2015. The observation was made in four separate spectrograph exposures, each integrated over 900 s. The photometric centres of the exposures are: 08:33:13, 08:50:31, 09:06:59 and 09:24:04 UTC. The Ceres spectra were cross-correlated with a spectral mask based on a list of solar lines at laboratory wavelengths with 3625 lines within the spectral range of the LFC. In the actual solar spectrum, these wavelengths are shifted by effects such as the convective blue shift and the gravitational redshift. Empirical use of this mask has shown that this shifts the derived radial velocities by $99.5 \, \mathrm{m \, s^{-1}}$ (average over multiple observations Lanza et al. 2016), which we subtract from the Doppler shifts that we measure. The solar spectrum is calibrated by creating a wavelength solution from the known optical frequencies of the LFC lines versus their observed positions on the detector. This is constructed as a piecewise 3rd order polynomial across each master block of 512 pixels in width. This allows us to take into account stitching errors from the manufacturing process of the CCD (Wilken et al. 2010; Molaro et al. 2013b). The calibration is then adjusted by the spectrograph drift as seen with the LFC on channel B during the observation of Ceres. After cross-correlating the calibrated Ceres spectrum with the mask, the cross-correlation function is fitted with a Gaussian function, whose centre indicates the average shift of the lines relative to their positions in the mask. With this we determine the Doppler shift to be $-21\,800.6 \, \mathrm{m \, s^{-1}}$ on average over the four exposures (individual values: $-21\,831.1$ $-21\,812.1$ $-21\,791.2$, and $-21\,767.8 \, \mathrm{m \, s^{-1}}$). From the known orbit of Ceres relative to the Sun, and from the known motion of the observer relative to Ceres, we compute the predicted Doppler shift to be $-21\,797.6 \, \mathrm{m \, s^{-1}}$ on average (individual values: $-21\,829.0$, $-21\,808.2$, $-21\,782.2$, and $-21\,765.8 \, \mathrm{m \, s^{-1}}$).

Figure 2.7: Photo of the laser frequency comb (LFC1) after its permanent installation on HARPS. The optical setup is contained in the blue enclosures on the left-hand-side, while the racks on the right accommodate control electronics and powers supplies, as well as a continuous wave laser, a wavemeter, and diode lasers for pumping of the power amplifier (see Figure 2.1).

2.3 Supplementary material

2.3.1 Instruments

The HARPS instrument

The High Accuracy Radial Velocity Planet Searcher (HARPS) is a fibre-fed, cross-dispersed echelle spectrograph with two channels (Mayor et al. 2003; Rupprecht et al. 2004). It is located at the La Silla Observatory in the Chilean Atacama desert and is operated by the European Southern Observatory (ESO). The spectrograph is contained in a vacuum vessel with a pressure of below $0.01\,\text{mbar}$ and a temperature stability of $10\,\text{mK}$ at its interior, which minimizes spectrograph drifts. HARPS is equipped with one object and one reference fibre for simultaneous calibration, currently using thorium-argon lamps as standard calibrators. It covers a wavelength range of 380–$690\,\text{nm}$ with 72 echelle orders and a resolution of 115 000. Its camera consists of a mosaic of two CCDs, each with a format of $4096 \times 2048\,\text{px}^2$. HARPS is scientifically devoted to the search for extrasolar planets via radial-velocity measurements. Through its excellent stability, HARPS has become the most successful planet hunter of its kind.

HARPS has played a key role in introducing laser frequency combs (LFCs) for astronomical applications. In a collaboration comprising ESO, the Max Planck In-

stitute of Quantum Optics (MPQ), and Menlo Systems GmbH, a total of five test campaigns have been conducted on HARPS to demonstrate spectrograph calibration with an LFC. The campaigns were carried out in January 2009 (Wilken et al. 2010), March 2010, November 2010 (Wilken et al. 2012), January 2011 (Wilken et al. 2012), and February 2012. The LFCs used in these tests were developed at MPQ and Menlo Systems, while similar systems were also developed by other groups and tested at other spectrographs (Braje et al. 2008; Quinlan et al. 2010; Phillips et al. 2012; Ycas et al. 2012; Glenday et al. 2015). The test campaigns on HARPS yielded several ground-breaking results, such as a new and improved atlas of solar lines (Molaro et al. 2013b) and the first planetary orbit measured with an LFC (Wilken et al. 2012). Absolute calibration of HARPS was demonstrated with previously unparalleled accuracy, which allowed accounting for structures in the CCD pixelation that previously went unnoticed (Wilken et al. 2010; Molaro et al. 2013b). Further, a previously unmatched calibration repeatability of 2.5 cm was demonstrated (Wilken et al. 2012). Very similar results for the repeatability were later also attained with LFCs at other facilities (Glenday et al. 2015; Probst et al. 2015a). LFCs at spectrographs other than HARPS have scientifically been applied in observations of the Sun as a star (Probst et al. 2015a; Dumusque et al. 2015a). Such observations are used to search ways to discern stellar activity from the radial-velocity signature of low-mass planets. In another scientific study, an LFC was utilized to create a new atlas of uranium-neon lines (Redman et al. 2012).

The laser frequency combs of the 2015 campaign

In May 2015, HARPS was permanently equipped with an LFC that is presently being prepared to become its new routine calibrator. During the installation run, the permanently deployed LFC (LFC1) was characterized relative to a second, temporarily installed LFC (LFC2), which was shipped back to Germany after the test. In October 2016, LFC2 arrived at the Wendelstein Observatory its final site of operation as the calibration system for the recently upgraded FOCES spectrograph (Brucalassi et al. 2016). Given the spectral resolution of HARPS of 115 000, an optical resolution element has a 5 GHz full-width at half-maximum (FWHM) in the spectrographs mid-spectral range. The FWHM being sampled by 3.3 pixels, the 15 μm width of a single pixel spans 1.5 GHz, or $0.8 \, \text{km s}^{-1}$. Theoretically, the optimal mode spacing of an astronomical LFC is 3 times the FWHM of an optical resolution element (Murphy et al. 2007). For HARPS, we have decided for a slightly wider spacing of 18 GHz for LFC1. This keeps the calibration lines well apart with virtually no residual overlap, which facilitates data analysis. The mode spacing of LFC2 is 25 GHz, adapted to the slightly lower resolution of the FOCES spectrograph of 70 000. Figure 2.7 shows a photograph of LFC1 after its permanent installation on the site. The configuration of the LFCs is explained in the main article. A more detailed description is found in Probst et al. (2014). It should however be noted that, contrary to the LFC described in Molaro et al. (2013b), the present systems LFC1 and LFC2 no longer employ a second-harmonic generator (SHG). Earlier versions used the SHG to transfer the infrared (IR) laser spectrum into the green before spectral broadening. This is now omitted, as we

Figure 2.8: Full echellogram with different LFCs on the two channels of HARPS. The upper half of each order is channel A (LFC1, 18 GHz mode spacing) and the lower part channel B (LFC2, 25 GHz mode spacing). The original grey-scale image as recorded with HARPS has been coloured, using the known wavelengths of the diffraction orders, to approximate the perception of the human eye. The number of the physical diffraction order is annotated above each echelle order.

have developed tapered photonic crystal fibres (PCFs) that directly broaden the IR spectrum far enough to cover the visible range (Probst et al. 2015b). Figure 2.8 shows an echellogram with the two LFCs on HARPS. 43 echelle orders are illuminated. The spectrum of LFC1 contains about 12 000 spectral lines over the observed spectral range, many of which appear twice due to the spectral overlap of the echelle orders. The echellogram contains a gap in the green region, where one order of channel A and two orders of channel B are lost. The gap results from the space between the active areas of the two CCD chips that make up the mosaic camera.

2.3.2 Undesired spectral components

Background

In our initial tests of the 2015 campaign, we noticed an unusually strong background component to the LFC spectrum, which was not known from previous campaigns. This is clearly seen in Figure 2.9, showing how the intensity between the lines forms a plateau instead of dropping to zero. We also observe a clear and consistent tendency of the background to rise towards shorter wavelengths. The only substantial physical difference to earlier versions of our LFC consists in the spectral broadening strategy. Skipping the SHG simplifies the setup, but requires more nonlinear spectral broadening to be driven in the tapered PCF, which is known to create phase noise from amplitude noise. The noise background is very likely to be seeded by amplified spontaneous emission (ASE) from the high-power Yb-fibre amplifier. The amplifier must now generate more optical power, thereby plausibly emitting more ASE.

In fact, this view is confirmed by our later investigations on this matter using HARPS, FOCES, and a home-built echelle spectrograph (Probst et al. 2017). Optimizing the core diameter of the gain fibre and its length can reduce the amount of ASE produced, which lowers the background. In case that the ASE spectrum does not fully coincide with the signal of the infrared laser source, it is possible to block a part of the ASE, which again lowers the background. Careful optimization of the amplifier, pre-amplifier and broadening schemes, as well as minimization of losses, should allow us to significantly reduce the background level. For spectrograph calibration, this has the benefit of reducing the photon noise, because the photons forming the background are an additional source of photon noise and decrease the available dynamic range of the CCD (Probst et al. 2017). Furthermore, we found that the spectral background can lead to systematic errors if the background has a slope that is not taken into account by the fit function. With a simple Gaussian fit, we could observe this very clearly on strongly modulated spectra. Since the background level locally seems to follow the signal strength of the comb lines, the background can acquire a distinct slope through fine modulations in the spectral line intensities that lie below the resolution of the spectral flattening unit.

Due to a defect in a PCF that was tested near the start of the campaign, we observed strong spectral modulations at around 500 nm over about 40 nm (modulation period: 0.2 nm, depth: up to 10 dB peak-to-valley). We initially analysed these sequences of spectrograph exposures with a fit function that assumed the
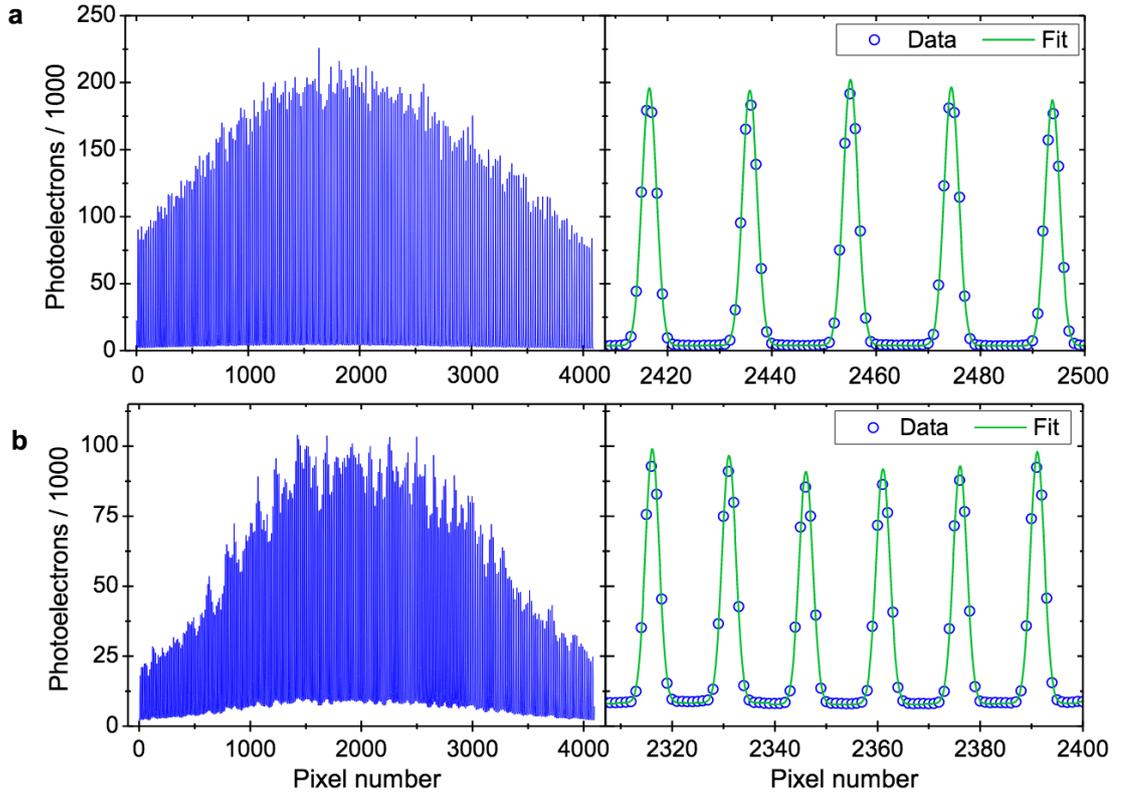
Figure 2.9: Echelle orders of channel B after data extraction, using LFC2 (25 GHz mode spacing). Left: data over the full order. Right: fit (green) to the data (blue) on a zoomed-in horizontal scale. Panel (a): Echelle order 100 (centre wavelength 612 nm). Panel (b): Echelle order 128 (centre wavelength 478 nm). Comparison to panel (a) shows a clear trend of the background to rise towards shorter wavelengths.

background to be constant over the data range used to fit a single line. With this model, we observed the calibration to experience global shifts of up to about $\pm 1\,\mathrm{m\,s^{-1}}$ in an oscillating fashion. This occurred in sync in both channels A and B. The shifts were about $\pm 4\,\mathrm{m\,s^{-1}}$ in the most strongly modulated parts of the spectrum, which was very clearly correlated with changes in fringe positions (the peaks moving to the positions of the valleys and vice versa). We concluded that the fringes imprinted a slope on the background, which followed the comb lines intensity, so our assumption of a constant background over the width of a line no longer held true. We thus changed our fit function to incorporate linear background model (see Section 2.2.3). This made the oscillatory calibration shifts disappear, even in the most structured parts of the spectrum. Both channels A and B were now stable within $14\,\mathrm{cm\,s^{-1}}$, limited by the passive stability of the spectrograph. We take this as proof that our linear model for the background in the fit function is a good approximation even in an extreme case of a structured background. In particular, this excludes that the $49\,\mathrm{cm\,s^{-1}}$ difference between the two LFCs (see Figure 2.5) is due to inaccurate modelling of the background.
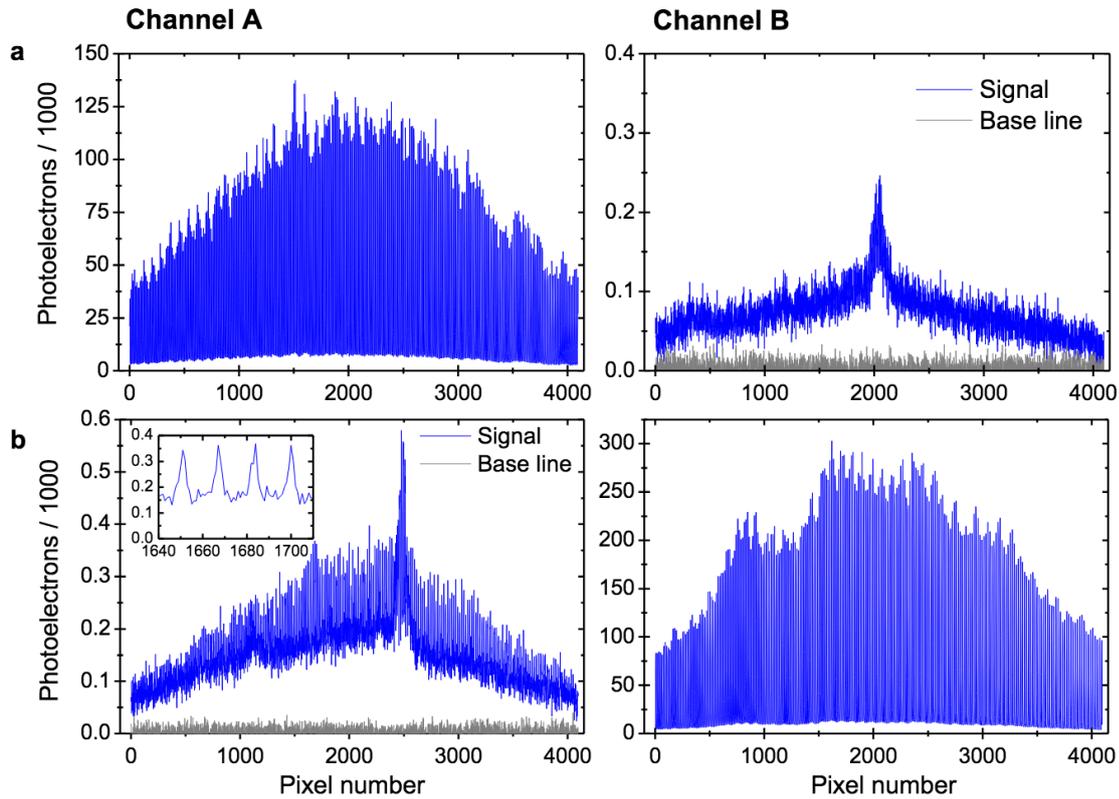
Figure 2.10: Crosstalk characterization. LFC1 is in channel A (left-hand side) and LFC2 in channel B (right-hand side). For measuring the crosstalk in each direction, an LFC is injected into just one channel, while observing the signal on the other one. The base line (grey) is a measurement without any input light on either channel. We deliberately show echelle orders that contain strong ghosting (71% relative to the strongest ghost found in channel A; 63% for channel B). Panel (a): Crosstalk from A to B in order 108 (centre wavelength: 567 nm). Panel (b): Crosstalk from B to A in order 111 (centre wavelength 551 nm). The inset shows that a comb spectrum is observed in channel A. Notice that in these measurements, LFC2 is more intense than LFC1.

Crosstalk and stray light

A minor contribution to the spectral background stems from stray light within the spectrograph, which originates mainly from imperfections of the grating. We characterize this contribution in the course of a more general examination of the crosstalk between the two channels. This is done by injecting light into one channel, and observing the signal on the other. Figure 2.10a shows the crosstalk from channel A to channel B. It consists of two components: stray light and ghosting. Ghosted diffraction orders cross the echelle orders at specific points, creating small artefacts in the spectrum. At their peak, this creates a crosstalk of 1.3×10^{-3} (fraction of the signal in channel A relative to channel B). The affected sections can be excluded from the analysis, which for us has so far never made a difference. The stray light component appears to be structureless and amounts to 8×10^{-4} as

a fraction of the signal measured in the other channel. Hence, it should not have an influence on spectrograph calibration.

As expected, the crosstalk in opposite direction (see Figure 2.10b) exhibits ghosting and stray light of the same level. However, contrary to the crosstalk from channel A to B, we here also find a third component that consists in spurious comb lines throughout the entire echelle order. This might either be explained through light pollution in the fibre feed or through charge spilling on the CCD. The effect creates a crosstalk of 5×10^{-4}. For the LFC this might cause a line shift of about $1 \, \mathrm{m \, s^{-1}}$ in a worst-case scenario. The combination of an $18 \, \mathrm{GHz}$ LFC with a $25 \, \mathrm{GHz}$ LFC leads to recurring relative line positions every $450 \, \mathrm{GHz}$. The effect is therefore not necessarily fully diluted when averaged over many lines. However, the effect should vanish if two LFCs with the same offset frequency and mode spacing were used, as the signal would only be polluted by another signal with identical lines. Hence, measurements using the same LFC is in both channels should not be impacted. The crosstalk can, however, at least partly be responsible for the absolute disagreement between the two different LFCs (Figure 2.5).

2.3.3 Optimization of fibre coupling and scrambling

For each channel, the light is fed through a series of three different multimode fibres before reaching the spectrograph (see Figure 2.3). The light of the LFC has high spatial coherence, as opposed to that of a thorium-argon lamp. Dynamic fibre scrambling is, therefore, of vital importance to reach a calibration that is stable within the photon noise limit. This scrambles the relative phase of the spatial modes propagating within the fibre, which blurs out their interference pattern (laser speckles) at the fibre output. We achieve this through agitation of the first and last fibre of the series. The shaker on the last fibre, however, was initially not used, because it is not foreseen for standard operations of the telescope and spectrograph. Instead, we attempted to fully rely on agitating the first fibre only. This requires special care, because in a sequence of several fibres, it is not guaranteed that the previously imposed relative dephasing of spatial modes is fully passed on to the next fibre(s). Depending on how several cascaded fibres are coupled to one another, some spatial modes of the later fibres might partly share a stable relative phase. The fewer spatial modes of a scrambled fibre are coupled to a non-scrambled fibre, the more the effect of the scrambling will be lost.

Finding a configuration that can best preserve the spatial mode scrambling was an important first goal of our campaign. This meant to optimize the alignment of the fibre illumination in the calibration unit. The optimization was first done with a single LFC on both channels (first with LFC1, then LFC2) before we moved on to testing the relative stability of the two LFCs. We characterized the performance of each configuration in sequences of exposures. The light injection into the fiber was optimized so as to minimize the standard deviation of the differential calibration shifts (AB) in the respective sequence. Figure 2.11 shows all exposures obtained in this way with LFC1 in both channels. The overall standard deviation in AB is $32.1 \, \mathrm{cm \, s^{-1}}$, while the average photon noise is $4.0 \, \mathrm{cm \, s^{-1}}$. The plot does not only reveal how some configurations provided better stability than others, but it also
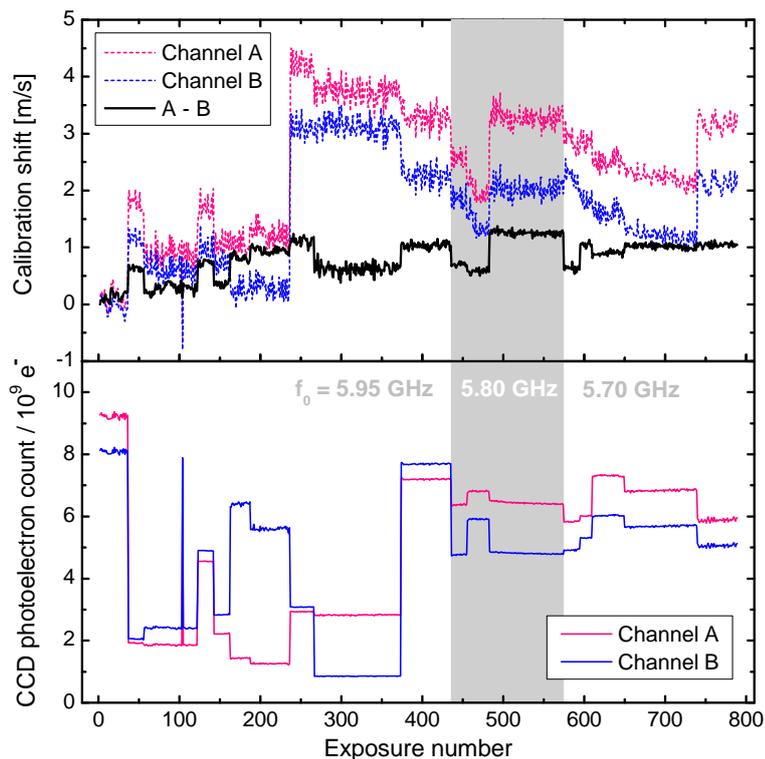
Figure 2.11: All spectrograph exposures with LFC1 on both HARPS channels. The graph concatenates various series of exposures with differing illumination of the multimode fibres. Panel (a): Calibration shifts relative to a common reference. As usual, calibration errors are to be seen from the differential shifts AB, since the channels A and B by themselves are influenced by spectrograph drifts. Panel (b): Total number of photoelectrons detected in each channel. The offset frequency is 5.95 GHz up to exposure 435, then 5.80 GHz until exposure 574, and finally 5.70 GHz for the remaining part of the data. Mode spacing: 18 GHz.

shows how a change in the alignment systematically shifts the calibration in AB. Changing the alignment also altered the throughput, which can in part explain the calibration errors. This is because charge-transfer inefficiency (CTI) of the CCD causes line shifts depending on signal strength (Bouchy et al. 2009; Zhao et al. 2014). However, although the calibration shifts and the changes in signal strength seem to have a common cause, they do not relate to each other in a simple and stable way. Trying to adjust the calibration shifts by assuming some simple relationship with signal strength (linear, inverse, exponential, or logarithmic) could reduce the standard deviation in AB to no less than $29.4\,\mathrm{cm\,s^{-1}}$. This suggests that the varying spatial mode occupation in the fibres is the dominant reason for the shifts, not the altered transmission. Both influences are side-effects of the changing alignment.

Although LFC1 was operated at three different offset frequencies, this is taken into account in Figure 2.11 and can by no means explain the shifts in AB. Moreover, the interpolation method that we use to relate LFC spectra with different parameters has been tested in an earlier campaign and shown to have no measur-
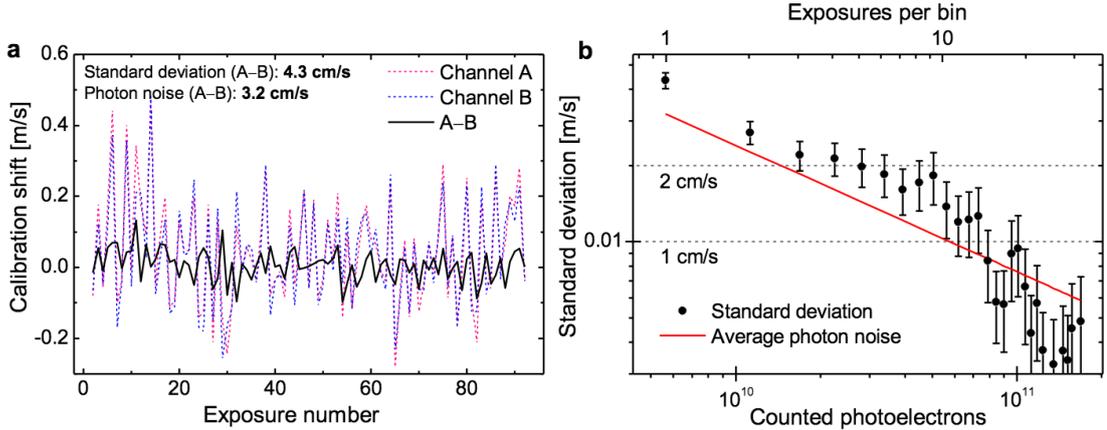
Figure 2.12: Calibration repeatability with LFC1 in both channels A and B. The alignment of the multimode fibre feed is kept stable during the complete series of exposures. The derived calibration shifts use exposure number 1 as a reference. The series is contained in Figure 2.11 as running from exposure number 483 to 573. Panel (a): Calibration shifts without binning of exposures. Panel (b): Combining several subsequent exposures into binned exposures. The standard deviation in AB is plotted as a function of the number of exposures per bin. The error bars quantify the uncertainty of the standard deviation estimated from the size of the statistical sample. One exposure was triggered every 92 s. Integration time: 40 s. Readout time: 22.6 s.

able impact on calibration down to at least $7 \, \mathrm{cm \, s^{-1}}$ (Probst 2015).

After having optimized the fibre injection in the calibration unit as explained above, we achieved nearly photon noise-limited results with both LFCs. A typical series with LFC1 is show in Figure 2.12. The standard deviation of the relative shifts is of $4.3 \, \mathrm{cm \, s^{-1}}$, which is 34% above the photon noise of $3.2 \, \mathrm{cm \, s^{-1}}$. The excess noise is probably caused by some residual fibre modal noise through less than optimal spatial mode scrambling. When proceeding to measuring the relative stability of the two LFCs, we were first facing a significantly higher excess noise, resulting in a standard deviation of up to $13.3 \, \mathrm{cm \, s^{-1}}$ with a photon noise of $3.2 \, \mathrm{cm \, s^{-1}}$. We then installed a dynamic scrambler to shake the last fibre in the sequence for better suppression of modal noise. With this enhanced scrambling configuration we achieved nearly photon-noise limited results also in the relative measurements (Fig. 4 of the main article).

Later tests with LFC1 permanently installed on HARPS showed that the additional fibre shaker can also be beneficial for closely approaching the photon-noise limit with just one LFC in both channels. However, efforts are still ongoing to make the added shaker obsolete through an optimized coupling of spatial modes in the fibre sequence.
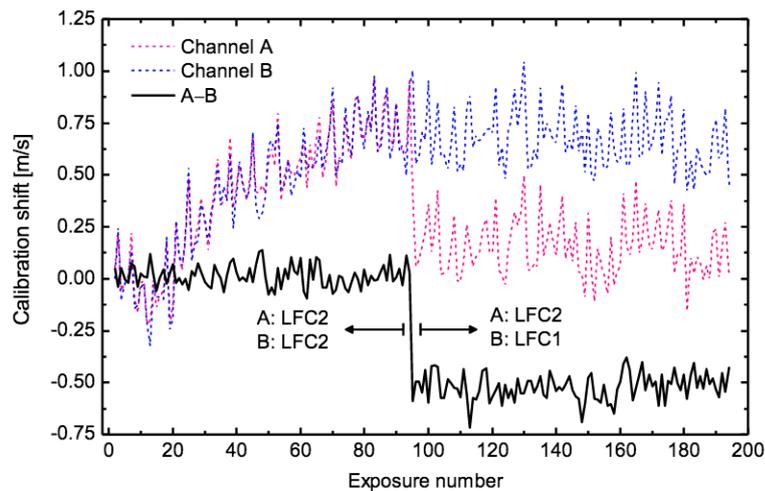
Figure 2.13: Calibration reproducibility with two different LFCs: mean shift in calibration polynomial. We supplement the analysis of calibration reproducibility in the main article by applying the standard astronomical calibration method of fitting polynomials to calibrate the spectrograph in absolute terms. We find an average shift of $53\,\mathrm{cm\,s^{-1}}$ between the calibrations performed by the two LFCs on the same dataset as in Figure 2.4. The slightly larger offset using this method is probably caused by altered sensitivity to systematic effects.

2.3.4 Absolute wavelength calibration in an astronomical context

Several science cases rely on the unprecedented wavelength accuracy and stability provided by LFCs. It is therefore important to quantify the agreement between the two absolute calibrators at our disposal. Our analysis in the main article reveals a $49\,\mathrm{cm\,s^{-1}}$ shift between the two LFCs when considering the average shift in the positions of comb lines. However, astronomical data analysis normally uses a more complex approach, which comprises calibrating scientific exposures by establishing an absolute wavelength scale. To reach the highest precision, this is usually complemented by monitoring line shifts on a second fibre channel. This tracks global shifts in the calibration caused by spectrograph drifts, which was the main focus in the previous parts of this Chapter. For absolute calibration of echelle spectra, the standard method entails finding a polynomial function relating the positions of a set of emission lines to their known wavelengths. This effectively assigns a wavelength value to each pixel in every echelle order. We emulate the standard astronomical procedure to assess the average agreement of the two LFCs in the important aspect of absolute calibration.

We determine the positions of the lines in the way described in the Section 2.2.3 and the wavelengths through the means of the LFC equation. The wavelength calibration is derived by fitting an eight order polynomial to this data. We compute the unweighted average shift in the wavelengths of all pixels relative to the reference exposure to quantify the overall shift in absolute wavelength calibration (Figure 2.13). This approach seamlessly handles the change to a different LFC

during a series of exposures. The results obtained are very similar to the results in the main article, with the change to another LFC causing a shift of $53\,\mathrm{cm\,s^{-1}}$ (as compared to $49\,\mathrm{cm\,s^{-1}}$ with the other method). The results are slightly noisier because the contributions of the lines are not weighted by their photon noise. The analysis is also more prone to the systematic influence of CTI, as it affects weak lines more heavily than strong lines (Bouchy et al. 2009; Zhao et al. 2014), which all receive the same weight. Differences in line intensities of the two LFCs may thus explain the $4\,\mathrm{cm\,s^{-1}}$ larger disagreement between the two LFCs as determined with this method.

2.3.5 All LFC vs. LFC exposures ever recorded with HARPS

Considering all campaigns that we have so far carried out with LFCs on HARPS, the available data now span a total of more than six years. Although made up from measurements with very different LFCs using different mode spacing and offset frequency, such a long time series is of great value for judging the long-term stability of HARPS and for identifying systematic influences. These aspects are essential when observing phenomena that evolve on extended time horizons such as long-period exoplanets.

Figure 2.14 shows all HARPS exposures ever recorded with LFC light in both channels. All exposures are related to a common reference of the year 2012. We use linear interpolation between the comb lines of the reference to relate LFCs with different structures (see Section 2.2.3). While A and B experience a drift by a total of about $35\,\mathrm{m\,s^{-1}}$, AB exhibits no clear continuous drift. Instead, AB seems to be dominated by effects from varying illumination conditions. This is most clearly seen in a section of the data that is shaded in grey in Figure 2.14. Here, one or both channels were attenuated by up to a factor of 100 through grey filters with different optical density. This causes shifts on the $\mathrm{m\,s^{-1}}$ scale, which can be explained through CTI of the CCD readout mechanism (Bouchy et al. 2009; Zhao et al. 2014) and through effects from the data extraction algorithm of the HARPS pipeline that depend on signal strength. The systematic errors evident from Figure 2.14 are of up to several $\mathrm{m\,s^{-1}}$. For measurements requiring $\mathrm{cm\,s^{-1}}$ precision, this highlights the importance of minimizing variations in signal strength and changes in fibre-coupling conditions. Future projects should, therefore, place particular emphasis on solutions for enhanced mode scrambling, modelling of CTI, and distortion-free data extraction techniques.
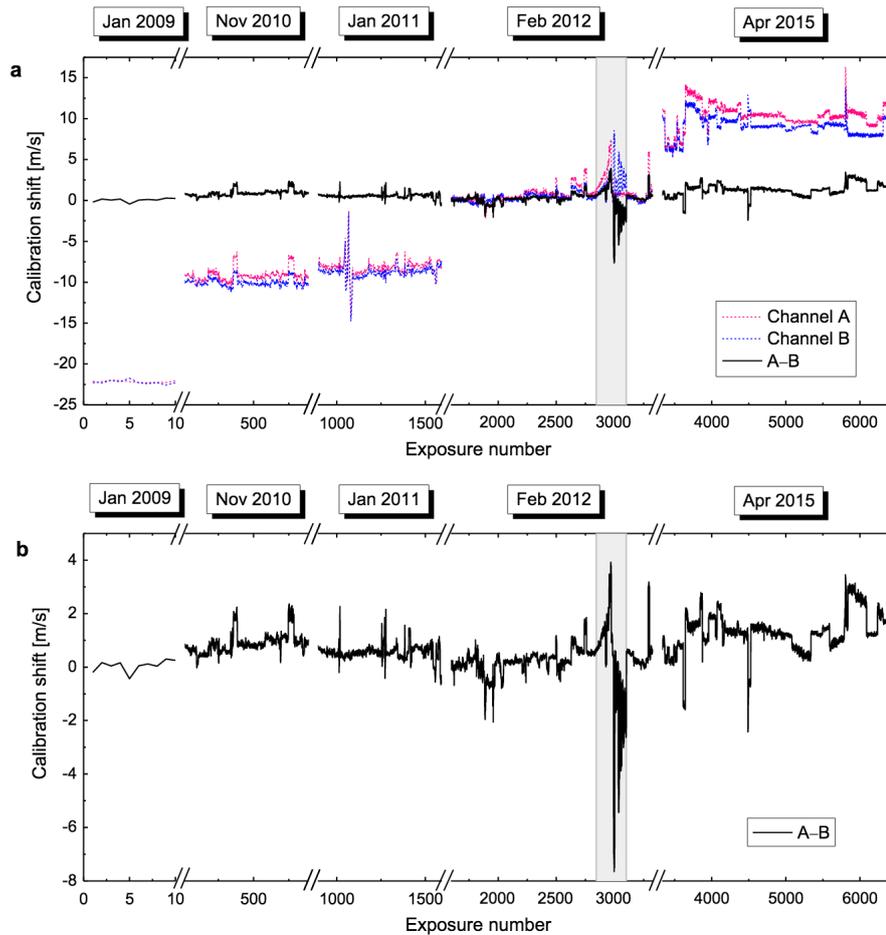
Figure 2.14: All HARPS exposures with LFCs on both channels recorded until April 2015. Panel (a): Evolution of the channels A, B and their difference (AB). Panel (b): AB only, on a magnified vertical scale. In the grey-shaded region, one or both channels are attenuated by neutral-density filters of varying optical density. The reference is an exposure recorded in the year 2012.

Chapter 3

Precision and consistency of astronomical laser frequency combs

3.1   Introduction

Measuring spectroscopic velocity shifts ($\Delta\lambda/\lambda$) in high resolution astronomical spectra is a powerful and widely used tool in a range of astronomical disciplines. It is used to detect planets outside of the Solar System (Mayor & Queloz 1995; Anglada-Escudé et al. 2016), look for variation in the values of fundamental constants (e.g. the fine structure constant, α, Dzuba et al. 1999b; Webb et al. 1999, 2011; King et al. 2012, and others). It is also the method proposed to measure the expansion of the Universe in "real time" and in a model-independent way (also known as the Sandage test or the "redshift drift" measurement, Sandage (1962); Loeb (1998); Liske et al. (2008)) and to map the gravitational potential of the Galaxy (Ravi et al. 2019a; Leão et al. 2019; Silverwood & Easther 2019). These science goals are important drivers for all 30-meter class telescopes planned for the 2020s, i.e. the Extremely Large Telescope (ELT, Tamai et al. 2018), the Thirty Meter Telescope (TMT, Simard et al. 2016), and the Giant Magellan Telescope (GMT, Fanson et al. 2018). This is why all will be equipped with a high resolution optical spectrograph.

The success of these projects relies not only on increasing the light gathering capability of telescopes, but also on the instrument stability, and the precision and accuracy with which velocity shifts in astronomical spectra can be measured. In this context, precision is the repeatability of subsequent wavelength measurements to each other and accuracy is the closeness of a measured wavelength to its true

value. The most demanding of the four aforementioned projects – the redshift drift measurement – requires velocity shifts be measured with a precision better than 3 part in 100 billion over spectral ranges of several hundred nanometers and a period longer than ten years. This corresponds to $\Delta\lambda/\lambda = 3 \times 10^{-11}$, or equivalently, 1 centimeter per second ($\mathrm{cm\,s^{-1}}$). The best way to ensure long term precision is to ensure that the measured wavelengths are also accurate. Instruments intended to perform these experiments are therefore designed to have wavelength calibration precision around $1\,\mathrm{cm\,s^{-1}}$ and accuracy of order 1 metre per second ($\mathrm{m\,s^{-1}}$) over a period of a decade (e.g. Liske 2014; Marconi et al. 2016). The success of these projects thus critically relies on the precision and accuracy of the wavelength calibration reaching these levels.

The currently most commonly used method of wavelength calibration in high resolution spectrographs uses hollow cathode lamp, most commonly Thorium (Th) and Uranium (U) lamps. The hollow cathode lamp calibration suffers from several major drawbacks limiting its precision to a few tens of $\mathrm{cm\,s^{-1}}$ over a period of one year and average accuracy of approximately $1\,\mathrm{m\,s^{-1}}$ in the same period (Lovis et al. 2006), thus falling short of the previously stated goals. It is now generally accepted that laser frequency comb systems (LFC, Udem et al. 2002; Hänsch 2006) can achieve substantially better results to push the precision and accuracy of wavelength calibration to the $1\,\mathrm{cm\,s^{-1}}$ level and ensure the feasibility of the scientific projects mentioned above. For a recent review about laser frequency comb spectroscopy, see Picqué & Hänsch (2019).

Laser frequency combs offer significant advantages over arc lamps: (i) they produce thousands of unblended emission lines of uniform intensity and equidistant in frequency; and (ii) the frequencies of LFC lines are *a priori* known with accuracy of the atomic clock to which the system is coupled (typically $\Delta f/f \approx 10^{-11}$ or $3\,\mathrm{mm\,s^{-1}}$). Murphy et al. (2007) discusses the advantages of LFCs in more detail. LFCs do not suffer from lamp aging in the same way as hollow cathode lamps, but their components (e.g. the photonic crystal fibre, PCF) degrade.

The European Southern Observatory (ESO) formed a consortium to develop and install an astronomical LFC (also known as an "astrocomb") on the HARPS instrument in 2008. The development saw several test campaigns after which the LFC was permanently installed on HARPS in May 2015. This LFC has already demonstrated short-term precision at the photon noise level, $\Delta\lambda/\lambda = 6 \times 10^{-9}$ or $2.5\,\mathrm{cm\,s^{-1}}$ (Wilken et al. 2012). Similar short-term precision using LFCs has also been demonstrated on several other high-resolution spectrographs (e.g. Ycas et al. 2012; Phillips et al. 2012; Doerr et al. 2012; Glenday et al. 2015; Brucalassi et al. 2016; McCracken et al. 2017b). All these previous studies referenced the LFC to itself, so any possible systematic effects arising in the LFC system itself may go undetected.

Definitive proof of LFC performance can only come from its comparison to another calibration source of the same (or higher) precision. This is why, in a campaign that took place in April 2015, two independent LFC systems were installed on HARPS: one constructed for the HARPS instrument itself and the other constructed for the FOCES instrument (Pfeiffer et al. 1992). The experiment had two goals. The first one was to understand whether the precision of a single LFC

is confirmed by an independent system, and the second one was to assess zero-point offsets in the wavelength calibration introduced by switching between the two LFCs.

Chapter 2 analyses the same dataset, but with a focus on describing the experimental setup, LFC hardware and its optimisation during the campaign before the LFC's deployment in May 2015. This Chapter, on the other hand, focuses on data analysis techniques and advanced algorithms that will provide wavelength calibration precision required by the ELT projects discussed above.

The Chapter is divided as follows: in Section 3.2 we briefly describe the experimental setup – the spectrograph and the two LFCs. Section 3.3 describes the dataset and general properties of the spectra. Section 3.4 gives details on our automatic algorithm to detect LFC lines and obtaining their centres and wavelengths. Wavelength calibration is discussed in Section 3.5, where we describe how we deal with instrumental effects that impact significantly on wavelength calibration precision and accuracy, e.g. defects associated with the CCD manufacturing process. In the same section, we describe our findings on the optimal wavelength calibration model for HARPS spectra. We present our findings on precision and consistency in Section 3.6, where we also consider and model the contribution of flux dependent velocity shifts. Finally, our results are presented in Section 3.7 and discussed in Section 3.8.

3.2   Experimental setup

3.2.1   The HARPS instrument

The High Accuracy Radial velocity Planet Searcher (HARPS, Mayor et al. 2003) was built for extreme stability and precision and is one of the most stable astronomical spectrographs in existence. HARPS is a fibre fed, high-resolution ($R = \lambda/\Delta\lambda = 115000$), R4 grism cross-dispersed echelle spectrograph installed on the 3.6m telescope at ESO's La Silla Observatory and the first instrument to be equipped with an LFC for regular operations (Probst et al. 2016). The light of the two input fibres is dispersed into 72 echelle orders on the detector, simultaneously covering the wavelength range between 378 and 691 nm. The spectrograph is enclosed in a thermal and pressure controlled vacuum vessel, with long-term temperature variations at the 0.01K level and operating pressure below 1×10^{-3} mbar.

Several thermal and mechanical effects can slightly shift the positions of the spectrum on the detector with time, an unavoidable effect that is eliminated through simultaneous referencing: drifts in the science fibre are tracked by simultaneously observing a spectrum rich in velocity information content (e.g. ThAr, an LFC, or a Fabry-Pérot etalon) in the secondary fibre (Baranne et al. 1996). Each fibre has a static double scrambler. A servo controller ("secondary guiding") ensures that the object image is always centered in the object fibre. In order to ensure light entrance stability and proper mode mixing in the fibres, a dynamical fibre scrambler that shakes the fibres was added to the setup, adding a temporal scrambling of light (Probst et al. 2020).

The detector

The HARPS detector is a mosaic of two EEV2k4 CCDs (red and blue). Each CCD is constructed by stitching together $1024 \times 512\,\mathrm{pix}^2$ segments: eight in the dispersion (x) direction and two in the cross-dispersion (y) direction. CCD pixels have a nominal size of $15 \times 15\mu\mathrm{m}^2$ but errors in the segment stitching process results in deviations from the nominal pixel size at segment boundaries, i.e. every 512 pixels in the x direction and every 1024 in the y direction.

The HARPS detector was the first one shown to suffer from the effect of imperfect CCD stitching (Wilken et al. 2010). Dumusque et al. (2015b) showed that lines which cross segment boundaries produce spurious velocity shifts as high as a few m s^{-1} with a period of one year in HARPS observations. Proper mitigation of this effect is therefore important for exoplanet detection and also for fundamental constant and redshift drift measurements. The way this has been done previously is to shift the measured positions of calibration lines (in pixel space) by the measured size of pixel size anomalies during wavelength calibration. Pixel size anomalies were measured by Bauer et al. (2015) using Fabry-Pérot etalon exposures and by Coffinet et al. (2019) using flat-field exposures. We perform a similar measurement using LFC exposures in Section 3.5.1. Therefore, in Section 3.5.2, we examine the effectiveness of different calibration methods in removing the effect of pixel size anomalies.

3.2.2 The laser frequency combs

Astrocombs are laser frequency comb systems built specifically to wavelength calibrate astronomical spectrographs (Steinmetz et al. 2008). They produce thousands of emission lines (or modes) of uniform intensity with precisely known frequencies. The nominal frequency of each mode is given by the "LFC equation":

$$f_n = f_o + n \times f_r, \tag{3.1}$$

where f_o and f_r are the "offset" and the "repetition" frequencies, and n is the mode number (a large positive integer). Both f_o and f_r are radio-frequencies referenced to an atomic clock and known with precision of $\Delta f/f = 5.6 \times 10^{-12}$ over the timescale of several hours (Probst et al. 2020). The frequency, and the wavelength, of each line can therefore be determined with the same precision.

The HARPS LFC development saw several test campaigns between January 2009 and April 2015. One of the goals of the April 2015 campaign was to characterise the performance of the HARPS LFC against a completely independent second one. The second LFC, built for the FOCES instrument (Pfeiffer et al. 1992), was loaned from the Wendelstein Observatory (operated by Ludwig-Maximillians-Universität, LMU) for this purpose. The HARPS LFC has an 18 GHz line separation. Since the HARPS spectral resolution is around 5 GHz in the middle of its spectral range, LFC lines are kept well apart with virtually no residual overlap. We will refer to this 18 GHz LFC as "LFC1" in further text. To accommodate the lower resolution of the FOCES instrument ($R = 70000$), the FOCES LFC was designed with a wider line separation of 25 GHz. We will refer to this LFC as

Fibre A	LFC2	LFC2
Fibre B	LFC2	LFC1
Exposure		
1	94	194
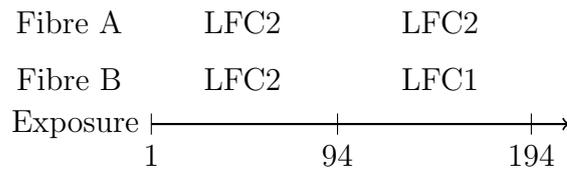
Figure 3.1: We analyze a time series of 194 LFC exposures. Fibre A carrying LFC2 light was used for simultaneous referencing throughout the series. Fibre B carried LFC2 light during the first 94 exposures, after which it was switched to carry LFC1 light. We measure velocity shifts of all exposures, compensating for unavoidable spectrograph drifts, in order to establish the precision of each LFC and the consistency of their wavelength calibrations.

"LFC2" in further text. Wavelength coverage of the two LFCs differs slightly due to different requirements for the HARPS and FOCES instruments. Relevant information about the two LFCs are tabulated in Table 3.1. See Probst et al. (2020) for a more comprehensive description of the LFC design and the setup during the April 2015 campaign.

3.3   Data

3.3.1   The dataset

We use a time series of spectra of the two LFCs described above for our analysis. The series consists of a total of 194 exposures. Each exposure was 30 seconds long with a read-out time of 22 seconds. The entire duration of the series is approximately six hours, with a two hour gap between the end of exposure 94 and the beginning of exposure 95. In the first 94 exposures, fibre B (the object fibre) was illuminated with LFC2, after which the LFC was changed and 100 exposures of LFC1 were taken in the same fibre. We therefore divide the dataset into two samples, depending on which LFC illuminated fibre B. We will refer to the "LFC2 sample" for exposures 1–94 and to the "LFC1 sample" for exposures 95–194. Meanwhile, LFC2 was used as a simultaneous reference in fibre A (simultaneous fibre), keeping track of spectrograph drifts throughout the whole series, for a total of 194 exposures in fibre A. Figure 3.1 gives a schematic of the series of exposures we use in our analysis.

We choose to work on 'E2DS' files – unmerged 1D spectra extracted from raw images by the HARPS pipeline (version v3.8) using optimal extraction after Horne (1986). Each E2DS file consists of 72 (71) echelle orders covering 4096 pixels of an exposure in fibre A (B). The detector covers echelle orders 89 to 161, with the exception of order 115 for fibre A and orders 115 and 116 for fibre B, which fall in between the two detector CCDs. A fraction of the LFC1 spectrum showing individual lines around 561 nm in echelle order 109 is plotted in Figure 3.2.

We limit our analysis to echelle orders 89–130, where the fluxes of the two LFCs are sufficiently high and comparable – as evidenced by the total number of counts detected in each echelle order (see Figure 3.3). This covers wavelengths between

	LFC1	LFC2
f_o	5.7 GHz	9.27 GHz
Native f_r	250 MHz	250 MHz
Mode filtering	72	100
f_r	18 GHz	25 GHz
λ_{min}	438.8 nm	455.4 nm
λ_{max}	691.5 nm	691.5 nm

Table 3.1: Basic parameters of LFC1 and LFC2. The two share the basic design but have been optimised for different instruments. See Probst et al. (2020) for a more comprehensive description.

468.1 nm and 691.5 nm, or 70% of the total HARPS wavelength range.

A quantity that is directly relevant for exoplanet detection studies, but not for varying constant or redshift drift measurements, is the uncertainty on the mean velocity shift that can be determined in a spectrum. This measure of uncertainty is determined by the photon noise and other detailed spectral attributes – see Bouchy et al. (2001) and Murphy et al. (2007) for details. We will refer to this quantity as "photon-limited velocity precision". We calculate the photon-limited velocity precision of individual exposures in our dataset across orders $89 - 130$. We find that the average photon-limited velocity precision of a single exposure in the LFC1 sample is $2.3 \, \mathrm{cm \, s^{-1}}$ ($3.0 \, \mathrm{cm \, s^{-1}}$) in fibre A (B); and in the LFC2 sample is $2.3 \, \mathrm{cm \, s^{-1}}$ ($2.5 \, \mathrm{cm \, s^{-1}}$) in fibre A (B). The reduced photon-limited velocity precision in fibre B is attributable to lower flux compared to fibre A for both samples.

3.3.2 Spectral background

The spectra obtained using both LFCs exhibit a strong background component (the red line in Figure 3.2 illustrates the background light in the LFC1 spectrum). This background originates mostly from amplified spontaneous emission (ASE) in the high-power Yb-fibre amplifier (Probst et al. 2013; Probst et al. 2020) and, to a lesser extent, from scattered light (measured to be $< 1\%$ at 590 nm in the HARPS spectrograph Rodler & Lo Curto 2019). The ASE from the high-power amplifier is further amplified by non-linear processes in the photonic crystal fibre, located just after it in the optical path. We refer the reader to Probst et al. (2020) for more technical details on the origin of the background.

The background is highly modulated, approximately tracing the variations in intensity of LFC lines. We characterize the contribution of the background to the total detected flux in terms of the background-to-envelope ratio (B2E). The background here refers to a function connecting the local minima in the spectrum. This is a piece-wise linear function of pixel number, x:

$$B(x)\Big|_{x_1}^{x_2} = F(x_1) + \frac{F(x_2) - F(x_1)}{x_2 - x_1} \times (x_2 - x) \tag{3.2}$$

where F is flux (in units counts) and x_1 and x_2 are locations of adjacent minima
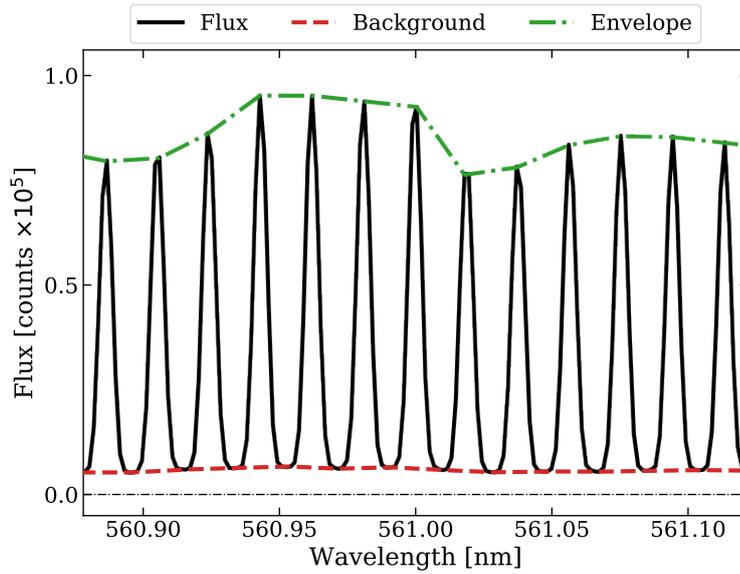
Figure 3.2: Zoom-in of the region between 560.8 and 561.1 nm of the 1D extracted spectrum of LFC1 in echelle order 109 (solid black line). Signal amplification gives rise to the background (dashed red line), which contributes approximately 13% of the total flux in a single exposure, as calculated from its ratio to the envelope (dot-dashed green line). The background is removed before line fitting.
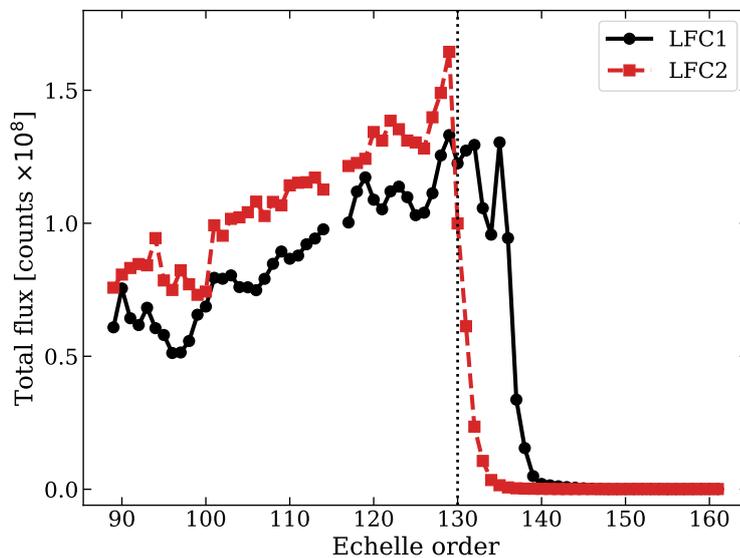
Figure 3.3: The light intensity of the two LFCs in fibre B drops sharply above order 136 for LFC1 (black circles) and above order 130 for LFC2 (red squares). The two LFCs cover more than 70% of the total HARPS wavelength range. We limit our analysis to orders 89–130, where the two LFCs have sufficient and comparable flux. Orders 115 and 116 fall in between the two detector CCDs and therefore have no measured flux.

in the spectrum. The envelope is the analogous quantity except maxima instead of minima are fitted. Both are illustrated in Figure 3.2.

Our analysis shows that B2E increases linearly with decreasing wavelength, with values between 2% and 16% in a typical exposure in both LFCs. The background is therefore non-negligible. To minimise any impact on estimating LFC line centres, we subtract the background from the total flux, propagating the errors to correctly modify the spectral variance array:

$$\sigma^2(x) = F(x) + B(x), \tag{3.3}$$

where we have assumed that the flux and the background are Poissonian and the detector noise (i.e. dark-current, read-out-noise) is negligible.

3.4   Laser frequency comb lines

3.4.1   Line detection

Our line detection algorithm automatically detects all LFC lines in the exposure. The algorithm relies on locating minima between individual lines as their natural limits. We found this practice to be preferred over using maxima, as the latter caused issues with falsely detected and skipped lines when using an automated detection routine. Line detection is done in three steps. In the first step, we smooth the recorded spectrum of an echelle order with a Wiener filter with a 3 (5) pixel wide window for LFC1 (LFC2). The smoothed spectrum makes identifying minima easier in the following step. In the second step, we identify local minima as points where the first derivative (with respect to the pixel number) switches sign, and the second derivative is larger than zero. This step sometimes falsely detects minima in the data, especially when the signal-to-noise ratio (S/N) is low. We therefore reject falsely detected minima in the third and final step, using the following two assumptions: (i) the distance between minima must not significantly deviate from typical distance between LFC lines; and (ii) the distance between minima within the echelle order increases approximately linearly with increasing wavelength.

We use the first assumption to remove minima closer together than 90% of the typical distance between lines in the same echelle order, where the latter is equal to period of the strongest peak in the periodogram of the order. Depending on the LFC and the echelle order, this number is between 11 and 20 pixels, with LFC2 always having larger values due to larger mode separation. The second assumption follows directly from Equation (3.1): the separation of consecutive lines in wavelength space approximately follows $\sim \lambda_n/n$, where λ_n is the wavelength of the nth LFC mode. This means that the distance between lines increases approximately linearly in pixel space within a single echelle order. We therefore remove $> 3\sigma$ outliers to the linear function best describing $\Delta x(x)$, where Δx is the distance between adjacent lines.

Our automatic line detection algorithm detects ≈ 13300 and ≈ 9800 in each LFC1 and LFC2 exposure (where the background was subtracted), respectively.

Across the entire dataset, we detect $N_A = 1898254$ lines in fibre A and $N_B = 2222168$ lines in fibre B, with a total $N = 4120422$ lines. The average signal-to-noise ratio (S/N) of lines is approximately 165, with $\approx 10\%$ differences between LFC1 versus LFC2 and fibre A versus fibre B.

3.4.2 Profile fitting

In this analysis we assume that each LFC line can be well-represented by a single Gaussian profile. Visual examination of the data suggests that the approximation is generally reasonable although asymmetries are seen depending on position on the CCD. We will explore different ways of modelling LFC profiles in forthcoming work.

We use a Levenberg-Marquardt algorithm[1] to derive best-fit Gaussian parameters for each line. Since the data are high signal-to-noise, and since slight non-linearities in the model function exist across individual spectral pixels, the model-fitting procedure avoids simply computing the Gaussian value at the centre of each pixel but instead performs an integration of the flux falling within each pixel. The expected number of counts in each pixel (with x_L and x_R its left and right boundaries) is given by:

$$\Phi(x_L, x_R) = A\,\sigma\,\sqrt{\frac{\pi}{2}}\left[\mathrm{erf}\left(\frac{x_R - \mu}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{x_L - \mu}{\sqrt{2}\sigma}\right)\right]. \qquad (3.4)$$

Here erf is the error function, and A, μ, and σ are the amplitude, the mean and the standard deviation of the Gaussian, respectively. The algorithm provides a line-centre uncertainty estimate for each line. The mean line-centre uncertainty across all ≈ 4M detected lines is 3 milli-pixel (mpix). An example fit for a single LFC1 line is shown in Figure 3.4.

Ultimately, the Gaussian approximation above is incorrect. This can be seen by eye from the asymmetric shape of LFC lines and is evident from the high values of reduced χ^2 values ($\chi^2_\nu = \chi^2/\nu$, with ν the number of degrees of freedom) we get for the Gaussian line fits. The χ^2_ν distribution derived from Gaussian fitting, shown in Figure 3.5, peaks at $\chi^2_\nu = 7.9$ and has a mean 57.6 across the detector, indicating an overall poor fit to the data. Large χ^2_ν values for Gaussian fits are concentrated in two regions: in the red half ($x > 2048$) of echelle orders 89–98 and the middle part of orders 125–130 ($x \approx 2048$). The χ^2_ν values do not correlate with any of the fit parameters or their errors. The same pattern is seen independently for both LFCs and in both fibres. This suggests that the χ^2_ν pattern must be due to variation in the line-spread function profile across the detector.

Whilst a Gaussian is clearly not the correct line shape, we show later that the Gaussian approximation nevertheless performs well in term of repeatability such that radial velocity studies are relatively unaffected by the correlated patterns seen in Figure 3.5. We note however that this issue will be important for other types of studies such as varying constants and redshift drift. We will later derive a model of the line-spread function – assuming it will also give us a more accurate estimation of the line centre.

---

[1] `leastsq` routine from `scipy.optimize`
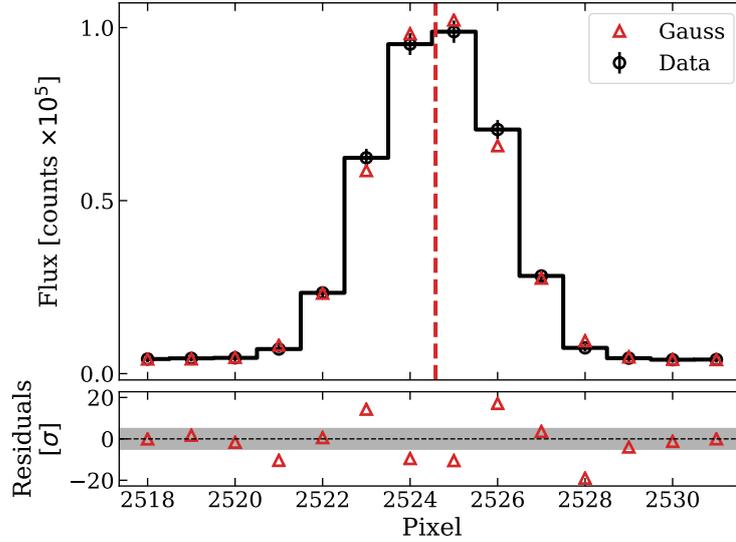
Figure 3.4: We fit Gaussian profiles to LFC lines after spectral background removal. The Gaussian function is integrated under each pixel during fitting. A smooth Gaussian curve is therefore not plotted. *Top*: The data (black circles) and the Gaussian model (red triangles). The best line centre estimate is given by the vertical red dashed line. Errors on the data are enlarged by a factor of 10 for visibility. *Bottom*: Normalised residuals show the Gaussian model is ultimately not the correct line profile. The shaded gray area shows the 5σ range.

3.4.3   Mode identification

LFC wavelength calibration requires another wavelength calibration source to establish the absolute scale by identifying a single line in each echelle order, after which the frequencies of all lines are known by counting. This is not a critical aspect as long as the error in the absolute calibration source is much smaller than half of the separation between LFC lines. The local accuracy of ThAr wavelength calibration, the HARPS standard, is between 10 and $80 \, \mathrm{m \, s^{-1}}$ or $\Delta f = 0.02$ to $0.15 \, \mathrm{GHz}$ at 550 nm, (respectively). This is well below the line separation of either LFC (Table 3.1). We therefore use the ThAr wavelength calibration to identify a single LFC mode per echelle order in the following way. The mode number, n, is the nearest integer to the number:

$$n = \mathrm{nint}\left(\frac{f_n^{ThAr} - f_o}{f_r}\right), \tag{3.5}$$

where f_n^{ThAr} is frequency of the LFC line determined from the ThAr wavelength calibration, f_o and f_r are the LFC offset and repetition frequencies, respectively. These frequencies are recorded by the LFC system[2]. We always use the same ThAr

---

[2]The experimental nature of the April 2015 campaign meant f_o of both LFCs was changed multiple times. We discover an un-noted shift in LFC1 of 100 MHz whilst analysing the daset. More details can be found in Appendix A.
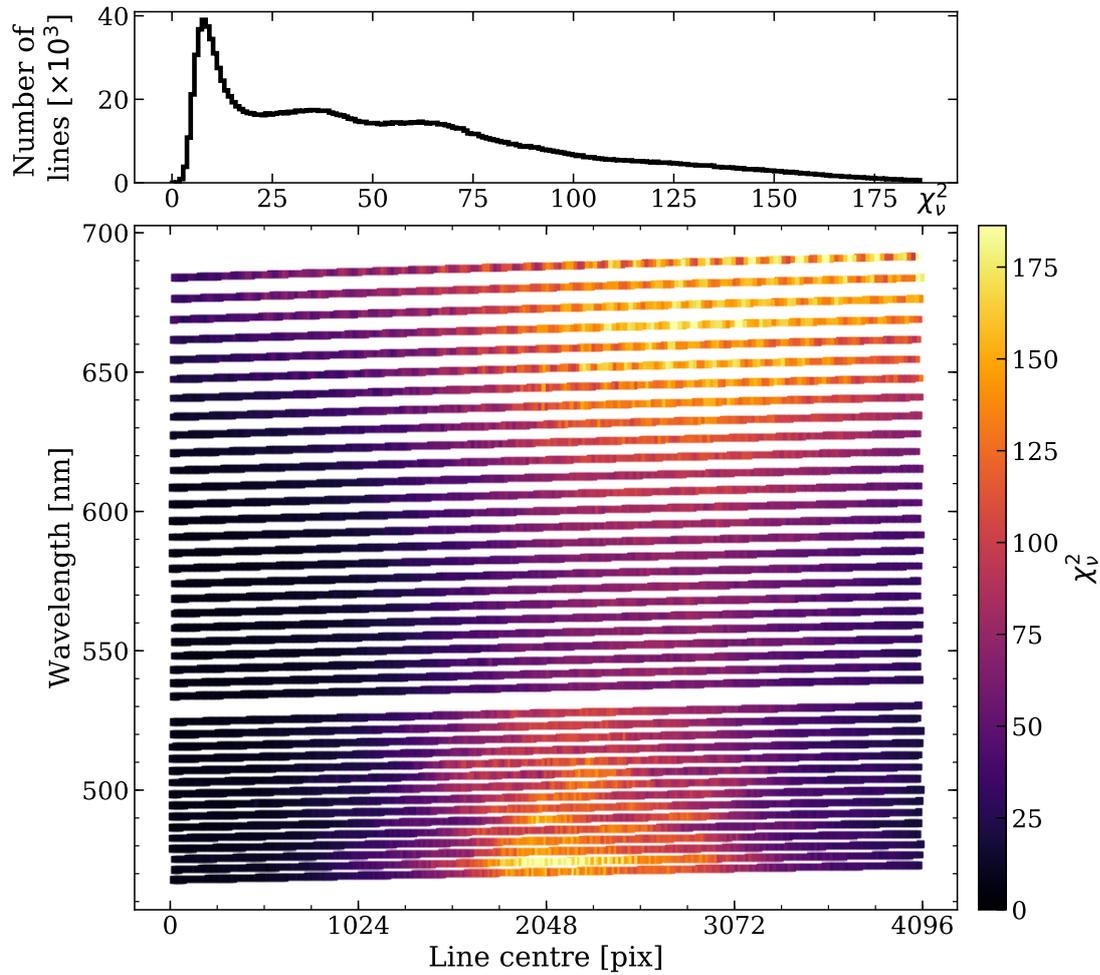
Figure 3.5: *Main panel*: The distribution of χ_ν^2 values for the Gaussian profile fits shows that a Gaussian profile does not provide a satisfactory fit to the data. There exists a concentration of high χ_ν^2 values in two places on the detector, showing that the LFC line profile changes systematically across the detector. *Top panel*: The histogram of values in the main panel, total number of lines $N = 1898254$. The distribution peaks at $\chi_\nu^2 = 7.9$.

coefficients to determine the mode number of the line closest to pixel 2048, where the ThAr calibration is expected to be the most accurate.

The wavelength and the corresponding uncertainty of each line are calculated from equations:

$$\lambda_n = \frac{c}{f_n}, \tag{3.6}$$

and

$$\sigma_\lambda = \frac{c}{f_n^2}\, \sigma_f, \tag{3.7}$$

with c the speed of light and σ_f is the frequency uncertainty for each line. Empirically, $\sigma_f/f = \sigma_\lambda/\lambda \approx 10^{-11}$. Uncertainties at this level are generally orders of magnitude below spectral line uncertainties in astronomical targets such as

quasars.

3.5   Wavelength calibration

Spectrograph wavelength calibration relates the measured positions of a set of calibration lines on the detector with their known laboratory wavelengths in a way that assigns a wavelength to each position on the detector. The most common approach in optical echelle spectroscopy is to fit a polynomial to a set of calibration lines in each extracted echelle order. The large number of LFC lines and the exquisite accuracy with which their wavelengths are known allow us to look for a more realistic model, e.g. by increasing the polynomial order (see e.g. Wilken et al. 2010). In this Section we examine a range of polynomial orders in an attempt to identify an optimal number of degrees of freedom.

Furthermore, it is generally assumed that all pixels have the same physical size, so that the physical distance between calibration lines on the detector can be expressed in pixel distance. This assumption has been proven invalid by the discovery of the HARPS detector pixel size anomalies (see Section 3.2.1). Distortions of the HARPS wavelength scale caused by the pixel size anomalies can be removed using one of the following two approaches: (i) *global polynomial*: producing a wavelength calibration spanning the entire echelle order in which the anomalies have been accounted for (see Section 3.5.1), and (ii) *segmented polynomials*: producing a separate wavelength calibration for each 512-pixel CCD segment that an echelle order crosses. Coffinet et al. (2019) take the former approach, whereas Wilken et al. (2010) and Molaro et al. (2013b) take the latter. In this analysis we use the Akaike information criterion (AIC, Akaike 1974), corrected for the finite sample size (AICc, Sugiura 1978), in order to asses which of these two approaches provides the best results (Section 3.5.2).

In what follows, we make extensive use of Weighted Orthogonal Distance Regression[3], an algorithm which allows us to account for both the positional and wavelength uncertainty of LFC lines in polynomial fitting.

3.5.1   Measuring HARPS detector pixel size anomalies

We derive our pixel size anomaly measurements from distortions of the HARPS wavelength scale revealed by the LFC lines. LFC wavelength coverage limits us to only three out of four y-blocks: two blocks of the red CCD (blocks 1 and 2 in further text) and a single block of the blue CCD (block 3 in further text).

Pixel size anomalies are measured for each block individually using the wavelength calibration residuals in the following way:

1. For a given detector y-block, we consider only those echelle orders which fall onto the block. It seems likely that distortions in the x direction are common to all orders that fall within the same 1024-pixel high CCD block in the y direction. If this is not the case, the effect would be to increase the

---

[3]Python package `scipy.odr`, based on `ODRPACK` (Boggs et al. 1992)

scatter in the y-direction within each 512×1024 block (the shaded region) in Figure 3.6;

2. We fit a global eighth order polynomial to pixel-wavelength pairs of LFC lines separately in each echelle order and each exposure. We calculate the residuals to the true line wavelengths (i.e. Equation (3.1)) and express them in $\mathrm{m\,s^{-1}}$;

3. We bin the residuals in 64 bins along the x axis (i.e. giving 8 points per 512 pixel-wide segment), excluding lines closer to segment borders than 10 pixels and those with residuals larger than $200\,\mathrm{m\,s^{-1}}$;

4. We fit a third order polynomial to the binned residuals. A typical error on each binned residual is of order $5\,\mathrm{cm\,s^{-1}}$ due to the large number of points in each bin ($\approx 10 - 20$k).

5. The discontinuity g is given by the difference between the polynomials in two adjacent segments at their boundary (in units $\mathrm{m\,s^{-1}}$):

$$g(k) = P_1(x)\Big|_{x=k} - P_2(x)\Big|_{x=k}. \tag{3.8}$$

Here, P_1 and P_2 are the polynomials in two adjacent segments and k is the position of the discontinuity (in pixels, multiple of 512);

6. We convert the discontinuity into the pixel size anomaly by dividing it by the size of the HARPS pixel in velocity units in the middle of the HARPS wavelength range: $1\mathrm{pix} = 829\,\mathrm{m\,s^{-1}}$.

This is illustrated in Figure 3.6 using data from fibre A. Measurements from fibre B (not illustrated) were found to be consistent with those from fibre A; there are a total of 21 pixel size anomaly measurements from each fibre. Differences between corresponding pixel size anomaly pairs were measured for all 21 pairs. The unweighted mean of those numbers is $\mu = -0.4 \pm 0.4\,\mathrm{mpix}$, corresponding to $-0.3 \pm 0.3\,\mathrm{m\,s^{-1}}$.

Although the two fibres produced completely consistent results, the results from fibre A seemed better than from fibre B, in that the scatter in the fibre B residuals (Figure 3.6) were more pronounced. We did not attempt to explore the reason for this and simply used fibre A to make the corrections. These are tabulated in Table 3.2 and used to adjust the positions of individual LFC lines. The corrected position of one line within one echelle order, in pixels, is given by:

$$x_c = x + \sum_{k}^{k<x} g_k, \tag{3.9}$$

where x_c and x are the corrected and the fitted line positions of the same line and g_k is the size of a pixel size anomaly located at the kth pixel.

Our measurements are in very good agreement with previous, independent, results. The average agreement with the results of Bauer et al. (2015) is $\mu =$
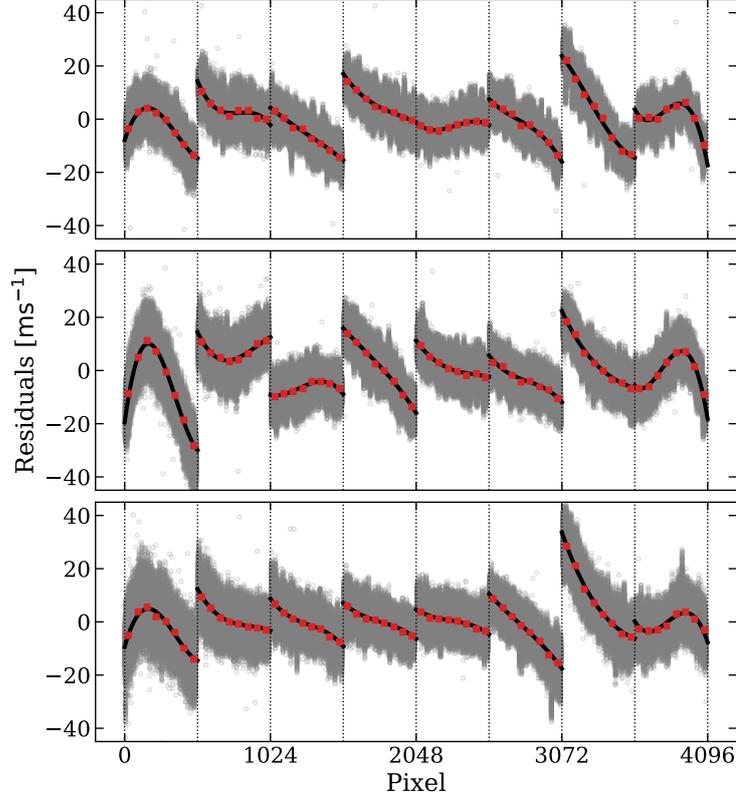
Figure 3.6: We use $\approx$ 2M LFC lines detected in fibre A to measure pixel size anomalies imparted during the CCD manufacturing process. The residuals to the global 8th order polynomial wavelength solution (gray points) reveal $\approx 20\,\text{m s}^{-1}$ discontinuities the edges of each 512 pixel-wide segment. We bin the residuals into 64 bins and calculate the mean in each (red squares). We fit the means in each segment with a 3rd order polynomial and use them to determine the size of the discontinuity. This is measured as the difference between the models (black lines) evaluated at segment borders (dotted vertical lines). The top two panels correspond to blocks 1 and 2 (comprising the red CCD) and the bottom panel corresponds to block 3 (comprising the blue CCD). Block 4 is not illuminated. The average error on each point is $\approx 5\,\text{cm s}^{-1}$ and therefore not visible.

3.1 ± 1.7 mpix ($2.6 \pm 1.6\,\text{m s}^{-1}$). Similarly, agreement with the results of Coffinet et al. (2019) is $\mu = 2.6 \pm 0.5$ mpix ($2.1 \pm 0.4\,\text{m s}^{-1}$). This corresponds an agreement between the two measurements at the 40 nm level on the detector. Unlike the flat-field method of Coffinet et al. (2019), we are not sensitive to the sizes of individual pixels at segment borders but only to the sum of their sizes. Our LFC method is, however, complementary to theirs and serves as a consistency check. Furthermore, it demonstrates the usefulness of LFCs for detector characterization necessary to obtain robust scientific results.

Table 3.2: The HARPS detector pixel size anomalies are calculated from LFC line wavelength residuals obtained using an eighth order global polynomial (see Section 3.5.1 for details). Units are milli-pixels (1 pixel = 15μm).

CCD	Red		Blue	
Orders	89–99	100–114	116–134	135–161
Block	1	2	3	4
N_{lines}	423k	698k	754k	0
Pixel	Pixel size anomaly (mpix)			
512	34.97	53.62	32.27	–
1024	7.67	-23.79	13.79	–
1536	39.12	30.88	19.72	–
2048	0.14	30.56	13.06	–
2560	11.77	9.82	18.40	–
3072	48.28	41.86	61.89	–
3584	21.78	3.18	7.07	–

3.5.2 Choosing a wavelength calibration model

We return to our aim of determining the model the optimal residuals number of degrees of freedom. We do this using the AIC corrected for small sample sizes (AICc, Sugiura 1978) and choose the model providing the smallest residuals using the smallest number of free parameters possible. The AICc is calculated as:

$$\text{AICc} = \chi^2 + 2\,p + \frac{2p\,(p+1)}{n-p-1}, \tag{3.10}$$

where p is the number of free parameters and n is the number of data points used in the fit. Theoretically, the model with the lowest AICc value is preferred.

We consider a total of 29 wavelength calibration models, grouped into two groups mentioned beforehand: the segmented and the global polynomial models. The segmented polynomial models range between 2nd and 12th order, whereas the global polynomials range between 3nd and 20th order. The former have $p = 8 \times (m+1)$ free parameters and the latter have $p = m+1$ free parameters, with m the highest order polynomial in the model.

For segmented polynomial models, we divide the echelle order into eight 512-pixel wide segments and fit a polynomial in each segment individually. We do not impose conditions on the continuity or smoothness of the polynomials at segment boundaries and leave the parameters in each segment independent, resulting in a discontinuous wavelength calibration model. In the case of global polynomial models, we first adjust the positions of individual lines to account for the CCD stitching pattern using Equation (3.9), after which we fit a single polynomial to all LFC lines in the echelle order.

We calculate AICc values for all 29 models, for each of the 43 echelle orders, for each of the 194 exposures, and for each fibre. There are therefore 66736 AICc values
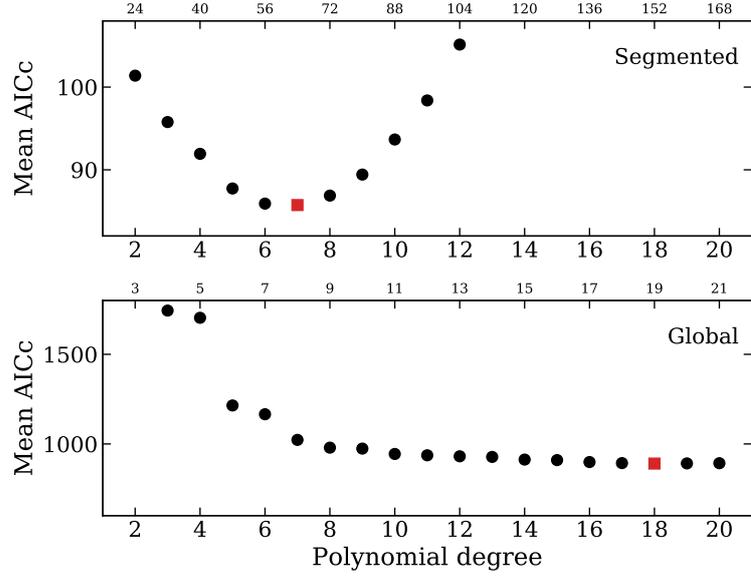
Figure 3.7: To find the optimal wavelength calibration model, we compute the mean AICc for all 29 models (black circles) divided into two groups: segmented (top panel) and global (bottom panel) polynomials. The preferred model (red square) has the lowest mean AICc value. These are the 7th order segmented and the 18th order global polynomial models. The top x-axis indicates the total number of free parameters for each model.

for each segmented polynomial model (i.e. for each parabolic, cubic, quadratic, quintic, etc.), and 8342 values for each global polynomial model, for each fibre. Considering the large number of individual fits, we look at AICc in a statistical sense when comparing models: the best model is the one with the lowest mean AICc, where that mean is averaged over all echelle orders and all exposures. The mean AICc for all 29 models is plotted in Figure 3.7, separately for the segmented (top panel) and global polynomials (bottom panel).

The segmented polynomial model with the lowest mean AICc is a seventh order polynomial ($p = 56$), whereas the best global polynomial model is an 18th order polynomial ($p = 19$). All global polynomials have AICc values 80-100 times higher than segmented polynomials, indicating that segmented polynomials are preferred in all cases. The AICc retains no spatial information. We therefore look at the histogram of the residuals and explore any possible correlations in the residuals with pixel number for the two wavelength calibration models preferred by AICc: the 7th order segmented polynomial and the 18th order global polynomial.

The histograms of the residuals for the two models are shown in Figure 3.8. Residuals to the segmented 7th order polynomial model are more centrally concentrated than those of the global 18th order polynomial model. The central 68% (95%) of residuals for the segmented 7th order polynomial model are smaller than $3.4 \, \mathrm{m \, s^{-1}}$ ($7.5 \, \mathrm{m \, s^{-1}}$). The central 68% (95%) of residuals for the global 18th order polynomial model are smaller than $4.3 \, \mathrm{m \, s^{-1}}$ ($9.1 \, \mathrm{m \, s^{-1}}$). Therefore, the lowest AICc segmented polynomial model provides, on average, smaller residuals compared to the lowest AICc global polynomial model.
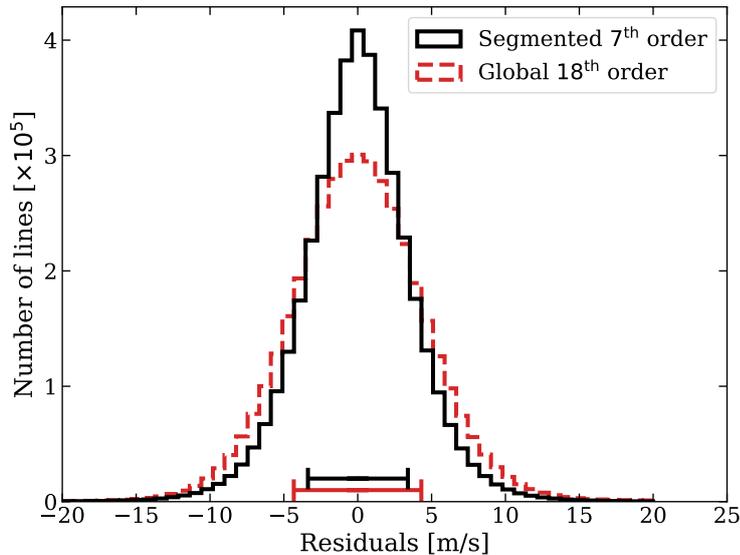
Figure 3.8: The histogram of residuals for the segmented 7th order polynomial in fibre A (solid black) is more centrally concentrated than that of the global 18th order polynomial (dashed red). Horizontal bars show the central 68% of the distribution ($3.4\,\mathrm{m\,s^{-1}}$ and $4.3\,\mathrm{m\,s^{-1}}$ for the black and red, respectively). The total number of LFC lines is $N = 4120422$.

We discover that the residuals obtained from the global polynomial model exhibit a structure that correlates with pixel number. In order to more clearly illustrate the effect, we bin the residuals into 64 bins along the pixel axis (8 bins per 512-pixel wide segment) and calculate the mean and its uncertainty in each bin. The correlation pattern is similar in all 3 blocks in the y-direction on the CCD for fibre A (red circles in Figure 3.9). The same patterns are seen in the fibre B data (not shown). The pattern has amplitudes as high as $4\,\mathrm{m\,s^{-1}}$ with a root-mean square (rms) of $\approx 1\,\mathrm{m\,s^{-1}}$. To test the sensitivity of the pattern to the lines falling close to segment edges, we remove lines falling within 32 and 64 pixels around the segment boundaries and recalculate the means. We find that the pattern persists and neither its amplitude or rms changes significantly. The same result is obtained using bins of different sizes. No such correlation is seen for the segmented polynomial (black squares in Figure 3.9).

The pronounced residuals seen in the red circles in Figure 3.9 are not associated with the pixel anomaly discontinuities illustrated in Figure 3.6, since these have been removed prior to polynomial fitting. However, looking at the characteristics of the continuous black line (and red squares), applying 7 offsets to move the curves together will not yield an overall trend *entirely* free of discontinuities. Thus we may not expect one global polynomial to provide a complete description of the data, even after correcting for the pixel anomalies. Examining figure 5 of Coffinet et al. (2019) indicates the same general phenomenon is found when using flat-fielding methods to quantify and remove the pixel anomalies.

Whether these small remaining discontinuities are sufficient to generate the residual correlations we have found is unclear. What is clear however, is that
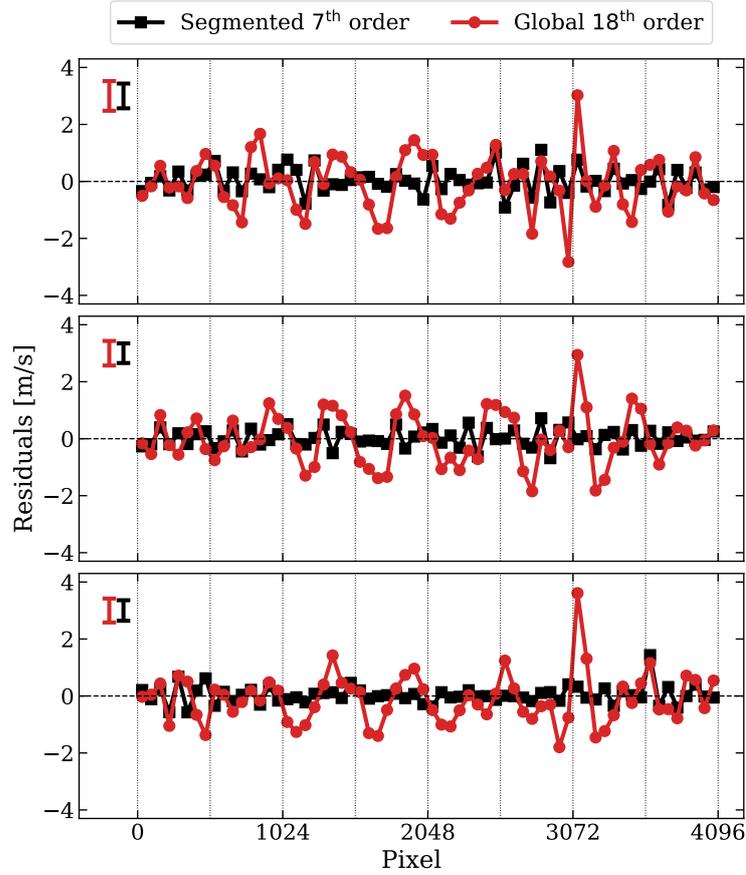
Figure 3.9: Mean values of the wavelength calibration residuals in 64 bins along the dispersion direction in fibre A. Points closer than 32 pixels to the segment boundaries (vertical dashed lines) have been excluded. The global polynomial fit (red line and circles) produces highly correlated residuals. The structure is present and similar across all three HARPS detector blocks. The segmented order polynomial (black line and squares) shows no such structure. The vertical bars in the top left of each panel illustrate 10 times the average error on each point.

global polynomials (i.e. a single polynomial per echelle order) should not be used to calibrate astronomical spectra used for spectroscopic velocity shift measurements. We therefore use the segmented 7$^{\text{th}}$ order polynomial for wavelength calibration in the rest of our analysis.

3.6   Precision and consistency

Two types of velocity shifts are present in LFC time-series measurements. The first impacts on each fibre identically (e.g. velocity drifts caused by pressure and temperature variations in the spectrograph). Velocity shifts like these, common to the two fibres, can be removed using the simultaneous referencing technique (Baranne et al. 1996). The second type of velocity shift operates independently on each fibre. Quantifying the precision and consistency of the two LFCs used in this work relies on measuring the second type of shifts in our dataset. We therefore
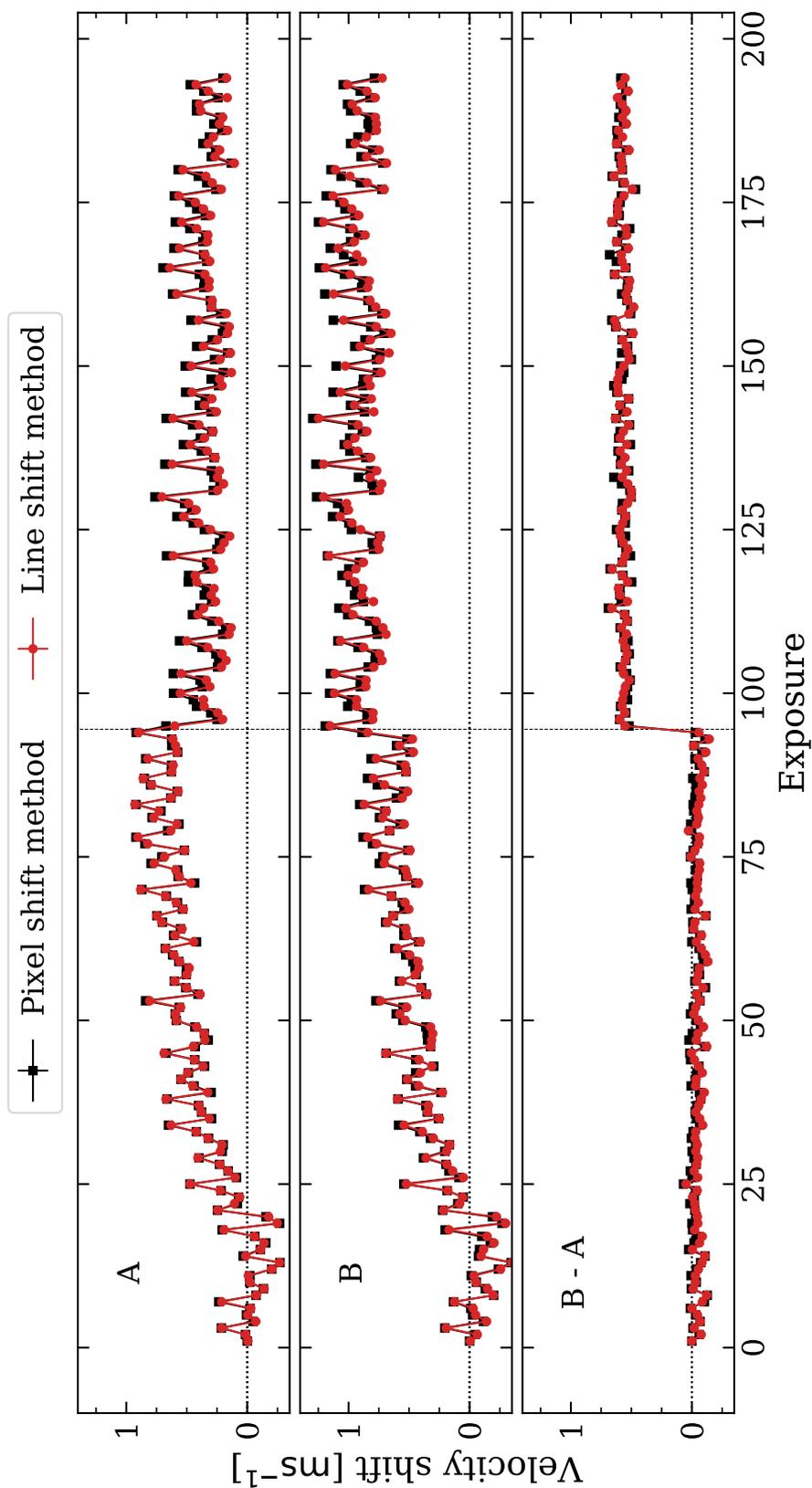
Figure 3.10: Velocity shift measurements for all exposures in fibres A (top panel) and B (middle panel) using the pixel shift method (black line and circles) and the line shift method (red line and squares). The velocities derived by both methods trace each other very well and are virtually indistinguishable by eye. Subtracting the spectrograph drift (B-A, bottom panel) reveals remarkable precision of each LFC (rms $\leqslant 4\,\mathrm{cm\,s^{-1}}$) but also a discrepancy in the absolute scales of the two LFCs of $61 \pm 0.6\,\mathrm{cm\,s^{-1}}$ level. The average photon-limited velocity precision is $2.3\,\mathrm{cm\,s^{-1}}$ ($2.2\,\mathrm{cm\,s^{-1}}$) for fibre A, $2.8\,\mathrm{cm\,s^{-1}}$ ($2.7\,\mathrm{cm\,s^{-1}}$) for fibre B, and $3.6\,\mathrm{cm\,s^{-1}}$ ($3.5\,\mathrm{cm\,s^{-1}}$) for B-A, using using the pixel (line) shift method.

calculate velocity shifts of all exposures in the two fibres and take their differences, effectively removing the first type of shifts.

We measure the mean velocity of each exposure relative to the zero-point set by the first exposure in the time-series. The velocity shift is calculated using all wavelength calibrated pixels, where the shift of the ith pixel in the jth exposure is given by:

$$\frac{\Delta v_i^j}{c} = \frac{\lambda_i^j - \lambda_i^{ref}}{\lambda_i^{ref}}. \tag{3.11}$$

Here, λ_i^{ref} is the ith pixel's wavelength in the reference (first) exposure. The velocity shift of an exposure is the unweighted average of velocity shifts of all pixels. The uncertainty on the velocity shift derived above is the photon-limited velocity precision of all wavelength calibrated echelle orders, calculated using the Bouchy et al. (2001) formalism. The velocity shift of the first exposure is by definition equal to zero. We refer to this method as the "pixel shift" method.

We cross-check velocity measurements of the pixel shift method using a second, independent, one. The second method, which we refer to as the "line shift" method, uses shifts in the LFC line positions on the detector to calculate the average velocity shift of an exposure. This requires a set of reference wavelength calibration coefficients: we use those of the first exposure of the series. The coefficients are used to infer wavelengths of lines in the exposure by evaluating the polynomial at the measured line positions. The velocity shift of the ith LFC line in the jth exposure is:

$$\frac{\Delta v_i^j}{c} = \frac{\lambda_i^j - \lambda_i}{\lambda_i}, \tag{3.12}$$

where λ_i^j is the inferred, and λ_i is the true line wavelength per Equation (3.1). The velocity shift of the exposure is the mean velocity shift of all LFC lines[4], weighted by the errors on the inferred wavelength. The uncertainty on the velocity shift of an exposure is the standard error of the weighted mean. Because of the definition of Equation (3.12), the velocity shift of the first exposure is not necessarily exactly equal to zero.

The results for the entire dataset, using both methods, are plotted in Figure 3.10. The top two panels, corresponding to measurements in fibres A and B, show that spectrograph shifts are not negligible: up to about $1\,\text{m}\,\text{s}^{-1}$ in the six hours of duration of the test, as measured by the shift of each fibre. However, shifts in the two fibres trace each other remarkably well, as can be seen in the bottom panel of Figure 3.10, showing their differential shift (B-A). A $\approx 60\,\text{cm}\,\text{s}^{-1}$ velocity offset in the differential shift occurs after exposure 94, corresponding to the change from LFC2 to LFC1 in fibre B. The differential shift, B-A, should contain only those shifts that are either inherent to the two LFC systems or which influence the two fibres independently – and is therefore relevant for quantifying the precision and consistency of LFCs.

---

[4]A cut-off velocity $200\,\text{m}\,\text{s}^{-1}$ was imposed to eliminate a very small number of spurious measurements (44 lines or 0.0001% of the sample), probably associated with large line-centre uncertainties for lines with very low flux

We estimate the precision of LFC1 and LFC2 as the rms of the differential shift in the LFC1 and LFC2 samples, respectively. The precision of LFC1 sample is $4.5\,\mathrm{cm\,s^{-1}}$ ($4.0\,\mathrm{cm\,s^{-1}}$) and of LFC2 sample is $3.7\,\mathrm{cm\,s^{-1}}$ ($3.5\,\mathrm{cm\,s^{-1}}$) using the pixel (line) shift method. This precision is $\approx 10\%$ above the average photon-limited velocity precision, which is $3.8\,\mathrm{cm\,s^{-1}}$ for LFC1 and $3.3\,\mathrm{cm\,s^{-1}}$ for LFC2 (see Section 3.3).

The consistency between the two LFCs – i.e. the jump recorded at exposure 95 in the differential shift – is $60.4 \pm 0.6\,\mathrm{cm\,s^{-1}}$ for the pixel shift method, and $61.8 \pm 0.6\,\mathrm{cm\,s^{-1}}$ for the line shift method. In order to understand the shift, one must consider that LFC1 and LFC2 have significant differences, namely different mode separations and offset frequencies. Changing from LFC2 to LFC1 is thus a major change in the calibration system akin to switching from a ThAr to a U hollow cathode lamp. Major changes in the calibration system are almost always associated with a jump in the instrumental zero-point. This implies that, in addition to the photon noise, all systematic effects associated with the wavelength calibration process will determine the consistency between LFC1 and LFC2. These include changes in the light injection into the fibres, insufficient temporal or spatial scrambling of the fibres, differences in the light path, line-spread function (LSF) variation across the detector, charge transfer inefficiency (CTI), fringing, data reduction techniques, and fitting of the data.

Zhao et al. (in prep.) analysed data in which tests of this nature were performed in 2012 on the HARPS LFC prototype. Their analysis of a series of 1713 exposures shows that extreme changes to the calibration system (e.g. exchanging the photonic crystal fibre, changing the light injection, disabling the mechanical scrambler, light scrambling using the integrating sphere, mechanical realignments, etc.) produce velocity shifts with a standard deviation of $45\,\mathrm{cm\,s^{-1}}$. Differences in illumination therefore cannot fully explain the observed $60\,\mathrm{cm\,s^{-1}}$ jump between the two LFCs. With the exception of CTI and LSF, none of the aforementioned effects can be modelled and corrected retroactively as no suitable data were collected during the campaign.

The impact of CTI on spectroscopic velocity measurement was first measured by Bouchy et al. (2009) on the SOPHIE spectrograph. The authors of this study used a series of ThAr lamp exposures finding a clear correlation between the measured velocity shift of an exposure and its flux. Whereas shifts are as high as several tens of $\mathrm{m\,s^{-1}}$ at low flux ($\lesssim 600\,e^-$) observations on SOPHIE, they estimate that the effect is 2-3 times less severe on HARPS because of improved CCD performances and smaller pixels. Optimal mitigation of CTI, however, requires the acquisition of proper calibration spectra and correction of the raw frames before software post-processing, and is hence beyond the scope of this work. In what follows, we use archival LFC observations to produce a simple model to correct flux dependent velocity shifts in HARPS spectra and apply it to our data. As far as LSF reconstruction is concerned, we will report on our work on reconstructing the LSF of HARPS in separate work, with a focus on wavelength calibration accuracy.
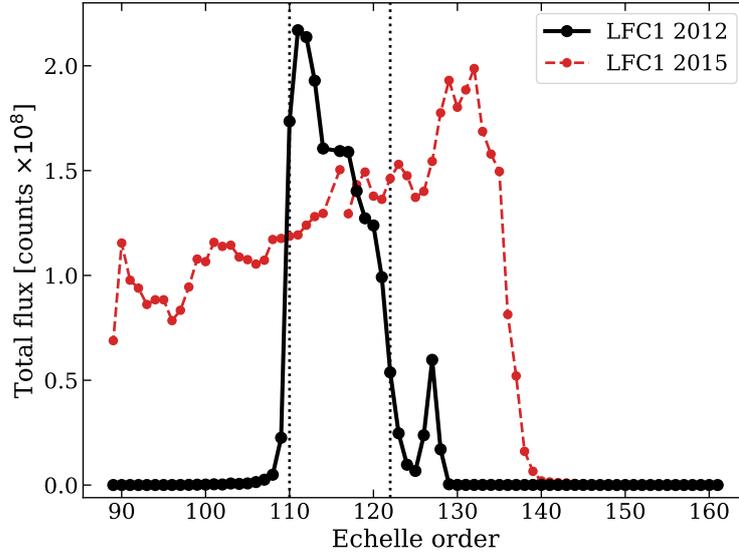
Figure 3.11: The wavelength range of the LFC1 prototype (2012, solid black line) is significantly shorter than that of the final design (2015, dashed red line). Our model of signal dependent velocity shifts therefore uses the average flux per line in place of the total exposure flux as in Bouchy et al. (2009). We use echelle orders 110 to 122 (marked by the dotted vertical lines), where the flux levels are sufficiently high to be sure not to miss or falsely detect lines.

3.6.1   Contribution of flux dependent shifts

Following the methodology of Bouchy et al. (2009), we look for correlations between flux in an exposure and its velocity shift using data collected during the LFC campaign of February 2012, when the LFC1 prototype was installed on HARPS. The prototype had minor differences with respect to the final LFC (installed during the April 2015 campaign), the most notable being the shorter wavelength range coverage. The prototype illuminated only $\approx 33\%$ of the total wavelength range of HARPS, between 475 nm and 580 nm (echelle orders 106-128). However, we use only orders 110 to 122, in which the flux is sufficiently high to be sure of not missing or falsely detecting lines. A comparison with the wavelength coverage of the final design is shown in Figure 3.11.

The spectral flattening unit of the LFC was not optimised at that time, resulting in strong and fast fluctuations in line amplitudes within a single echelle order (Probst 2015). We also see a much smaller background component in the 2012 spectra, with an average B2E ratio of less than 1%. As noted before, the background is likely caused by the amplification of laser light before entering the photonic crystal fibre, in which the background is further amplified by non-linear processes. The background levels are lower in the 2012 data because the power in the amplification stage was significantly lower in the 2012 setup.

Data and methods

We use fifteen sets of 10 exposures each of the LFC1 prototype, taken on 15th February 2012. The first and the last sets were taken under nominal conditions, whereas the thirteen sets in between had neutral density filters of different values inserted into the light path (see section 6.3 in Probst 2015). The exposures were taken over a time span of 7 hours. Exposure time was $40\,\mathrm{s}$ with $22\,\mathrm{s}$ read-out.

The February 2012 exposures were reduced by version v3.5 of the standard HARPS pipeline and made public via the ESO archive. We re-reduce a part of this data using pipeline version v3.8 (used to reduce the rest of our data) and find no significant differences between the products of the two pipeline versions. We therefore use the archived data (i.e. version v3.5) in this subsection.

We detect, fit, and identify all lines in echelle orders 110 to 122 in all exposures (see Section 3.4) and wavelength calibrate them (see Section 3.5). We then calculate velocity shifts using both the pixel shift and line shift methods described previously in this section. Given the different wavelength ranges covered by the prototype and the final design of LFC1, the total flux of an exposure is not a relevant quantity with which we can quantify flux dependent velocity shifts in the 2015 data. We therefore use the average flux per line taken over echelle orders 110 to 122. Individual line fluxes vary significantly not only across orders but also within each individual order in this data. Nonetheless, a clear trend between the average flux per line and average velocity shift is present for both fibres (Figure 3.12).

Flux dependent velocity shifts are different for the two fibres. This is unexpected and currently not understood. This indicates that other effects, in addition to CTI, affect the velocity-shift dependency on flux. The shift to negative velocities in the last set of calibrations (without filter) are not due to flux, but spectrograph drifts with time over the duration of the test (see Figure 3.12). Assuming a linear drift with time, we correct each exposure for the temporal component of the velocity drift by fitting a straight line to the mean observing time of the first and the last set of exposures (both without filter). We subtract this temporal drift prior to focusing on the flux dependence.

Model

We model the flux dependency of velocity shifts with a simple exponential model of flux:

$$v(f) = a \exp\left(-f/b\right) \ [\mathrm{m\,s^{-1}}], \tag{3.13}$$

where v represents the velocity shift of an exposure with an average flux per line f. We correct each exposure for temporal drift and subsequently bin them into fifteen sets of ten prior to fitting. We use least-squares fitting to determine the values of parameters a and b from the data, producing four separate models: one for each combination of fibre and velocity shift measurement method (Table 3.3). An example of the fit for fibre A and the line shift method is shown in Figure 3.13.

Table 3.3: Parameters of the exponential model (Equation 3.13) for each fibre and each velocity shift method.

Fibre	Method	$a\,(\mathrm{m\,s^{-1}})$	$b\,(\times10^4)$
A	Pixel shift	3.00 ± 0.17	9.69 ± 0.90
	Line shift	2.52 ± 0.15	10.17 ± 1.02
B	Pixel shift	1.61 ± 0.09	11.80 ± 1.25
	Line shift	1.40 ± 0.10	11.75 ± 1.55
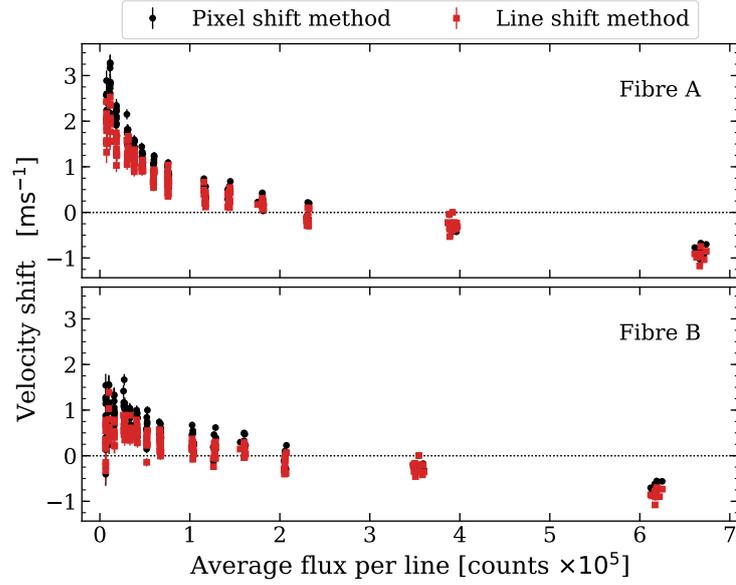
Figure 3.12: There is a clear trend in velocity shift with the average flux per LFC line. The amplitude of the shift is different for fibres A and B, as well as for the two methods we use to calculate shifts. Negative velocity shifts at the highest fluxes are due to spectrograph drifts over the duration of the series. This is because flux dependent velocity shifts are negligible for those points. This temporal shift is removed before modelling the flux dependency.

Application to 2015 data

We apply the flux dependency models (Table 3.3) derived from the 2012 data to the measured velocity shifts in our 2015 data. The average flux per line in the 2015 data is calculated over the same orders as are used to derive the model, plotted on Figure 3.14 as a function of exposure number. Figure 3.15 shows the impact of the flux correction on the measured velocity shifts. Whereas LFC2 sample velocities are mostly unchanged due to their high average flux, LFC1 sample velocities shift by $\approx -15\,\mathrm{cm\,s^{-1}}$, with an uncertainty from the model of $\approx 2\,\mathrm{cm\,s^{-1}}$. Applying the correction improves the precision of LFC1 by $\lesssim 5\%$ and improves the consistency between LFC1 and LFC2 to $43.1 \pm 0.6\,\mathrm{cm\,s^{-1}}$ ($46.9 \pm 0.6\,\mathrm{cm\,s^{-1}}$) using the pixel (line) shift method. This is an improvement in absolute value of $\approx 25\%$ and is perhaps surprising given the simplicity of the model. This result demonstrates that the effect of flux on the precision and consistency of LFC calibrations is not

Table 3.4: A tabular overview of our LFC precision and consistency results. We calculate velocity shifts of LFC exposures for LFC1 and LFC2 samples using two independent methods. We tabulate the average photon-limited velocity precision of the sample (μ_{PL}, Bouchy et al. (2001), see Section 3.6) in column 5. The empirical precision achievable from a single exposure (rms) of the sample is tabulated in column 6. The difference between the mean velocities of each sample (μ in column 7 and the corresponding error σ_μ in column 8) reveal an offset in the velocity zero-point when switching from LFC2 to LFC1 in fibre B after exposure 94. The consistency between the absolute calibrations of the two LFCs (column 9) is thus $\approx 61 \pm 0.6\,\mathrm{cm\,s^{-1}}$. Allowing for flux-dependent velocity shifts in the data improves the consistency by $\approx 25\%$, bringing it to $\approx 45 \pm 0.6\,\mathrm{cm\,s^{-1}}$.

Method	Flux corrected	Sample	Exposures	μ_{PL} ($\mathrm{cm\,s^{-1}}$)	Precision ($\mathrm{cm\,s^{-1}}$)	μ ($\mathrm{cm\,s^{-1}}$)	σ_μ ($\mathrm{cm\,s^{-1}}$)	Consistency ($\mathrm{cm\,s^{-1}}$)
Pixel shift	No	LFC2	1 - 94	3.4	3.7	-3.2	0.4	$+60.4 \pm 0.6$
		LFC1	95-194	3.8	4.5	$+57.2$	0.5	
Line shift	No	LFC2	1 - 94	3.3	3.5	-5.2	0.4	$+61.8 \pm 0.6$
		LFC1	95-194	3.7	4.0	$+56.6$	0.4	
Pixel shift	Yes	LFC2	1 - 94	3.4	3.7	-3.2	0.4	$+43.1 \pm 0.6$
		LFC1	95-194	3.8	4.4	$+39.9$	0.4	
Line shift	Yes	LFC2	1 - 94	3.3	3.5	-5.2	0.4	$+46.9 \pm 0.6$
		LFC1	95-194	3.7	3.8	$+41.7$	0.4	
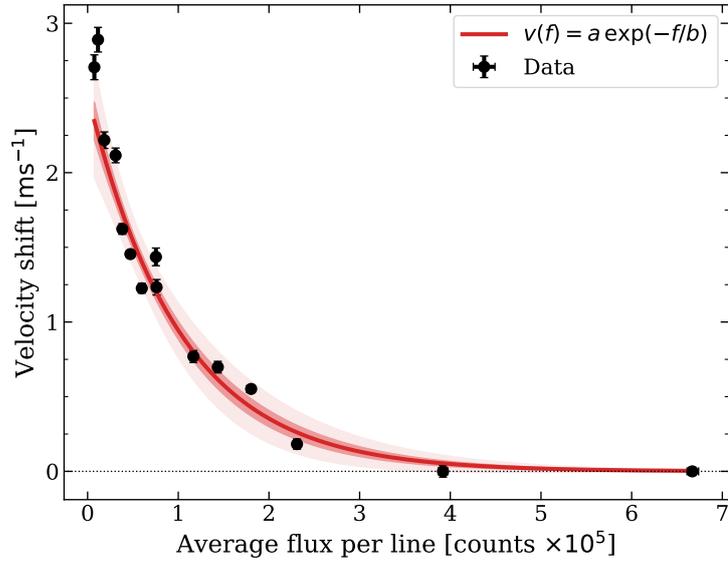
Figure 3.13: We fit an exponential function to the mean velocity shift of exposures taken through each neutral density filter used (black circles, 10 exposures per point) from the 2012 data, after correcting for a temporal shift component. We subsequently apply the model derived here to the 2015 data. The plot shows an example for a model in fibre A and the line shift method. The dark and light shaded areas correspond to 1σ and 3σ uncertainties on the model.
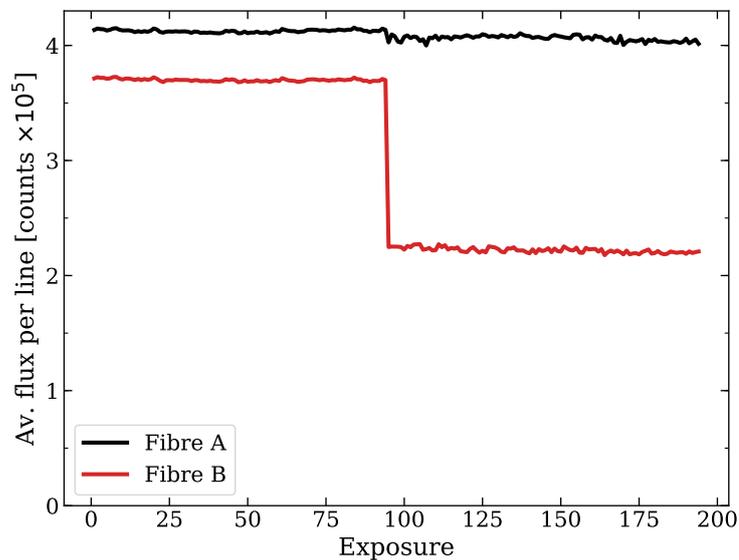
Figure 3.14: The average flux per line in orders 110–122 from our 2015 dataset. Fibre A (black) carried LFC2 light throughout the series. Fibre B (red) carried LFC2 light for the first 94 exposures, after which it carried LFC1 light. The average flux per LFC line in fibre B is 90% of the flux in fibre A for the first 94 exposures, after which it drops to 55%.
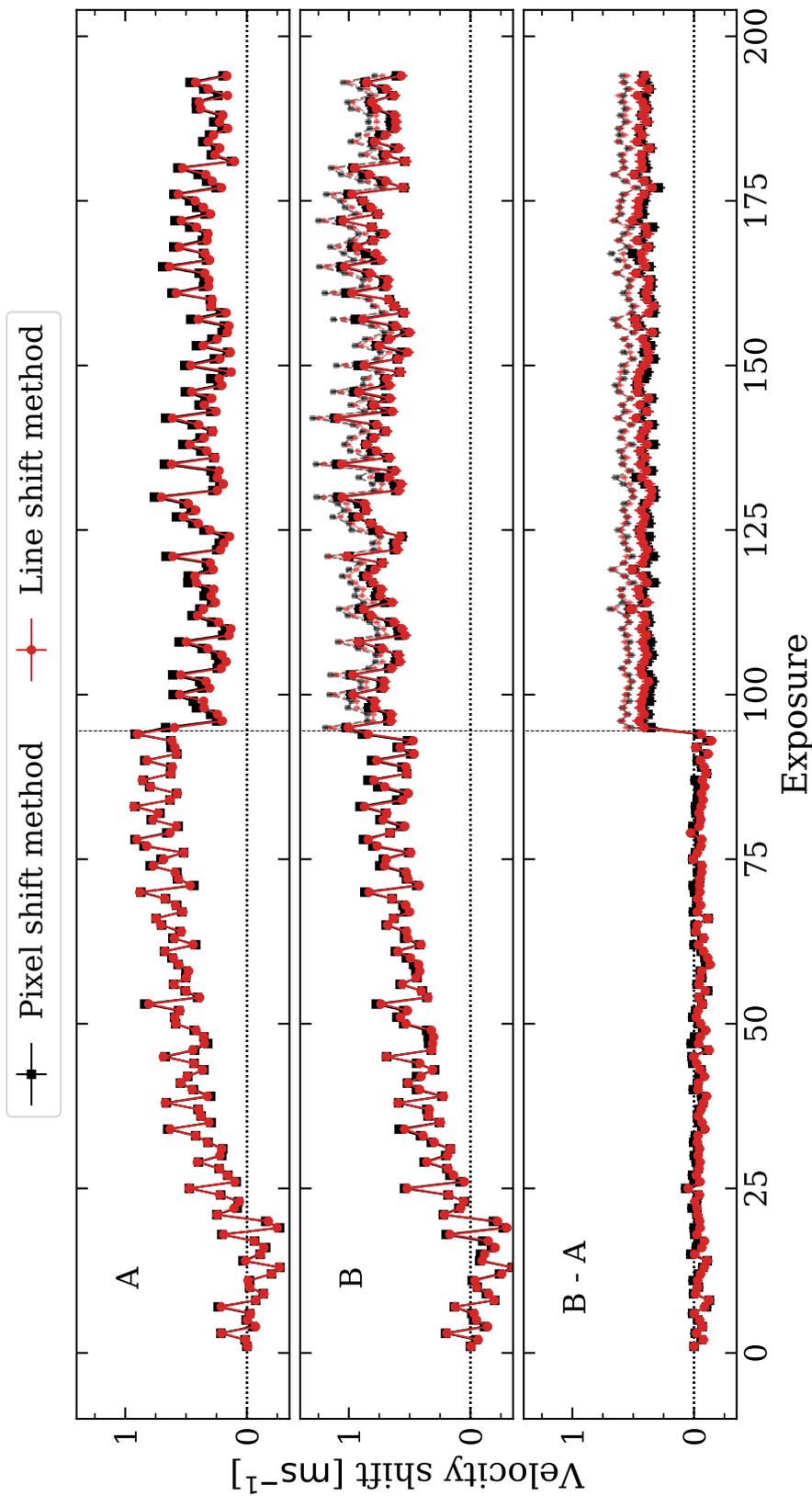
Figure 3.15: The same as Figure 3.10, but after correcting for flux dependent velocity shifts. This improves the consistency between LFC1 and LFC2 by $\approx 25\%$, bringing it to $\approx 45 \pm 0.6\,\mathrm{cm\,s^{-1}}$. The original, uncorrected, measurements are shown as a thin dotted red and black lines.
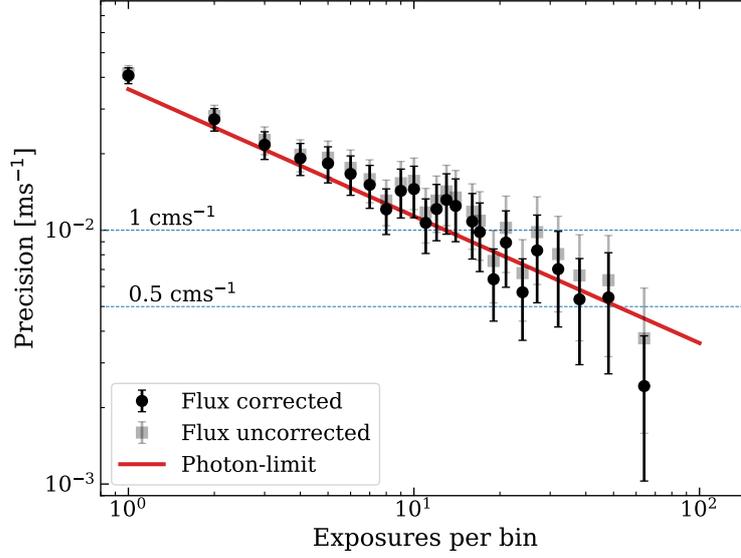
Figure 3.16: The calibration precision of LFC velocity measurements (as measured by the standard deviation) improves as we bin the measurements in bins of increasing size in a way that is in agreement with the photon-limited precision (red line). The precision is additionally improved after correcting for flux dependent effects.

negligible. This strongly motivates the importance of detailed measurements and parameterisation of such dependencies when attempting astronomical calibrations and observations approaching $\approx 1 \, \mathrm{cm\,s^{-1}}$.

3.6.2   Achievable precision

In applications such as extrasolar planet radial velocity measurements, it is of interest to explore the limiting calibration precision achievable in a reasonable observing time. To examine this, we begin by taking the data illustrated in Figure 3.15 but remove the break in mean velocity shift seen at exposure 94. This is done simply by calculating the means either side of exposure 94 (see Table 3.4) and removing both i.e. normalising to zero means either side of exposure 94. The data have previously been flux-corrected as described in Section 3.6.1 to account for the different flux levels of the two LFCs. The rms velocity shift is then calculated for all 194 points. We then bin the number of exposures in increasingly large bins, starting with 2 exposures per bin, and increasing the number of points per bin. This is illustrated in Figure 3.16, which shows that for maximal binning, a radial velocity calibration precision $\approx 0.5 \, \mathrm{cm\,s^{-1}}$ can in principle be achieved. Comparing this with the expectation based on the photon-limited velocity precision (continuous red line in Figure 3.16), we see good agreement between the theoretical prediction and observations.

The 194 exposures used for the procedure above correspond to a total integration time of 1.6 hours (taken over a period of 6 hours - see Section 3.3). We conclude from this that, given the corrections applied above, a realistic achievable calibration precision is of order $1 \, \mathrm{cm\,s^{-1}}$.

3.7 Results

Our main results are:

1. Using global polynomials for wavelength calibration produces residuals which correlate with pixel number (see Figure 3.9) even when high (18th) order polynomials are used. The pattern is highly modulated, has amplitudes as high as $4\,\mathrm{m\,s^{-1}}$, and is present across the detector.

2. We find absolute velocity shifts between the wavelength solutions measured using two independent LFCs of $\approx 61 \pm 0.6\,\mathrm{cm\,s^{-1}}$ when only HARPS instrumental drifts are removed. This unexpected result can be partly attributed to differences between the two LFC flux levels, since a clear (non-linear) correlation is seen between the mean flux and a mean shift velocity in spectral line positions (Figure 3.13). However, even allowing for this, a significant absolute shift remains: $\approx 45 \pm 0.6\,\mathrm{cm\,s^{-1}}$ or $\Delta\lambda/\lambda = 1.3 \times 10^{-9}$ (Figure 3.15).

3. The precision of each LFC in a single exposure is $\lesssim 4\,\mathrm{cm\,s^{-1}}$ ($\approx 10\%$ higher than the measured photon-limited precision). Precision remains unchanged when a different comb is injected in the second fibre demonstrating it remains unchanged by using two independent systems. A realistic achievable velocity calibration precision is of order $1\,\mathrm{cm\,s^{-1}}$ provided systematics are carefully measured and removed (see Section 3.6.2).

3.8 Discussion

We set out to determine the limiting precision with which spectroscopic velocity shifts in high resolution spectra can be measured using current methodology, in the context of the science goals motivating the construction of future large optical observatories. In this sense, we have achieved our goals. Firstly, we demonstrated that LFCs can achieve repeatabilities of around $1\,\mathrm{cm\,s^{-1}}$ using advanced methods in conjunction with a second, independent, LFC. We thus go beyond the results of Wilken et al. (2012) where only one LFC was used. Secondly, by comparing measurements from two independent LFCs, we discovered unexpected and substantial wavelength zero-point offsets between LFCs, the causes of which are not yet completely understood.

We identify systematics introduced into the LFC measurements by the combined effects of the detector morphology, the CTI during data read-out, and imperfect LSF modelling (i.e. the Gaussian approximation). We have not identified any systematic effects associated with the LFC itself. This implies that improvements in the former three will enable precision improvements that approach theoretical limits. A tunable LFC, capable of scanning the full separation between two LFC modes would be ideal to better understand the system. LFCs with large ($\approx 10\,\mathrm{GHz}$) native mode separations should be able to provide such a feature, e.g. those based on electro-optic combs (Obrzud et al. 2019).

It was only possible to quantify the zero-point offset resulting from the change of the LFC because two LFCs were used simultaneously on HARPS. This would

not have been possible if only one LFC had been used, but instead the first LFC had simply been replaced by a second. This point merits careful consideration when designing astronomical measurements requiring long-term stability. Whilst observing a set of radial velocity standards before and after the LFC change might be sufficient for exoplanet detection studies, achieving the stability required for the redshift drift measurement warrants a different approach.

Lastly, we discovered highly correlated wavelength residuals resulting from employing global polynomials for wavelength calibration – the default method in essentially all previous echelle spectroscopy. The discovery was made whilst investigating different calibration algorithms (global versus segmented polynomials, Section 3.5) and could only be made due to the large number of LFC lines available.

The expected effect of the correlated residuals is to introduce spurious velocity shifts in the data. The severity of this effect depends both on the science goal of observations and on individual characteristics of the target: the number of useful lines and where they fall with respect to the correlated structure. For example, the most precise redshift measured to date for any single heavy element absorption line at high redshift, using optical spectroscopy, has a redshift uncertainty of around 4×10^{-6}, or around $5 \, \mathrm{m \, s^{-1}}$. If global polynomials are used, correlated calibration residuals may emulate a varying fine structure constant at a level $\Delta\alpha/\alpha$ of around 10^{-6} in this single line. This is of the same order as the statistical error in this system. Radial velocity measurements from stellar spectra will be influenced in a similar way, where the signal could emulate periodicity. The exact period of the spurious signal will depend on time sampling of observations in addition to which lines are used for the measurement. Therefore, the correlated residuals also have the potential to emulate spurious exoplanet detections. Finally, the expected signal in the redshift drift measurement is of order $1 \, \mathrm{cm \, s^{-1}}$ (see Fig. 2 in Liske et al. 2008). Correlated residuals at the level of $4 \, \mathrm{m \, s^{-1}}$ would therefore render detection of redshift drift impossible. However, the results presented here are rather encouraging: provided segmented polynomials are used (assuming existing technology), the calibration precision of $\approx 1 \, \mathrm{cm \, s^{-1}}$ has now just about been reached.

Chapter 4

A new era of fine structure constant measurements at high redshift

4.1 Introduction

Fundamental constants, such as the fine structure constant ($\alpha \equiv \frac{1}{4\pi\epsilon_0}\frac{e^2}{\hbar c}$) and the proton-to-electron mass ratio ($\mu \equiv \frac{m_p}{m_e}$), are expected to vary in some modifications of General Relativity. A scalar field ϕ coupling to the baryonic matter can produce temporal and/or spatial α variations (Bekenstein 1982; Sandvik et al. 2002; Shaw & Barrow 2005; Barrow & Lip 2012; Copeland et al. 2004; Marra & Rosati 2005). α may also vary with gravitational potential (Dicke 1959; Sandvik et al. 2002; Mota & Barrow 2004b,a), or via interactions of baryonic matter with dark matter candidates (Olive & Pospelov 2002; Stadnik & Flambaum 2015), or if the vacuum expectation value of ϕ depends on the local density (Silva et al. 2014). In theories with extra spatial dimensions (e.g. Kaluza-Klein and string theories), expansion (or contraction) of higher dimensions can produce observed changes to the coupling constants in our 4-dimensional space time. Recent reviews of varying constants are given by Uzan (2011); Martins (2017).

Variations in α and μ have been explored both on Earth through atomic clock measurements (Rosenband et al. 2008), isotope ratio studies (Damour & Dyson 1996), and in space using astronomical observations of white dwarfs (Berengut et al. 2013; Bainbridge et al. 2017), galaxies (Bahcall et al. 2004), quasars (Webb et al. 1999; Murphy et al. 2003a; Wilczynska et al. 2015; Ubachs 2018), stars around the supermassive black hole in the Galaxy (Hees et al. 2020) and the Cosmic

Microwave Background (Avelino et al. 2001; Planck Collaboration et al. 2015). A comprehensive analysis of 317 quasar absorption systems using the *Many Multiplet method* (Dzuba et al. 1999b; Webb et al. 1999) hinted at a spatial variation of α, modelled as a dipole with amplitude $\Delta\alpha/\alpha = (\alpha_Q - \alpha_0)/\alpha_0 = 1.1 \pm 0.2 \times 10^{-6}$, where α_Q are quasar absorption measurements and α_0 is the terrestrial value (Webb et al. 2011; King et al. 2012; Wilczynska et al. 2020).

Echelle spectrographs, using slit-based observations and calibrated using ThAr methods, are prone to long-range wavelength distortions (Molaro et al. 2008; Rahmani et al. 2013; Evans et al. 2014). Such distortions, if present and left uncorrected, can significantly contribute to the total $\Delta\alpha/\alpha$ measurement uncertainty (Evans et al. 2014; Kotuš et al. 2017). Correction techniques include using additional external calibration information from asteroid observations (Molaro et al. 2013b; Rahmani et al. 2013), iodine cells (Griest et al. 2010; Whitmore et al. 2010), solar-twin observations (Whitmore & Murphy 2015), or by using additional model parameters (Dumont & Webb 2017). To a reasonable approximation, the best wavelength correction that could be achieved with any of these methods has an accuracy no better than $\lesssim 30\,\mathrm{m\,s^{-1}}$. For comparison, the best laboratory accuracy of UV wavelengths used for $\Delta\alpha/\alpha$ measurements is $0.01\,\mathrm{m\,s^{-1}}$, three orders of magnitude better. Laser Frequency Comb (LFC, Udem et al. 2002; Hänsch 2006; Steinmetz et al. 2008) wavelength calibration methods provide a vastly superior calibration than the correction methods above as they provide $3\,\mathrm{m\,s^{-1}}$ accuracy for individual line center measurements (Chapters 2 and 3).

In this Chapter, we report a set of high redshift α measurements from new observations of the quasar HE0515$-$4414. The observations (described in Section 4.2) are of very high quality. These data are the first quasar spectral observations where the wavelength calibration has been carried out using an LFC. This means that any wavelength scale distortions present will be negligible. A second spectrum was produced from the same quasar observations but calibrated using ThAr methods. The two spectra enable a unique set of comparative tests to quantify uncertainties in searches for fundamental constant variations.

We use new automated analysis methods (Lee et al. 2020) to produce models for each spectrum and measure α using the Many Multiplet method (Section 4.3). We introduce a new method, measuring α for each absorption component (rather than an average across an entire absorption complex). This provides considerably more detail and also offers a substantial advantage by enabling systematics to be more readily identified. We summarise our main findings in Section 4.4 and discuss them in Section 4.5.

4.2   Data

4.2.1   Data acquisition

The spectrum used in this work was produced from high-resolution ($R = \frac{\lambda}{\Delta\lambda} = 115000$) observations using the High Accuracy Radial velocity Planet Searcher echelle spectrograph (HARPS, Mayor et al. 2003). HARPS is a double-channel echelle spectrograph built for extremely precise spectroscopic measurements. We

observed the quasar HE0515−0414 (abbreviated as HE0515) between 3rd and 11th December 2018 using HARPS in the classical fibre spectroscopy mode, where channel A recorded the HE0515 spectrum and channel B recorded the sky spectrum. We obtained 36 exposures totalling 52h 31m (Table 4.1). Each exposure was bracketed by ThAr and LFC exposures for wavelength calibration. The sky was dark and the seeing conditions varied between 0.45 and 1.98 arcsec throughout the observing run. The median seeing (i.e. the median of column 3 in Table 4.1) is 1.34 arcsec. This has no influence on the final spectral resolution. The secondary guiding system ensures the object is consistently centered on the object image up to 0.01 arcsec and octagonal fibres ensure that the light evenly illuminates the spectrograph pupil. Therefore, telescope guiding and fibre illumination are not expected to introduce spectroscopic velocity shifts larger than $0.12\,\mathrm{m\,s^{-1}}$ (Lo Curto et al. 2015).

Light entering the spectrograph is recorded on the detector (a mosaic of two EEV2k4 CCDs) for which the read-out mechanism is located on one of its sides (Mayor et al. 2003; Rodler & Lo Curto 2019). By design, charge transfer occurs in the cross-dispersion direction to minimise effects of charge transfer inefficiency (CTI). Left uncorrected, CTI can introduce spurious spectroscopic velocity shifts up to $3\,\mathrm{m\,s^{-1}}$ for very low flux exposures (Section 3.6.1). However, as no appropriate CTI model yet exists for HARPS, we do not correct for this effect.

Table 4.1: The final co-added spectrum of HE0515−4414 is formed from co-adding 36 HARPS exposures taken in classic spectroscopy mode, totalling 52h 31m. Columns 1 and 2 give the observing time start (in UTC) and the exposure time, respectively. Column 3 gives the average of the telescope seeing recorded at the beginning and the end of observation. Column 4 gives the S/N per extracted pixel at the center of order 111 ($\approx 5500\,\text{Å}$). All quantities are determined from values recorded in headers of `e2ds` HARPS pipeline products.

Observing time (UTC)	Exp. time (s)	Seeing (arcsec)	S/N (pix^{-1})
2018-12-04T00:27:52.031	5400	1.48	6.4
2018-12-04T02:12:04.582	5400	1.35	11.0
2018-12-04T03:50:40.736	5400	1.54	8.3
2018-12-04T05:36:27.032	5400	1.71	5.6
2018-12-04T07:14:20.953	2700	1.25	3.2
2018-12-05T03:14:39.850	5400	1.91	7.3
2018-12-05T04:52:04.530	5400	0.45	7.1
2018-12-05T06:42:00.250	5400	N/A	7.5
2018-12-06T00:41:08.634	5400	1.98	5.8
2018-12-06T02:30:00.882	5400	1.64	9.0
2018-12-06T04:08:04.226	5400	1.40	4.6

(the table continues on the next page)

Observing time (UTC)	Exp. time (s)	Seeing (arcsec)	S/N (pix^{-1})
2018-12-06T05:46:25.189	5400	1.58	8.2
2018-12-06T07:25:04.433	5098	1.59	6.4
2018-12-07T00:23:24.209	5400	1.50	7.1
2018-12-07T02:01:11.070	5400	1.33	9.1
2018-12-07T03:38:28.641	4905	1.66	7.2
2018-12-07T05:32:14.425	5400	1.31	7.9
2018-12-07T07:09:39.678	5400	1.40	5.3
2018-12-08T00:32:12.597	4214	1.32	5.0
2018-12-08T02:15:25.854	5400	1.44	5.9
2018-12-08T03:54:16.167	5400	1.32	9.0
2018-12-08T05:32:25.299	5400	1.39	11.4
2018-12-08T07:10:06.569	5400	1.26	9.6
2018-12-09T00:37:46.416	5400	1.29	8.2
2018-12-09T02:22:05.279	5400	1.21	10.1
2018-12-09T03:59:32.258	5400	1.17	9.3
2018-12-09T05:36:04.415	5400	1.10	11.0
2018-12-10T00:24:18.778	5400	1.30	9.2
2018-12-10T02:15:06.228	5400	1.69	9.1
2018-12-10T03:54:46.885	5400	1.23	8.6
2018-12-10T05:32:28.955	5400	1.07	10.6
2018-12-10T07:10:20.256	5400	0.80	13.2
2018-12-11T00:28:14.801	5400	1.04	8.9
2018-12-11T02:12:13.372	5400	1.56	11.2
2018-12-11T03:50:37.414	5400	1.27	10.7
2018-12-11T07:31:21.735	4795	1.05	8.6

Raw images were reduced using the HARPS pipeline (version 3.8, Rodler & Lo Curto 2019). The pipeline extracts 1d spectra of individual echelle orders following optimal extraction by Horne (1986) and Robertson (1986). Order tracing and pixel weights are determined from tungsten-lamp frames taken at the beginning of each night. Pipeline products previously demonstrated a $0.01\,\mathrm{m\,s^{-1}}$ precision (Section 3), so we do not expect spectroscopic velocity shifts associated with its use.

4.2.2   Wavelength calibration and data addition

Wavelength calibration was obtained from LFC and ThAr frames taken immediately before each quasar exposure. The LFC has an offset frequency of $4.58\,\mathrm{GHz}$ and $18\,\mathrm{GHz}$ line separation. LFC wavelength calibration was performed using

eight 7$^{\text{th}}$ order polynomials per echelle order. Each echelle order spans eight 512-pixel blocks on the CCD (Wilken et al. 2010; Molaro et al. 2013b). The accuracy of the LFC wavelength calibration is $3\,\text{m\,s}^{-1}$, measured by the root-mean-square (rms) of calibration residuals (i.e. known LFC line frequency minus the frequency determined from the wavelength solution at line position on the detector, see Section 3.5.2). The ThAr wavelength calibration was produced by the HARPS pipeline using a single third order polynomial per echelle order, with an accuracy of $27\,\text{m\,s}^{-1}$. The average difference in the two calibrations, considering wavelengths $\lambda \geqslant 5000\,\text{Å}$, is $-1.13\,\text{m\,s}^{-1}$ (LFC minus ThAr).

Comparing the true LFC line wavelengths to the ThAr-calibrated wavelengths at their location on the detector reveals a distortion pattern in the ThAr calibration, illustrated on Figure 4.1. The pattern shows no long-range wavelength trends, but contains discontinuities associated with stitching of the HARPS detector (Wilken et al. 2010; Molaro et al. 2013b), not accounted for by the pipeline calibration procedure. We discuss the impact of this distortion pattern on α measurements in Section 4.5.

Over the 8 nights of our run, the spectrograph stability is $0.52\,\text{m\,s}^{-1}$, as illustrated in Figure 4.2[1]. This number was obtained by measuring the average shifts of LFC line positions in individual exposures with respect to their position in the first exposure and calculating the rms. Applying the same method to the ThAr lines gives an rms of $2.87\,\text{m\,s}^{-1}$, six times larger.

Although all LFC exposures were taken under the same nominal conditions and with same exposure times, it turned out that one exposure was substantially better (in terms of flux) than all others (Figure 4.3). Therefore, after careful consistency checking between multiple LFC exposures, this highest flux LFC exposure was used to wavelength calibrate all quasar exposures for wavelengths $\lambda \geqslant 5000\,\text{Å}$ (the LFC data cuts off below this wavelength). There are no saturated LFC lines. We do not follow the same procedure for ThAr calibration, but calibrate each quasar exposure using the ThAr frame taken immediately beforehand.

We transform the LFC and the ThAr wavelength scales to the Solar system barycentre rest-frame using the barycentric velocity shift correction provided by the HARPS pipeline, independently for each quasar exposure. The barycentric correction provided by the pipeline is based on Bretagnon & Francou (1988) and uses the flux-weighted average time of observation. This value agrees down to several $\text{mm\,s}^{-1}$ with our independent calculation, using the same information and the `astropy` module[2].

Finally, we rebin the individual extracted spectra onto a common wavelength grid using a custom routine and sum them together, weighting each pixel by its error estimate (which includes the Poissonian error term, the read-out noise, and the dark current). The error array extracted during this procedure agrees with the estimate derived from flux rms over $\approx 1\,\text{Å}$ range. The final co-added spectrum has an average signal-to-noise ratio (S/N) near 50 per $0.015\,\text{Å}$ pixel in the continuum. This data extraction process was performed for the LFC and ThAr calibration

---

[1]This is not the same as the precision which can be achieved in the simultaneous referencing observing mode.

[2]https://docs.astropy.org/en/stable/coordinates/velocities.html
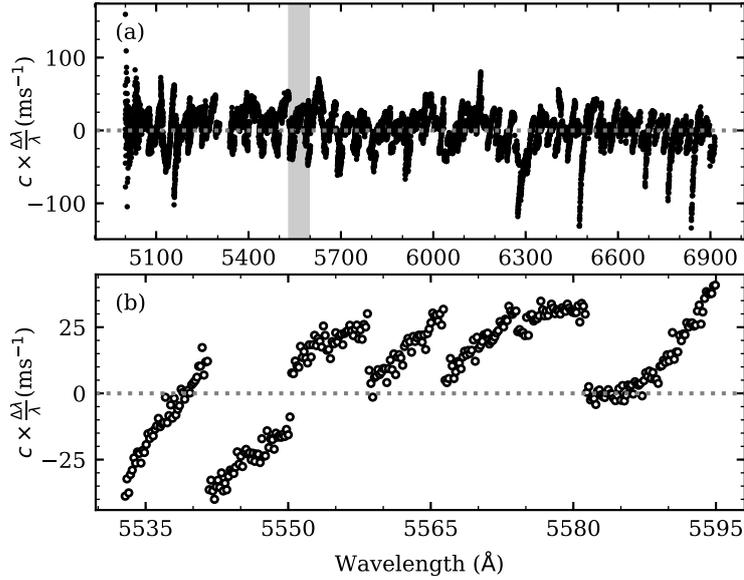
Figure 4.1: Distortions in the ThAr calibration revealed by comparing with LFC lines. Each dot is a single LFC line. Distortion amplitude generally increases at order edges. Clear discontinuities associated with the HARPS detector stitching pattern (Wilken et al. 2010; Molaro et al. 2013b) are seen for all orders. The distortions do not show long-range wavelength dependency, have a $-1.1\,\mathrm{m\,s^{-1}}$ mean offset with respect to the LFC scale, and an overall scatter of $27.9\,\mathrm{m\,s^{-1}}$ rms. Panel (b) shows the region covered by echelle order 110 ($\lambda \approx 5500\,\text{Å}$), also marked by a grey rectangle in panel (a).

separately, producing two spectra from the same observations.

4.3   Modelling procedure

The spectrum shows a damped Lyman-α absorption complex spanning $700\,\mathrm{km\,s^{-1}}$, at redshift $z_{abs} \approx 1.15$ (Reimers et al. 1998), from which numerous previous measurements of α have been made (Quast et al. 2004; Levshakov et al. 2005, 2006; Chand et al. 2006; Molaro et al. 2008; Kotuš et al. 2017). There are at least twenty-six transitions useful for an α measurement in this system. The *Many Multiplet* analysis in this work makes use of transitions covered by the LFC calibration, listed in Table B.1. None of the transitions we use blend with any other systems nor with transitions from the $z = 0.28$ absorption complex identified by Bielby et al. (2017). The LFC-calibrated spectrum showing the Mg II $\lambda 2796$ transition is plotted as a black histogram in Figure 4.4.

We use the most recent set of laboratory wavelength measurements, transition probabilities, oscillator strengths, and isotopic structures for the relevant transitions. These are given in Table B.1. The isotopic abundances were assumed to be solar (Asplund et al. 2009). The sensitivity coefficients that relate atomic line shifts to a change in α are from Dzuba et al. (1999a, 2002); Dzuba & Flambaum (2009). All atomic data is provided as online supplementary material.
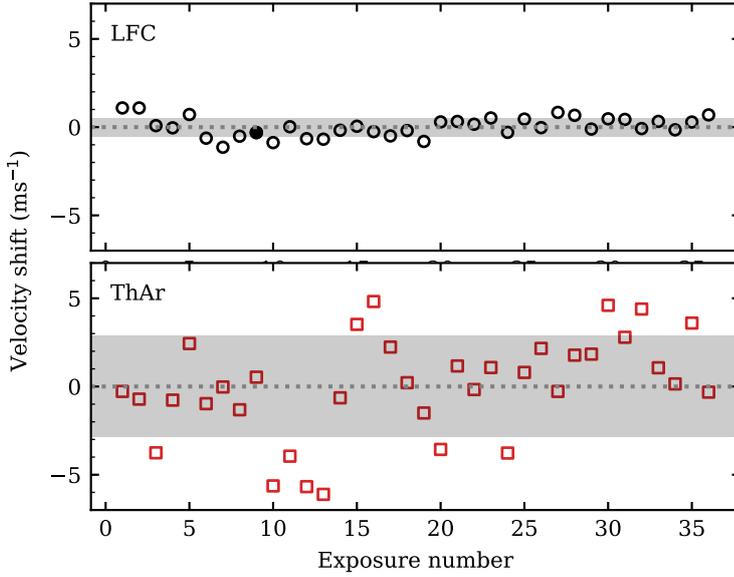
Figure 4.2: LFC calibration was found to drift by $0.52\,\mathrm{m\,s^{-1}}$ rms over the 8 nights of our run (grey shaded region in the top panel), as measured by shifts of LFC lines on the detector. Over the same period, ThAr calibration experienced drifts of $2.87\,\mathrm{m\,s^{-1}}$ rms (grey shaded region in the bottom panel). Shifts in individual LFC (ThAr) exposures are shown as unfilled black circles (red squares) in the top (bottom) panel. The zero line represents the mean value of all points. Shifts were calculated using echelle orders 88 to 121 only (those orders best covered by LFC). The filled black circle in the top panel represents the LFC exposure used for wavelength calibration of all quasar exposures (see text).

We use a fully automated modelling procedure, `AI-VPFIT`, to produce a model of the absorption system (Lee et al. 2020). `AI-VPFIT` is a development of the approach introduced in Bainbridge & Webb (2017) and Bainbridge et al. (2017). Model complexity is increased by placing absorption components ("trial lines") at a random location in the velocity structure and checking if the newly introduced parameters are justified by the data. For the analysis described in the present Chapter, the optimal number of model parameters are derived using the corrected Akaike Infomation Criterion (AICc Akaike 1974; Sugiura 1978). Performance tests using simulated data are described in Lee et al. (2020). Redshifts and b-parameters of components appearing in multiple species are tied during fitting (Section 4.3.2). Column densities are free parameters. We include additional parameters for the unabsorbed continuum level for all transitions and for zero-level adjustment for saturated ones. α is also kept as a free parameter but this has been treated in two different ways (see Sections 4.3.3 and 4.3.4). The basis of the AI algorithm is a genetic process in which a model is built up in 6 well-defined stages. An initial model for the absorption system is generated using a "primary" species, that is, one atomic transition (or atomic species), selected to maximise line strength but avoiding line saturation. Subsequent stages incorporate further atomic species, with appropriately tied parameters, refine parameter errors, check for overfitting, and allow for "interlopers", i.e. unidentified lines from other redshift systems that
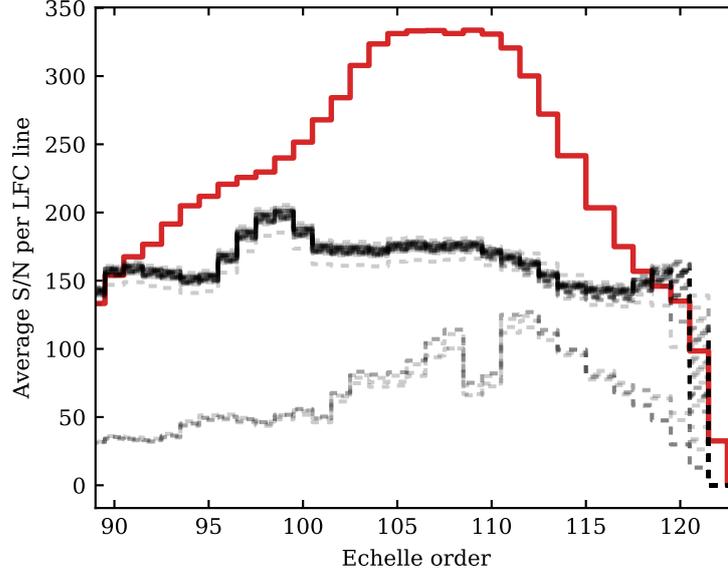
Figure 4.3: Each histogram shows the average S/N per LFC line in a single LFC exposure as a function of echelle order. Five exposures have significantly lower S/N than others (the reason is unknown). A single exposure (full red histogram) reaches S/N of ≈ 330 per LFC line and covers one additional echelle order (122). This exposure provides the most accurate wavelength calibration over the broadest wavelength range. Given the exquisite stability of HARPS (see Figure 4.2), we choose to use this exposure to calibrate all quasar exposures.
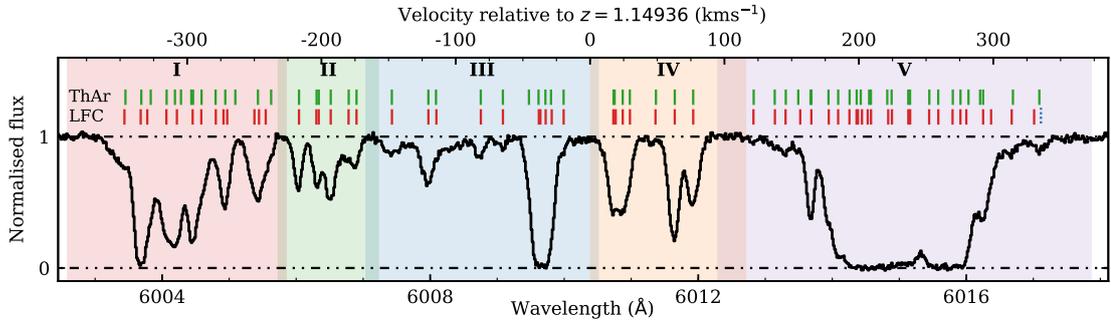
Figure 4.4: The LFC-calibrated spectrum of HE0515−4414 showing the Mg II $\lambda2796$ transition at $z \approx 1.15$ (black histogram). Five coloured areas mark individual regions (denoted by I-V) for which we produce AI-VPFIT models and measure $\Delta\alpha/\alpha$. Small overlaps between neighbouring regions enables better continuum level estimation in each region. The solid red (green) ticks above the data indicate individual metal absorption lines in the best-fit LFC (ThAr) model. The top x-axis shows the velocity with respect to the average redshift of all metal lines in the LFC model, $z = 1.14936$. The blue dotted line at $v \approx 330\,\mathrm{km\,s^{-1}}$ marks the location of an unidentified absorption line in the LFC model. More detailed plots, showing the data, the model, and the residuals can be found in Appendix C (LFC) and Appendix D (ThAr), split by region and transition.
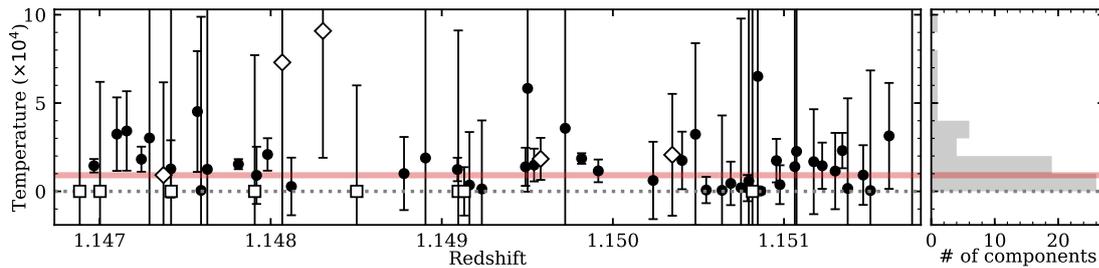
Figure 4.5: Temperatures of individual absorption components. Error bars represent 1σ uncertainties from the covariance matrix at the best-fit solution. Black filled circles show components having both thermal and turbulent broadening. Thermal-only fits are marked by white diamonds. Turbulent-only fits are shown as white squares. The red shaded area marks the weighted average temperature, $T = 9.12 \pm 1.08 \times 10^3$ K. Individual measurements are scattered around this value with $\chi_\nu^2 = 1.042$, where $\nu = 61$. The panel to the right is the histogram of temperatures.

are needed to derive a statistically acceptable overall model.

We produce models for the LFC-calibrated and ThAr-calibrated spectrum independently. All relevant settings during `AI-VPFIT` modelling are kept the same (such as the number of attempts `AI-VPFIT` will make to increase model complexity before proceeding to the following stage, default parameter values for trial lines, line dropping criteria, finite derivative step sizes, etc.), ensuring that all the differences between the final models are a direct consequence of differences in the input data. We refer to models produced from the LFC- and the ThAr-calibrated spectrum as the LFC and the ThAr models, respectively. Figures showing the data, the models, and the residuals for all transitions and all regions are in Appendices C (for LFC) and D (for ThAr).

4.3.1 Instrumental profile

The nominal HARPS instrumental profile has a FWHM of 2.61 km/s. The average b-parameter for individual absorption components in the absorption system analysed in this Chapter is ≈ 5 km/s. The observed quasar lines are thus well-resolved. When matching models to the observed data we must convolve Voigt profiles with the HARPS instrumental profile (IP). To do this we used a Gaussian IP. However, slight departures from a Gaussian have been reported. Moreover, these are found to vary with both flux and position on the detector (Section 3.4). A numerical profile determined directly from HARPS calibration data would thus provide a slightly more accurate IP. However, this was not possible due to insufficient available data[3].

---

[3]The IP is known to be flux-dependent. We do not have a suitable set of LFC exposures to determine the IP as a function of flux level. Data from Chapters 2 and 3 is not useful for this purpose because the HARPS fibres were exchanged since that data were taken, thus changing the IP (Lo Curto et al. 2015)

4.3.2   Temperature as a free parameter

The choice of a line broadening mechanism heavily influences the final model. We found that using turbulent broadening (i.e. not including temperature as a free parameter) impacts significantly on the analysis. For example, imposing a turbulent model forces b to be the same for all species irrespective of atomic mass. If turbulent broadening does not apply in practice, the consequence of the assumption is that additional velocity components are unavoidably included in order to achieve a satisfactory fit to the data. The converse is true - i.e. if a pure thermal model is imposed in the modelling procedure, additional velocity components may also be required to compensate if the model is inappropriate. We explored this by computing models for all three cases, i.e. turbulent, thermal, and mixed-b.

In the mixed-b model, the total line b-parameter is:

$$b^2 = b^2_{\text{turb}} + \frac{2kT}{m}, \tag{4.1}$$

where the right-hand-side terms are the turbulent and thermal contributions, respectively. In the thermal contribution, k is the Boltzmann's constant, T is gas temperature, and m is the appropriate atomic mass. The contribution of each broadening mechanism is determined by the relative widths of transitions of different atomic masses.

The interesting outcome was that a mixed-b model generally requires fewer components and also avoids spurious double-components in line centers (see Lee et al. 2020). Further, once mixed-b models have been derived, it becomes apparent that temperature parameters are genuinely required by the data. The weighted average temperature is $T = 9.12 \pm 1.08 \times 10^3 \, \text{K}$. Individual temperature measurements are shown in Figure 4.5. The temperature is poorly estimated for some components (due to line blending and/or weak lines). Similar results are obtained for the ThAr-calibrated spectrum (not reported). The normalised χ^2 of temperature measurements from 62 velocity components is 1.042, so the data appear to be consistent with a single temperature applying to all components.

4.3.3   Subdividing the absorption complex – 5 regions

We initially divide the system into five regions, I to V (coloured regions in Figure 4.4). Partitioning occurs where the normalised continuum recovers to unity. There is slight overlap between continuum regions in order to optimise continuum estimates. This partitioning has the benefit of simplifying computations and providing independent α measurements, whilst avoiding any potential bias that could occur if unidentified line blending corrupts part of the data.

The five $\Delta\alpha/\alpha$ measurements derived from splitting the absorption system into regions are tabulated in Table 4.2. The quoted uncertainties are derived from the covariance matrix at the best solution. Other relevant statistical information, i.e. the number of free parameters for each region, the number of metal components and their average redshift, and the reduced χ^2 of the model ($\chi^2_\nu = \chi^2/\nu$, where ν is the number of degree of freedom in the model), are also given. The weighted

Table 4.2: Measurements when a single $\Delta\alpha/\alpha$ parameter is used per region (Section 4.3.3), tabulated separately for the LFC and the ThAr models. Column 1 indicates the spectral region (see Figure 4.4). Columns 2 and 3 give the number of metal components (N_c) and the number of free parameters (N_p) in each model. The average redshift of the metal components is in column 4. Columns 5 and 6 give the values of $\Delta\alpha/\alpha$ and their 1σ uncertainties from the best-fit covariance matrix, respectively. Both are in units 10^{-6}. The normalised χ_ν^2 for the fit is in column 7. The lower row gives the average over all five regions for the relevant quantities.

LFC

ID	N_c	N_p	$\langle z \rangle$	$\frac{\Delta\alpha}{\alpha}$	σ_{stat}	χ_ν^2
I	13	125	1.14708	-3.90	4.42	0.9892
II	6	68	1.14788	17.74	4.30	0.9859
III	10	125	1.14870	18.45	15.07	0.9836
IV	7	89	1.14983	-6.39	4.12	0.8595
V	26	267	1.15080	-2.59	3.37	0.9860
All			1.14936	0.94	1.97	

ThAr

ID	N_c	N_p	$\langle z \rangle$	$\frac{\Delta\alpha}{\alpha}$	σ_{stat}	χ_ν^2
I	14	134	1.14735	14.68	4.13	0.9652
II	6	68	1.14788	18.03	4.27	0.9730
III	10	116	1.14872	4.71	15.67	0.9662
IV	7	95	1.14984	-3.04	4.15	0.8343
V	26	245	1.15078	-2.89	3.12	0.9868
All			1.14949	4.82	1.92	

average of $\Delta\alpha/\alpha$ measurements over all five regions for the LFC-calibrated spectrum is $0.94 \pm 1.97 \times 10^{-6}$. The same quantity for the ThAr-calibrated spectrum is $4.82 \pm 1.92 \times 10^{-6}$.

For the LFC-calibrated spectrum, the α measurements from regions I, III, IV, and V are consistent with no variation in α. However, region II produces the seemingly anomalous result of $\Delta\alpha/\alpha = 17.74 \pm 4.30 \times 10^{-6}$ (i.e. a 4.1σ deviation from zero). For the ThAr-calibrated spectrum, regions III, IV, and V are all consistent with $\Delta\alpha/\alpha = 0$. However, regions I and II are not. Region II produces a non-zero result that is similar to the LFC spectrum. Region I also gives a strongly positive result. We discuss ways in which such anomalies can arise in Section 4.3.4.

Table 4.3: The weighted average of the 47 individual $\Delta\alpha/\alpha$ measurements, grouped by region. Column 1 identifies the spectral region. Column 2 gives the total number of $\Delta\alpha/\alpha$ measurements in the region. Column 3 gives the weighted average redshift. Column 4 and 5 give the weighted average $\Delta\alpha/\alpha$ and associated error on the mean in units 1×10^{-6}. Column 6 indicates which absorption components were LTS trimmed (see Figures in Appendices C and D). Columns 7 and 8 give the weighted average $\Delta\alpha/\alpha$ and uncertainty after applying LTS. The lower row provides weighted averages over all 47 measurements (40 after LTS). Superscripts in Column 2 and below the table identify which absorption components were removed by LTS. Where a component lies in a group, the entire group was discarded.

LFC

ID	N	$\langle z \rangle$	$\langle \frac{\Delta\alpha}{\alpha} \rangle$	σ_{stat}	LTS	$\langle \frac{\Delta\alpha}{\alpha} \rangle_{\text{LTS}}$	$\sigma_{stat}^{\text{LTS}}$
I	13	1.14707	-5.40	5.47	ap,au	-5.32	5.53
II	6	1.14784	14.17	4.71	ac	24.45	9.51
III	10	1.14877	19.51	12.48	au	11.53	12.58
IV	4^{a}	1.14984	-6.77	4.25	aj	-0.50	5.00
V	14^{b}	1.15077	-2.98	3.45	ab,ar	-2.38	3.48
All	47	1.14943	-0.18	2.11		-0.27	2.41

ThAr

ID	N	$\langle z \rangle$	$\langle \frac{\Delta\alpha}{\alpha} \rangle$	σ_{stat}	LTS	$\langle \frac{\Delta\alpha}{\alpha} \rangle_{\text{LTS}}$	$\sigma_{stat}^{\text{LTS}}$
I	13^{c}	1.14732	21.22	8.51	aw,ar	16.47	8.55
II	6	1.14784	13.72	4.69	ac	24.67	10.23
III	10	1.14871	0.72	12.60	al	11.92	13.51
IV	4^{d}	1.14984	-3.08	3.89	ae	-5.52	7.02
V	14^{e}	1.15078	-4.10	2.87	as,ax	-3.70	2.90
All	47	1.14957	0.88	1.99		-0.15	2.44

[a] Grouped: (ag,am,aa,al)
[b] Grouped: (ab,aj,al,bm), (bs,bi,ah,ac), (ak,ao), (bq,bc), (ag,an), (ai,as,am), (au,bo)
[c] Grouped: (at,aa)
[d] Grouped: (ag,ak,aa,am)
[e] Grouped: (as,bd,ah,bo), (at,aq), (ao,ay,ab), (ak,bg), (an,bq), (ag,ad), (ai,bc,bi), (ap,bm)

4.3.4 Further subdivision – 47 measurements of α

Instead of dividing the complete absorption complex into five segments and obtaining five measurements of $\Delta\alpha/\alpha$, we can instead solve for the best fit model using one free $\Delta\alpha/\alpha$ parameter for each individual absorption component in the complex. Doing so provides considerably more detail and can identify any $\Delta\alpha/\alpha$ outliers that might "corrupt" an α measurement derived over a whole region or complex. The cost is obviously that the number of free parameters is increased.

To do this we accept the best-fit models for each of the 5 regions and use these parameters as a starting point. However, additional parameters are included to allow α to vary independently for each velocity component. Optimisation is done using `VPFIT`. In other words, we do not recommence the entire `AI-VPFIT` fitting process from scratch. The whole absorption complex (i.e. all five regions illustrated in Figure 4.4) comprises a total of 62 velocity components for the LFC-calibrated spectrum (63 for the ThAr-calibrated). An initial trial fit showed that some badly-blended (and/or weak) velocity components provided only very poor constraints. In those cases we grouped components on small scales, resulting in a total of 47 individual measurements of $\Delta\alpha/\alpha$.

The 47 $\Delta\alpha/\alpha$ measurements obtained this way are shown in Figure 4.6 for both the LFC (top panel) and the ThAr (middle panel) models. The weighted average $\Delta\alpha/\alpha$ across each of the five regions is tabulated in Table 4.3, together with their statistical uncertainties. The results are in good agreement with the results obtained previously, i.e. the weighted average of $\Delta\alpha/\alpha$ measurements within each region falls within 1σ of the results in Section 4.3.3.

Unlike the results obtained in Section 4.3.3, the weighted average over the 47 measurements for the LFC-calibrated spectrum ($\Delta\alpha/\alpha = -0.18 \pm 2.11 \times 10^{-6}$) and for the ThAr-calibrated spectrum ($\Delta\alpha/\alpha = 0.87 \pm 1.99 \times 10^{-6}$) are consistent with each other. As expected, given the large number of free model parameters, there is generally a slight increase in the $\Delta\alpha/\alpha$ error estimates (compare Tables 4.2 and 4.3).

4.3.5 Consistency between 47 α measurements

We now explore the differences between the LFC and ThAr models in more detail. We bin the individual $\Delta\alpha/\alpha$ measurements in redshift bins $\Delta z = 1 \times 10^{-4}$ (the 47 measurements fall into 37 bins) and calculate the weighted average in each. Their differences (LFC minus ThAr) are illustrated in the lower panel of Figure 4.6. The LFC and ThAr measurements agree well everywhere except in region I. Most of the LFC-ThAr differences in region I are located around -2×10^{-5}, with two bins (at the high-redshift end) at more negative values (-1.8 and -7.6×10^{-4}). The top two panels of Figure 4.6 suggest this is caused by velocity structure differences between the LFC and ThAr models. Discarding the points in these two bins and taking the weighted average of the remaining points in region I, we get $-4.90 \pm 5.49 \times 10^{-6}$ for LFC and $15.68 \pm 8.54 \times 10^{-6}$ for ThAr, i.e the two remain inconsistent.

The most significant deviation from zero occurs in region II ($\approx 3\sigma$). Measurements from the LFC-calibrated and the ThAr-calibrated spectrum are in excellent agreement in this region.
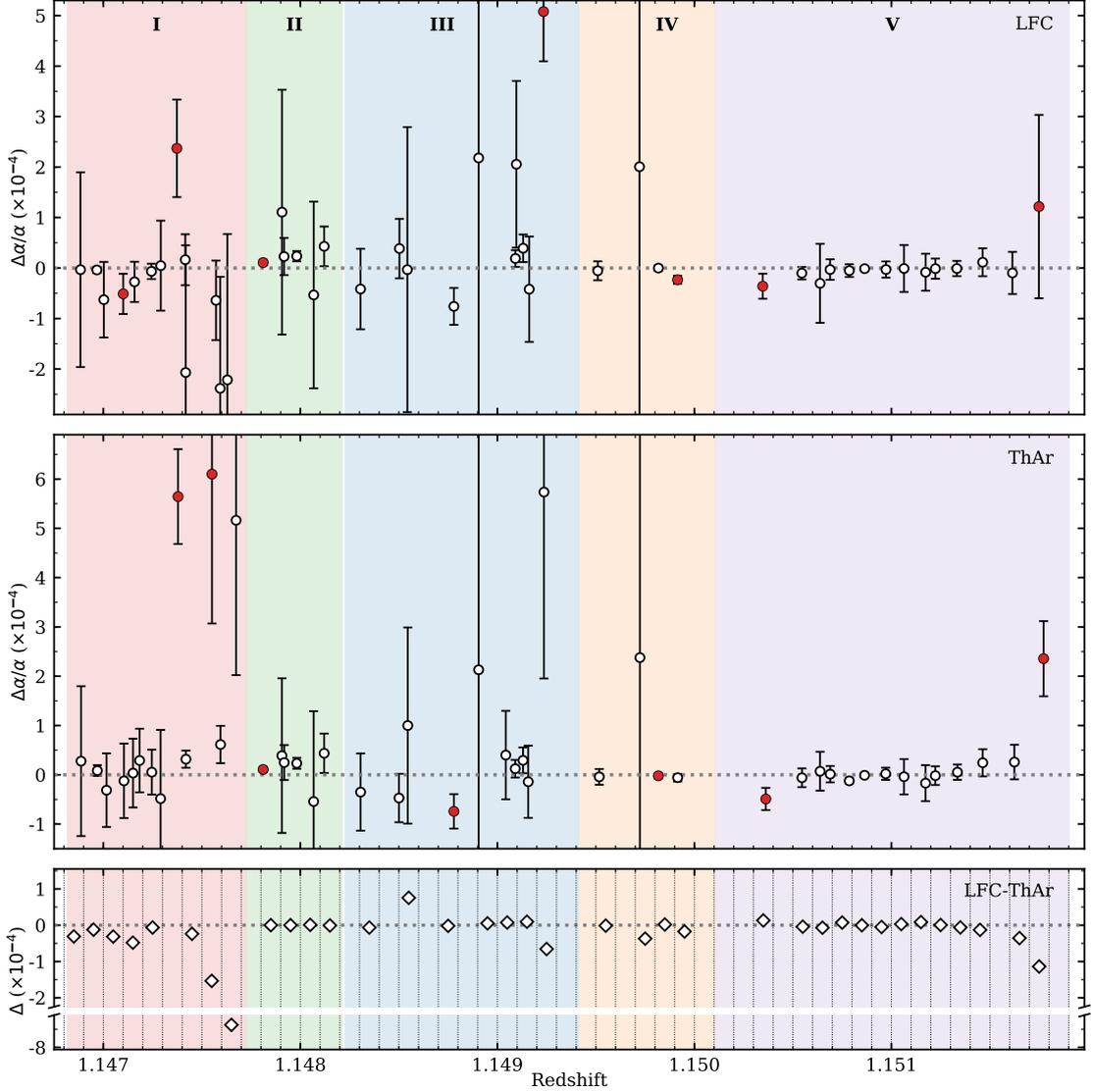
Figure 4.6: The 47 $\Delta\alpha/\alpha$ measurements. Some components were grouped together (Section 4.3.4). Measurements from the LFC-calibrated spectrum are shown in the top panel. ThAr results are shown in the middle panel. The lower panel shows the difference between the LFC and the ThAr measurements after averaging in bins of $\Delta z = 1.5 \times 10^{-6}$ (vertical dotted lines). Weighted averages of points in each region are tabulated in Table 4.3. Filled red circles indicate measurements removed by least trimmed squares. The weighted averages for each region, after discarding those points, are also given in Table 4.3. The data illustrated in the top two panels are available as online supplementary material.
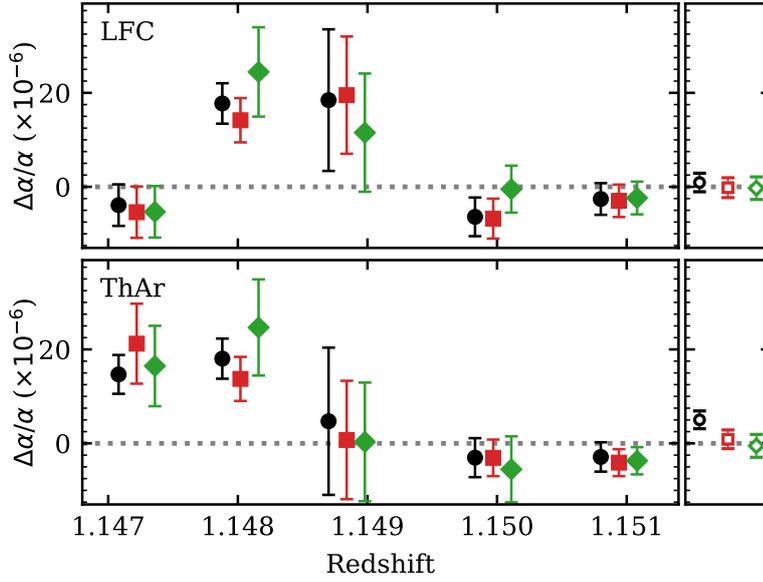
Figure 4.7: Comparing the different approaches to measuring α from the LFC-calibrated spectrum (top) and the ThAr-calibrated spectrum (bottom). Measurements from Section 4.3.3 are shown as black points. The weighted average of the 47 measurements, derived in Section 4.3.4, are shown as red squares. The green diamonds are after outlier removal. Points were offset along the x-axis for clarity. Panels on the right show the weighted average over the entire absorption complex.

A substantial advantage of deriving $\Delta\alpha/\alpha$ measurements for individual absorbing components (or small groups) is that it may help to identify and filter out any possible rogue measurements. The least trimmed squares (LTS) method is frequently used for this and provides a more robust estimate of the mean. We thus apply LTS here, discarding 15% of the data in each region. The discarded components are listed in column 6 of Table 4.3. The weighted averages for each region and the entire sample after LTS trimming are tabulated in columns 7 and 8 of the same table. Interestingly, the consequence of removing the most outlying measurement in region II was to move the region's average towards more positive values (but doubling the error and therefore making the result less significant).

The measured scatter (i.e. the empirical standard deviation) of the remaining 40 LFC-calibrated measurements is $\sigma = 9 \times 10^{-5}$. This is slightly smaller than the average error on individual measurements, $\langle \sigma_{stat} \rangle = 14 \times 10^{-5}$. The scatter of the ThAr-calibrated measurements is $\sigma = 13 \times 10^{-5}$ (the same as the average error).

4.3.6   Consistency with other recent studies

Recently, a detailed study of this same absorption complex was carried out (Kotuš et al. 2017) using spectra from the UVES spectrograph (Dekker et al. 2000) on the Very Large Telescope. Those spectra are higher S/N although the spectral resolution is lower. Comparing with those results reveals good consistency. In this Chapter we split the data into five distinct regions whereas Kotuš et al. (2017) use three. However, combining our regions {I+II} ("left") and {III+IV} ("centre")

enables the comparison (our region V corresponds to Kotuš' "right" region). Using the $\Delta\alpha/\alpha$ results from the LTS trimmed sample in our Table 4.3 (LFC-calibrated) and Kotuš' table 4, and combining all random and systematic errors appropriately, the $\Delta\alpha/\alpha$ solution differ by 1.15, 1.02, and 0.92σ respectively, left to right.

Prior to our study, Kotuš et al. (2017) was the most detailed study. However, several prior analyses also exist (Quast et al. 2004; Levshakov et al. 2005, 2006; Chand et al. 2006; Molaro et al. 2008). All produced results consistent with no change in α, with somewhat larger uncertainties than derived from our analysis or that of Kotuš et al. (2017).

4.4   Results

Figure 4.7 shows the $\Delta\alpha/\alpha$ measurements tabulated in Tables 4.2 and 4.3. The two large panels show the results from the LFC-calibrated (top panel) and the ThAr-calibrated spectrum (bottom panel). The weighted average for the entire $z \approx 1.15$ absorption complex are plotted in the small panels to the right of the main panels. Our main results are summarised as follows:

1. In the analysis presented in Section 4.3.3, we obtain five $\Delta\alpha/\alpha$ measurement (one per spectral region) from the LFC-calibrated and from the ThAr-calibrated spectrum independently. For the LFC-calibrated spectrum, the average over the five regions is consistent with no variation in α. Applying the same methods to the ThAr calibrated spectrum, we obtain a 2.5σ deviation from zero. These results are tabulated in Table 4.2 and plotted as black points in Figure 4.7.

2. Including $\Delta\alpha/\alpha$ as a free parameter for 47 individual absorption components (or appropriately grouped components) allows us to identify regions of data significantly affecting the overall measurement (Section 4.3.4). These results, tabulated in Table 4.3 and plotted as red squares on Figure 4.7, are in excellent agreement with the results from Section 4.3.3.

3. To explore robustness, we apply LTS, removing 15% of the sample in each region, obtaining the results plotted as green diamonds in Figure 4.7. This reduces the total number of measurements to 40. The weighted average over the 40 measurements is consistent with zero for both calibrations: $\Delta\alpha/\alpha = -0.27 \pm 2.41 \times 10^{-6}$ (LFC) and $\Delta\alpha/\alpha = -0.15 \pm 2.44 \times 10^{-6}$ (ThAr).

4. When using the approach in Section 4.3.3, the wavelength scale distortions imparted by the ThAr calibration methods (see Figure 4.1) have a small, but measurable, effect. The same distortions appear to have no effect on the measurements in Section 4.3.4.

5. The HE0515$-$4414 absorption complex modelled spans approximately 700 km/s. If this system represents a line of sight through a cluster of order 1 Mpc across, we can place an upper limit, for the first time, on small-scale α variations, using the empirical scatter in the 40 $\Delta\alpha/\alpha$ measurements.

The upper limit on small-scale α variations over scale-lengths ≈ 25 kpc, is $\approx 9 \times 10^{-5}$ (Section 4.3.5).

6. Averaged over all absorption components, we derive a gas temperature of $T = 9.12 \pm 1.08 \times 10^3$ K (Section 4.3.2). This value is in agreement with the results from Carswell et al. (2012) who found $T = 12 \pm 3 \times 10^3$ K in a quasar absorption system at $z_{abs} = 2.076$. As seen, the new data presented in this Chapter provide a more stringent constraint and also suggest that all individual absorption components are consistent with a single gas temperature.

4.5 Discussion

In this work we have analysed the first LFC-calibrated quasar spectrum. Quasar spectra of similar quality to the one presented here will be routinely produced by the new ESPRESSO spectrograph installed on the Very Large Telescope (Pepe et al. 2014) and by the future HIRES instrument on the Extremely Large Telescope (Maiolino et al. 2013). Both of these instruments have LFCs for wavelength calibration. We hope that the new methods introduced in this Chapter wil be beneficial in analysing future observations.

We have demonstrated that careful modelling procedures play a crucial part in the analysis of high resolution spectroscopic data. Tools such as `AI-VPFIT` eliminate any potential human bias and yield optimal models of the data in a reproducible and objective manner.

Choosing the correct line broadening mechanism, i.e. including temperature parameters for individual components, is important. Models produced assuming an incorrect broadening mechanism tend to generate artificial close blends of absorption lines. Modelling simulated data, based on the HE0515 spectrum used in this work, shows that using an incorrect broadening mechanism biases α measurements (Lee et al. 2020).

Examining the scatter of the five measurements obtained from the LFC-calibrated spectrum in Section 4.3.3, we find they have $\chi^2 = 21.99$ ($\nu = 4$). For a χ^2 distribution with four degrees of freedom, the probability of observing χ^2 values at least this large is $p = 0.02\%$. The five measurements are therefore highly inconsistent with each other. Performing the same for the 47 measurements from Section 4.3.4, we find the LFC-calibrated measurements have a $\chi^2 = 72.86$ ($\nu = 46$, $p = 0.7\%$). We assume this is not caused by small-scale spatial variations in α across the redshift range covered by the absorption system. After LTS trimming, the scatter in the remaining 40 measurements are consistent with their individually estimated errors ($\chi^2 = 23.76$, $\nu = 39$, $p = 97\%$). Similar results are obtained for the ThAr spectrum.

The analysis presented here is based on the assumption of solar relative isotopic abundances. Significant deviations from solar values translate to large shifts in $\Delta\alpha/\alpha$ (Webb et al. 1999; Ashenfelter et al. 2004a,b; Fenner et al. 2005; Berengut et al. 2012; Webb et al. 2014). Very approximately, when simultaneously modelling Mg II and Fe II, the measured $\Delta\alpha/\alpha$ may shift towards negative values by as much

as $\approx 5 \times 10^{-6}$ for 100% ^{24}Mg and by the same amount in the positive direction for 100% $^{25+26}$Mg. A discussion as to the validity of the solar isotopic assumption is deferred to a subsequent work.

For these particular observations, LFC calibration methods have not yielded significantly different $\Delta\alpha/\alpha$ measurements compared to the ThAr methods. The probable reason for this lies in the fact that we have combined a large number of individual exposures to form a final co-added spectrum. Due to different barycentric velocities for each observation, the position of relevant transitions falls differently with respect to the complicated distortion pattern each time, effectively smearing it out. This is less likely to occur for more efficient spectrographs, such as ESPRESSO and HIRES.

Chapter 5

Conclusions

Fundamental constants are expected to vary in many generalizations of General Relativity positing the existence of additional scalar fields or extra spatial dimensions, as discussed in detail in Chapter 1. Their existence could provide a physical explanation for the observed effects of dark matter and dark energy, as well as provide a mathematical framework for a Grand Unified Theory. In many of these models, the fine structure constant, α, is elevated from a pure numerical constant into a dynamical field which is allowed to evolve in time and/or space, or couples to the Standard Model degrees of freedom. Finding evidence for departure of α from its terrestrial value would therefore lead to a breakthrough in fundamental physics.

Spectroscopic measurements in high-resolution quasar spectra provide some of the tightest constraints on the variation of α on cosmological timescales and distances. Precise α measurements from quasar spectra require a way to meaningfully compare astronomical and laboratory spectra, on the same wavelength scale and with extreme accuracy. This is best achieved by providing an extremely accurate and stable wavelength calibration reference, such as the laser frequency comb (Udem et al. 2002; Hänsch 2006; Steinmetz et al. 2008).

Preliminary analysis of the data collected during the development of the HARPS LFC demonstrated the exquisite performance of the LFC in tracking spectrograph drifts and wavelength calibration (Wilken et al. 2010, 2012). This presented a major advance for precise spectroscopic measurements, as the large number of LFC lines, together with their extremely well known frequencies meant that varying constant studies could reach unprecedented precision. The LFC technology is still young but initial results, some of which are presented in this thesis, show that LFCs are indispensable for all high precision spectroscopic measurements.

Extremely high spectroscopic precision and accuracy is required not only for varying constant studies, but also for detecting Earth-mass exoplanets around Solar-mass stars, measuring the accelerations of stars in the Galaxy (Ravi et al. 2019a; Silverwood & Easther 2019), and measuring the small, systematic redshift drift of objects in the Hubble flow, (Sandage 1962; Loeb 1998; Darling 2012; Liske 2014). These projects are major science drivers for the construction of the Extremely Large Telescope (ELT, Tamai et al. 2018) and its future high-resolution spectrograph HIRES (Hook 2009; ESO ELT team 2010, 2011; Marconi et al. 2016).

Other 30-m class telescopes, the Thirty Meter Telescope (Simard et al. 2016) and the Giant Magellan Telescope (Fanson et al. 2018), will also have LFC-calibrated, high-resolution spectrographs for extremely precise measurements. Unlike varying constant studies, the other aforementioned studies require repeated, multi-epoch observations of the same objects spanning over periods of years or decades. This imposes an additional requirement that the wavelength calibration must be stable at the $1 \, \mathrm{cm \, s^{-1}}$ level over the period of the measurement.

5.1 Summary

5.1.1 Spectroscopic precision at the $1 \, \mathrm{cm \, s^{-1}}$ level

In April 2015, HARPS became the first high-resolution astronomical spectrograph to be permanently equipped with an LFC (Probst et al. 2015a). Several hundred LFC spectra were taken at that time, including those using a second, completely independent LFC which was later installed on the FOCES instrument (Pfeiffer et al. 1998; Brucalassi et al. 2016). The data was collected by the team leading the development and construction of the HARPS LFC. Data analysis, led by Rafael A. Probst, aimed to quantify the performance of the HARPS LFC, and investigate a number of effects pertaining to the LFC hardware and the configuration of the system. The HARPS LFC was shown to provide $1 \, \mathrm{cm \, s^{-1}}$ precision over a wavelength range of approximately 220 nm and over a period of ≈ 2 hours. This was substantiated through a comparison with the second LFC, implying that there are no systematic effects limiting the performance of LFCs at the desired level of precision, over such a short period of time. The same analysis, however, revealed a $\approx 50 \, \mathrm{cm \, s^{-1}}$ shift between the absolute wavelength calibrations provided by the two LFCs. An independent wavelength calibration procedure, performed by me, confirmed the existence and the magnitude of this shift (Section 2.3.4). These results were reported in a paper titled "A crucial test for astronomical spectrograph calibration with frequency combs", published in *Nature Astronomy* (Probst et al. 2020), and are presented in Chapter 2.

5.1.2 Significant improvements to spectroscopic calibration

Following this, I performed a more detailed analysis of the same collection of spectra, focusing on the development of optimal wavelength calibration methods for astronomical spectrographs (Chapter 3). Various effects, such as the wavelength dependent modulation of the LFC signal, stitching of the detector, and flux dependent charge transfer effects, were examined in significantly more detail than before. The results of this analysis were illuminating. Firstly, the LFC background, seen as a non-vanishing continuum between individual LFC lines, was shown to arise from the same physical process as the LFC spectrum (Milaković et al. 2017), consistent with the idea that it is seeded by the amplified spontaneous emission (Probst et al. 2020). Secondly, LFCs have been shown to be excellent tools for detector characterization: LFC lines can measure the sizes of anomalous pixels

coinciding with the detector stitches and can track small line centroid shifts due to flux-dependent effects. Thirdly, using an objective statistical criterion (AICc), I have demonstrated that optimal wavelength calibration is achieved when segmented polynomials are used (polynomial edges follow the division imparted by the detector stitching). Wavelength calibration performed in this way results in distortions smaller than $\lesssim 3\,\mathrm{m\,s^{-1}}$, similar to the statistical uncertainty on LFC line positions (Section 3.5).

Using single polynomials for wavelength calibration was shown to leave a pattern of correlated wavelength calibration residuals with amplitudes as large as $4\,\mathrm{m\,s^{-1}}$. The pattern persists even when high order polynomials are used and the impact of detector stitching is accounted for (Section 3.5.2). The expected effect of this pattern is to introduce distortions which may emulate periodicity (and therefore lead to spurious planet detections) or manifest as variation in fundamental constants, as well as corrupt redshift drift measurements. Furthermore, the analysis demonstrated that the $\approx 50\,\mathrm{cm\,s^{-1}}$ shift between the zero-levels of two LFCs's wavelength calibration, can only partly be attributed to flux-dependent effects (Figure 3.15).

Finally, once shifts of the sort described above have been properly accounted for, then I have shown that the precision of spectroscopic measurements is as good as $1\,\mathrm{cm\,s^{-1}}$. The achievable precision was also shown to benefit from applying the corrections for flux-dependent effects (Figure 3.16). These results were reported in a paper titled "Precision and consistency of astrocombs", published in the Monthly Notices of the Royal Astronomical Society (Milaković et al. 2020).

5.1.3   The first LFC-calibrated quasar spectrum

The first application of LFC wavelength calibration to quasar spectra pertain to the new HARPS observations of the quasar HE0515−4414 (Chapter 4). I was the principal applicant for a telescope proposal to use HARPS and observe HE0515 in ESO period 102. The proposal was ranked in the top quartile of all observations in this period and was awarded 8 nights of observing time. The goal of the proposal was to measure $\Delta\alpha/\alpha$ in an absorption system at $z_{abs} = 1.15$ seen towards this quasar, using LFC for wavelength calibration. Additionally, ThAr calibration frames were also taken and were used to produce a second spectrum from the same observations. The LFC-calibrated spectrum is virtually free from wavelength distortions, and its calibration residuals are six times smaller than those of ThAr (Section 4.2.2).

We measured variations in α in the $z_{abs} = 1.15$ absorption complex using the Many Multiplet method (Section 1.2.3), but this was treated in two different ways (Section 4.3). The spectrum was first divided into five spectral regions and `AI-VPFIT` models were produced. Initially, a single free parameter per region was used to measure $\Delta\alpha/\alpha$, yielding a total of ten measurements (one for each of the five regions and the two calibrations). The average of the five LFC-calibrated α measurements was found to be consistent with zero: $\Delta\alpha/\alpha = 0.94 \pm 1.97 \times 10^{-6}$. The average of the ThAr-calibrated measurements, however, produced a result which deviates from zero with $\approx 2\sigma$ significance: $\Delta\alpha/\alpha = 4.82 \pm 1.92 \times 10^{-6}$.

The high quality of the data allowed for the introduction of additional parameters for α, one for each individual absorption component (or appropriately grouped components). This yielded 47 measurements in this absorption complex. Measuring α over such small scales allowed us to identify a small number of anomalous values. We removed 15% of the data using least trimmed squares, a statistically robust method for outlier removal. The average of the 40 remaining measurements is consistent with zero for both calibrations: $\Delta\alpha/\alpha = -0.27 \pm 2.41 \times 10^{-6}$ for LFC and $\Delta\alpha/\alpha = -0.15 \pm 2.44 \times 10^{-6}$ for ThAr. The prediction of the dipole reported by King et al. (2012) (Section 1.2.4) for the sky position of HE0515 is $\Delta\alpha/\alpha = 1.2 \times 10^{-6}$. Therefore the HE0515 measurement reported here does not test the existence of the dipole.

The observed scatter in the 40 measurements allowed us to place the first ever upper limit on small-scale variations of α. The observed scatter in $\Delta\alpha/\alpha$ is $\sigma_{\delta\alpha/\alpha} \lesssim 9 \times 10^{-5}$ over $\approx 20\,\mathrm{km\,s^{-1}}$ scales. If the entire absorption complex arises in a galaxy cluster of ≈ 1 Mpc in size, this would correspond to variations over scale-lengths of ≈ 25kpc (Section 4.3.4).

5.1.4 Artificial Intelligence methods for varying constant studies

During my studies, I was involved in the development of a method for producing objective, reliable, and reproducible models of quasar absorption systems: `AI-VPFIT`. A short outline of `AI-VPFIT` is given in Section 1.3.1, whilst a more comprehensive description has been submitted for publication to the Monthly Notices of the Royal Astronomical Society (Lee et al. 2020). The scope of my involvement included testing `AI-VPFIT` at various stages of its development and providing feedback aimed to optimize its performance. I have also contributed to the discussions on particular improvements which were then implemented. The high quality of the HARPS spectrum of HE0515 is well suited for testing different approaches to model-building, as it contains a number of spectral features of varying complexity. These tests gave rise to several important conclusions which influenced the current version of `AI-VPFIT`.

For example, a comparison of models derived using different line broadening mechanisms, produced by me on real data, proved important. We have seen that, when free parameters for gas temperature were not included, `AI-VPFIT` introduced spurious double components (i.e. two coincidental metal absorption lines) to obtain statistically satisfactory fits to the data. Although we cannot exclude the possibility that such double components really exist, the frequency with which they appeared in the model suggested they were likely to be model artefacts. This inspired an investigation into the impact of the choice of the line broadening mechanism on $\Delta\alpha/\alpha$ using simulated data. The outcome of this investigation conclusively demonstrated that high quality data requires the inclusion of temperature parameters to produce optimal results, and that not doing so biases $\Delta\alpha/\alpha$ measurements. This work has resulted in two papers being submitted for publication to the Monthly Notices of the Royal Astronomical Society. The first one, titled "Artificial intelligence and quasar absorption system modelling; application
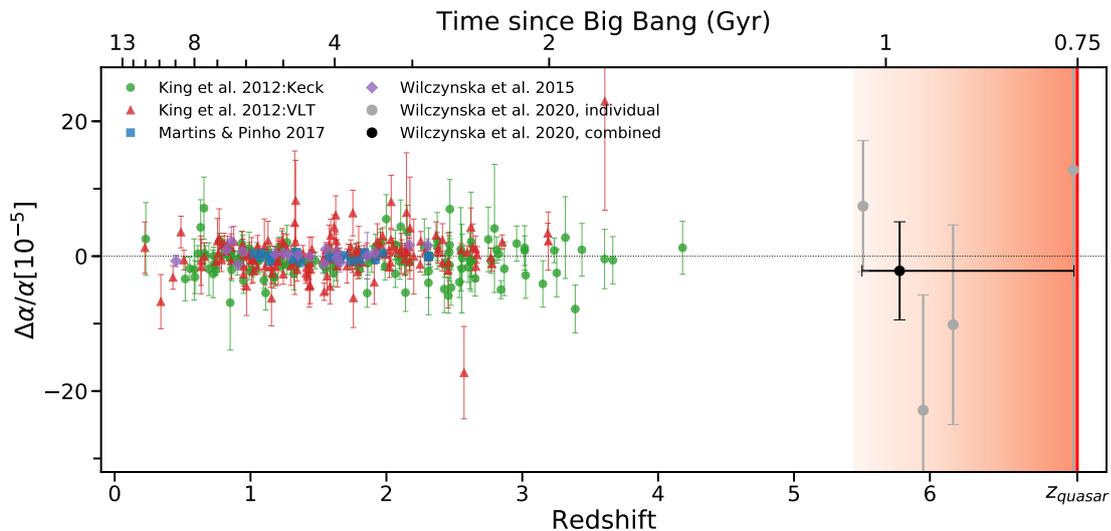
Figure 5.1: Direct measurements of $\Delta\alpha/\alpha$, taken from quasar absorption measurements King et al. (2012); Wilczynska et al. (2015); Martins (2017). Where measurements reported in Wilczynska et al. (2015) were re-analyses of the same systems from King et al. (2012), the former were used. Error bars include systematic contributions (although we note the heterogeneous nature of this combined dataset and point out that systematic errors were not all estimated in a consistent manner so error bars are not necessarily directly comparable in all cases). The point in black at $z = 5.87$ illustrates the weighted mean of the four measurements described in Wilczynska et al. (2020). Its horizontal bar indicates the redshift range spanned by those four measurements. The red shaded area shows the redshift range from the quasar emission redshift ($z_{em} = 7.085$) down to the lowest possible redshift for a $\Delta\alpha/\alpha$ measurement ($z_{abs} = 5.443$) assuming we retain the lowest rest-wavelength anchor line, Si II $\lambda 1526$. Figure taken from Wilczynska et al. (2020).

to fundamental constants at high redshift" is available as a pre-print (Lee et al. 2020). The second paper, titled "Getting the model right; an information criterion for spectroscopy" will appear shortly.

5.1.5 Four direct α measurements at $z \approx 6$

Another project I was involved with during my PhD studies is the measurement of α from X-SHOOTER (Vernet et al. 2011) observations of a $z = 7.085$ quasar J1120+0641. The data, described in Bosman et al. (2017), revealed four absorption systems with transitions suitable for measuring α. The four systems lie at redshifts $z_{abs} = 7.059, 6.171, 5.951$, and 5.507. This particular analysis used GVPFIT (Bainbridge & Webb 2017; Bainbridge et al. 2017) to model the spectrum and obtain a measurement of $\Delta\alpha/\alpha$ in each absorption system, shown as grey dots in Figure 5.1. Their exact values, ordered from the highest to the lowest redshift system, are: $\Delta\alpha/\alpha = (12.79 \pm 48.66 \pm 19.74) \times 10^{-5}$, $(-10.16 \pm 14.80 \pm 0.42) \times 10^{-5}$,

$(-22.85 \pm 17.11 \pm 0.32) \times 10^{-5}$, and $(7.42 \pm 9.60 \pm 1.52) \times 10^{-5}$. The first error term in the measurements above refers to the statistical uncertainty from the modelling procedure and the second term refers to the systematic uncertainty due to any possible wavelength scale distortions present in the X-SHOOTER spectra (although none were detected). The weighted mean of the four measurements is $\Delta\alpha/\alpha = -2.18 \pm 7.27 \times 10^{-5}$ and is shown as a black dot on Figure 5.1. These are the first direct measurements of $\Delta\alpha/\alpha$ from quasar spectra at such high redshift. The full analysis is reported in an article titled "Four direct measurements of the fine-structure constant 13 billion years ago", published with *Science Advances* (Wilczynska et al. 2020). I was involved with data analysis and have produced Figure 5.1.

5.2   Future work

5.2.1   Understanding fundamental limits to spectroscopic accuracy
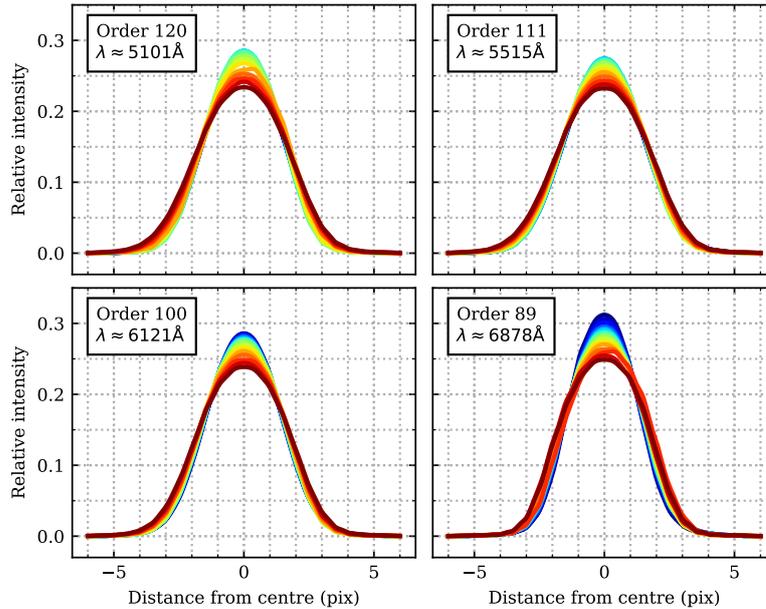
Figure 5.2: Each panel shows the HARPS instrumental profile at 16 different locations in a single echelle order. Different coloured profiles represent different locations, with the bluest (reddest) profile corresponding to the blue (red) edge of the order. The profile is seen to change shape within and across echelle orders.

Significant improvements to wavelength calibration accuracy are required to maintain the $1\,\mathrm{cm\,s^{-1}}$ precision over arbitrarily long periods, in order to enable the projects outlined at the beginning of this Chapter. Several systematic effects need to be overcome before this goal is achieved, the most important of which is a $\approx 50\,\mathrm{cm\,s^{-1}}$ offset in the absolute wavelength calibration when one LFC is replaced by another. The implication of this effect is that, even when using the
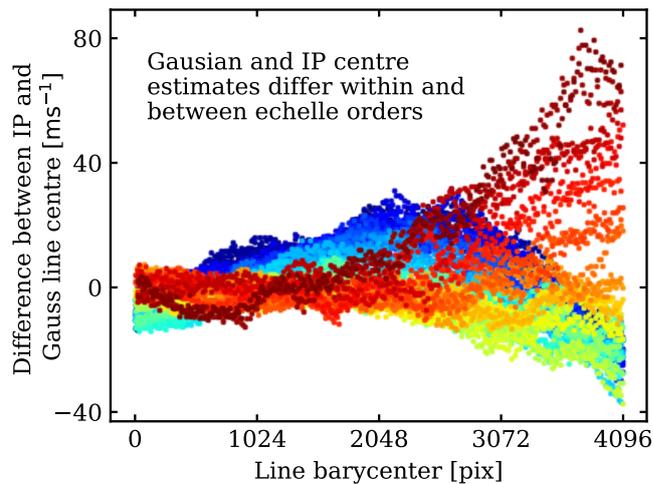
Figure 5.3: Replacing the Gaussian approximation with the empirical instrumental profile (Figure 5.2) results in line centre estimates which differ by as much as $80\,\mathrm{m\,s^{-1}}$. The data plotted correspond to ≈ 10000 LFC lines. Different colours represent different echelle orders.

best currently available methods and LFCs, the $1\,\mathrm{cm\,s^{-1}}$ stability can only be maintained for a short period of time. The $1\,\mathrm{cm\,s^{-1}}$ precision over significantly longer periods of time can, in principle, be achieved by introducing redundancy, but this requires substantially higher cost and greater complexity (this point is discussed in Section 4.5). Although the sources of this $\approx 50\,\mathrm{cm\,s^{-1}}$ offset are not yet well understood, it is reasonable to expect that further work will help to identify what limits the accuracy of wavelengths measurements in spectroscopic data and where improvements can be made.

One improvement that can already be made is to replace the Gaussian approximation of the LFC line shape with the empirical instrumental profile (IP). The high quality and density of LFC lines allows for modelling the spectrograph's IP, tracing its variation across the detector and identifying other effects such as dependence on intensity (Figure 5.2). Preliminary results on the HARPS instrument, show that centroids of lines derived using the empirical IP differ by as much as $80\,\mathrm{m\,s^{-1}}$ from those derived using a Gaussian profile approximation. As illustrated in Figure 5.3, this effect changes within and between echelle orders.

5.2.2 Measuring α with ESPRESSO

One of the science drivers for the newly commissioned Echelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations (ESPRESSO, Pepe et al. 2014) is to further explore indications of the α-dipole reported by Webb et al. (2011); King et al. (2012) and others (see Section 1.2.4). ESPRESSO is a high-resolution spectrograph for extremely precise spectroscopic measurements, installed at the coudé room of the Very Large Telescope. It was designed to collect the light from of any one of the VLT Unit Telescopes (UT) or to combine the light of several UTs and increase efficiency. Alongside the standard ThAr lamps and novel
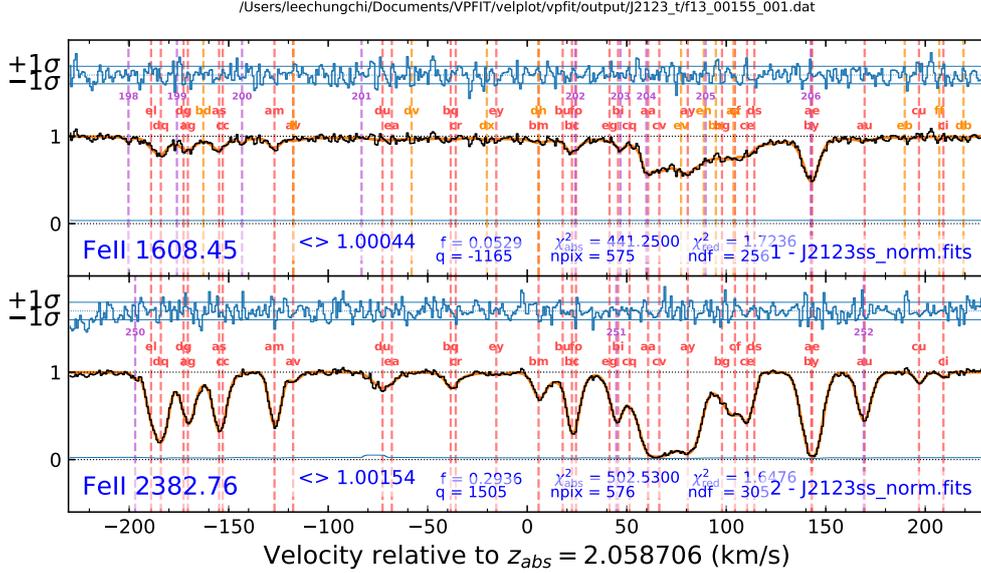
Figure 5.4: Details of the ESPRESSO spectrum of a sub-damped Lyman-α system at $z_{abs} = 2.06$ towards J2123-0050. Many transitions suitable for measuring $\Delta\alpha/\alpha$ are detected in an absorption complex spanning about $400\,\mathrm{km\,s^{-1}}$. The Fe transitions falling within the spectral range covered by our 4UT ESPRESSO data include Fe II $\lambda 1608$, Fe II $\lambda 2344$, $\lambda 2374$, $\lambda 2383$, $\lambda 2600$. The figure illustrates two of these, 1608 and 2383. This figure appears in the ESO Messenger article describing the ESPRESSO science verification (Leibundgut et al., ESO Messenger 181, in press).

Fabry-Pérot Interferometers (FP), ESPRESSO is also equipped with an LFC for wavelength calibration.

My co-investigators and myself obtained 5h of ESPRESSO observations of the quasar J2123-0050 during the science verification of its 4UT observing mode, in which all four UTs observed the quasar simultaneously. Previous observations of this quasar, using VLT/UVES and Keck/HIRES, identified eighteen transitions useful for measuring α in a system at $z = 2.06$. The ESPRESSO data are of high quality and calibrated using an LFC. Preliminary results of solving for α indicate an overall error budget around the 10^{-6} level on $\Delta\alpha/\alpha$, making this one of the most precise measurements to date. This work is featured in the ESO Messenger article describing the ESPRESSO science verification (Leibundgut et al., ESO Messenger 181, in press).

5.2.3   My involvement with the ESPRESSO collaboration

I am proud that my expertise in wavelength calibration was recognised by the ESPRESSO working group on varying constant studies, and resulted in an invitation for collaboration. The ESPRESSO team has been awarded guaranteed time to observe a sample of quasars which would then be used to measure the values of fundamental constants (PI: Paolo Molaro, Osservatorio Astronomico di Trieste).

However, several systematic effects pertaining to the wavelength calibration and the intermittent operation of the LFC on ESPRESSO prevent full exploitation of this high quality data. I was invited to apply my expertise on wavelength calibration methods to ESPRESSO LFC data, aiming to investigate the following:

1. Given the intermittent operation of the LFC on ESPRESSO, it would be useful to wavelength calibrate science spectra using LFC frames taken significantly earlier/later in time. The feasibility of this process relies on the instrument stability being sufficiently good, but this has not yet been characterized. This investigation would involve using calibration frames taken at different epochs, identifying any systematic effects influencing the said stability, and removing them.

2. A comparison of LFC and ThAr wavelength calibrations on ESPRESSO revealed that the two disagree at the $\approx 10\,\mathrm{m\,s^{-1}}$ level (Tobias Schmidt, private communication). The origin and the temporal characteristics of these distortions is unknown, but their removal is crucial for exoplanet and fundamental constant studies as they may emulate the signal sought in the data. Preliminary results show that the distortions are different for the two fibres and the two slices (ESPRESSO uses an image slicer and the light from each of its two fibres is projected twice on the detector). It is also important to understand which calibration (LFC or ThAr) is more strongly affected by systematic effects. Studying ThAr spectra calibrated in both ways, and comparing the measured wavelengths of ThAr lines to their laboratory values, could offer some clues in this respect.

3. The LFC does not cover the full wavelength range of ESPRESSO, requiring the use of an alternative calibration source in the region not covered by the LFC. Currently, this is done by combining the ThAr with the Fabry-Pérot Interferometer spectra. Similarly to the LFC, the FP produces a series of densely spaced emission lines which can be used to identify wavelength distortions and track spectrograph drifts. However, unlike LFC, FP is not an absolute calibrator and its accuracy is limited by the accuracy of the ThAr calibration. One must therefore understand how to best calibrate the spectrum in regions where LFC calibration is not possible by complementing it with ThAr/FP calibration frames. In the case the distortions described in point (2) above originate in the ThAr/FP calibrations, we need to understand how these behave in regions not covered by the LFC. A comparison of a solar spectrum obtained with ESPRESSO with the best available Fourier Transform Spectrometer solar spectrum could possibly provide some clues on the distortions in these regions (a study along those lines was done by Molaro et al. 2013b).

The results from the investigations above would then be applied to ESPRESSO quasar spectra with the purpose of measuring the change in α. It would be interesting to analyse the combined HARPS and ESPRESSO observations of HE0515, as the combined dataset should increase the total S/N by a factor of two.

5.2.4   Preparing for the Extremely Large Telescope

The success of several forthcoming large optical observatories, such as the Extremely Large Telescope (Tamai et al. 2018), the Thirty Meter Telescope (Simard et al. 2016), and the Giant Magellan Telescope (Fanson et al. 2018), depends on the development of new technologies, observing strategies, advanced data analysis methods (including observation simulators), as well as on the successful removal of various systematics effects. This can already be done using currently available instrumentation and methods.

This is why ESO invited a large number of world experts, specialists in various key aspects important for the success of the ELT project, to work together in anticipation of the ELT first light in mid 2020's. The effort is divided into themes, each covered by one or more working groups. Topics include detector characterization, point-spread-function reconstruction, advanced instrument simulations of observations, cataloguing stars suitable for use with adaptive optics systems, telluric correction methods, impact of weather on telescope scheduling, development of new photometric and wavelength calibration methods, obtaining new laboratory line measurements, and others.

The most demanding projects planned with the ELT, as far as the required spectroscopic precision and accuracy is concerned, are studies of fundamental constants, discovery of planets similar to Earth, and the redshift drift measurement. As demonstrated in this thesis, LFCs will be indispensable for all spectroscopic studies aiming for the highest possible precision. My expertise in this field is why I was invited (by Michele Cirasuolo, ELT Programme Scientist) to join two working groups: the "Line calibration" and the "Detector effects characterization". I am the most junior member and the only PhD student across all working groups.

The "Line Calibration" working group aims to produce a comprehensive list of recent laboratory measurements of transitions relevant for varying constant studies, high-redshift measurements of the CMB temperature, and primordial light-element abundances; making recommendations when improvements are necessary. The group consists of 20 members with expertise in laboratory measurements of atomic transitions, astronomical studies of varying constants and quasar absorption systems, atomic theory experts, laser system experts, experts in data analysis methods, and others. The frequency of group meetings so far has been one meeting approximately every two months, with a single joint meeting with other working groups in May 2020. The group is chaired by Carlos J. A. P. Martins (University of Porto).

The goal of the second ELT working group I am a member of, "Detector effects characterization" is to identify a set of detector effects which may limit measurement precision from ELT data. This ELT working group will also provide information to other teams who also carry out research to reduce the impact of these effects. Doing these things needs close collaboration between all interested parties, including both astronomers and engineers. The aims are to characterize and quantify the detector effects and to evaluate the impact of any such systematics on the science. The chair of this group is Elizabeth George (ESO, Detector Systems Group).

Appendix A

Mode identification issue with the 18 GHz LFC

During our analysis in Chapter 3, we noticed a systematic velocity shift of approximately $45\,\mathrm{m\,s^{-1}}$ between the LFC1 and LFC2 wavelength calibrations and between LFC1 and attached ThAr calibrations. A shift of this magnitude cannot be explained by spectrograph drifts, which made us suspect we have misidentified an astrocomb line. Coffinet et al. (2019) used the same dataset in their analysis and noted that the offset frequency of LFC1 was probably different by $100\,\mathrm{MHz}$ from that reported in the observing log. It is likely that the change in the offset frequency was made by the operator and not noted in the system. In what follows, we provide a definitive proof of the shift's existence and measure its value.

We return to the mode identification algorithm and perform an exercise to verify that we are indeed assigning modes correctly. To this end, we use LFC1, LFC2, and ThAr spectra taken within a short time period from each other to ensure spectrograph drifts are small. We select several echelle orders in the ThAr exposure and wavelength calibrate them ourselves by fitting a third order polynomial through several ThAr lines for which wavelengths are tabulated in the Palmer & Engleman (1983) atlas. We then identify, by eye, LFC1 and LFC2 lines that coincide with ThAr lines to within 1 pixel on the detector. We now take those particular astrocomb lines (one LFC1 and one LFC2 per order) and estimate their mode number from the wavelength of the coinciding ThAr line. Knowing the modes of the astrocomb lines, we use them to wavelength calibrate each echelle order as described in Section 3.5, but we change the frequency of LFC1 lines during fitting in steps of $20\,\mathrm{MHz}$ in the range $\pm440\,\mathrm{MHz}$. For each frequency step, we calculate the average velocity shift between ThAr, LFC1, and LFC2 wavelength calibrations using the pixel shift method (see Section 3.6 for details). We find that the LFC1 offset frequency needs to be shifted by $\Delta f_o = 100\,\mathrm{MHz} \pm n \times 250\,\mathrm{MHz}$ (with n an integer number) in all echelle orders examined to achieve agreement between all three calibrations (see Figure A.1). The frequency shifts are degenerate with $250\,\mathrm{MHz}$, which is the repetition frequency of the LFC1 spectrum before mode filtering Chapter 2.

Assuming the smallest possible shift, we change the LFC1 offset frequency by $\Delta f_o = 100\,\mathrm{MHz}$ during mode identification step of our analysis, i.e. Equa-
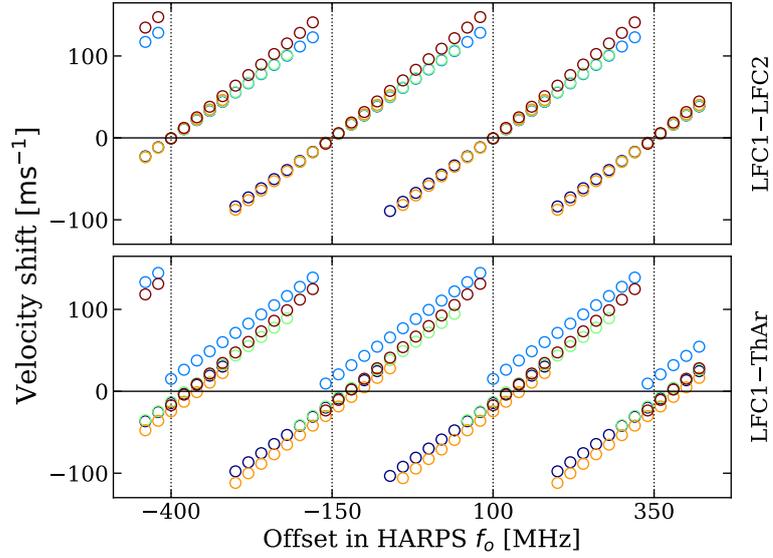
Figure A.1: We manually change the offset frequency of LFC1 in steps of 20 MHz and calculate the average velocity shift with respect to LFC2 (top panel) and ThAr (bottom panel) wavelength calibrations in several echelle orders (different colours). We find that LFC1 offset frequency needs to be shifted by 100 MHz from what was reported in the observing log in order to be consistent with the ThAr and LFC2 calibrations. This is probably due to logging error.

tion (3.1). The LFC1 offset frequency in Table 3.1 (5.7 GHz) already reflects this change.

Appendix B

Atomic data

Table B.1:   Atomic species and transitions, with isotopic structure, used in this analysis. Terrestrial isotopic relative abundances are assumed. Column 4 (λ_0) is rest-frame wavelength. Column 5 (f) is oscillator strength or relative abundance (%). The latter are from Rosman & Taylor (1998). Column 6 (Γ) is the sum of the spontaneous emission rates. Column 7 (q) gives the sensitivity coefficients to a change in the fine structure constant α. Citations to original measurement papers are given at the foot of the table. An atomic data compilation including the data in this table is given in Murphy & Berengut (2014).

Fe II	2344	55.845	$2344.212747(76)^{a,b}$	0.114	$2.680^{c,d}$	$1375^{e,f}(300)$
		58	2344.2113616^{f}	0.282%		
		57	2344.2120103^{f}	2.119%		
		56	2344.2126822^{f}	91.754%		
		54	2344.2141007^{f}	5.845%		
	2374	55.845	$2374.460064(78)^{a,b}$	0.03130	$3.090^{c,g}$	$1625^{e,f}(100)$
		58	2374.4582998^{f}	0.282%		
		57	2374.4591258^{f}	2.119%		
		56	2374.4599813^{f}	91.754%		
		54	2374.4617873^{f}	5.845%		

(the table continues on the next page)

Ion	Tran	A	λ_0 (Å)	f or %	Γ ($10^8\,\mathrm{s}^{-1}$)	q (cm^{-1})
Fe II	2382	55.845	2382.763995(80)[a,b]	0.320	3.130[c,g]	1505[e,f](100)
		58	2382.7622294[f]	0.282%		
		57	2382.7630560[f]	2.119%		
		56	2382.7639122[f]	91.754%		
		54	2382.7657196[f]	5.845%		
	2586	55.845	2586.649312(87)[a,b]	0.0691	2.720[c]	1515[e,f](100)
		58	2586.6475648[f]	0.282%		
		57	2586.6483830[f]	2.119%		
		56	2586.6492304[f]	91.754%		
		54	2586.6510194[f]	5.845%		
	2600	55.845	2600.172114(88)[a,g]	0.239	2.700[c]	1370[e,f](100)
		58	2600.1703603[f]	0.282%		
		57	2600.1711816[f]	2.119%		
		56	2600.1720322[f]	91.754%		
		54	2600.1738281[f]	5.845%		
Mg I	2852	24.3050	2852.962797(15)	1.83	5.000[h,i,j,k,l][m,n,o,p]	90[q,r](10)
		26	2852.959591(20)[s]	11.01%		
		25	2852.961407(20)[s]	10.00%		
		24	2852.963420(14)[s]	78.99%		
Mg II	2796	24.3050	2796.353790(16)	0.6155	2.625[t]	212[u](2)
		26	2796.34704565(42)[v]	11.01%		
		25	2796.353449(50)[v,w,x]	4.17%		
		25	2796.349030(50)[v,w,x]	5.83%		
		24	2796.35509903(42)[v]	78.99%		
	2803	24.3050	2803.530982(16)	0.3058	2.595[t]	121[u](2)
		26	2803.52420938(42)[v]	11.01%		
		25	2803.530941(50)[v,w,x]	4.17%		
		25	2803.525985(50)[v,w,x]	5.83%		
		24	2803.53229720(42)[v]	78.99%		

(the table continues on the next page)

Ion	Tran	A	λ_0 (Å)	f or %	Γ ($10^8\,\mathrm{s}^{-1}$)	q (cm^{-1})
Mn II	2576	54.9380	2576.87512(11)[a,b,y]	0.361	2.820[z,A,B,C]	1276[D](150)
		55	2576.890898	28.571%		
		55	2576.879368	23.801%		
		55	2576.869849	19.030%		
		55	2576.862494	14.286%		
		55	2576.856181	14.312%		
	2594	54.9380	2594.49643(11)[a,b,y]	0.280	2.780[z,A,B,C]	1030[D](150)
		55	2594.512068	28.579%		
		55	2594.500587	23.841%		
		55	2594.491191	19.078%		
		55	2594.483901	14.289%		
		55	2594.477608	14.213%		
	2606	54.9380	2606.45877(11)[a,b,y]	0.198	2.270[z,A,B,C]	869[D](150)
		55	2606.478271	28.563%		
		55	2606.463977	23.793%		
		55	2606.452264	19.052%		
		55	2606.443176	14.282%		
		55	2606.435406	14.310%		

[a]Aldenius (2009); [b]Nave (2012); [c]Biemont et al. (1991); [d]Guo et al. (1992); [e]Dzuba et al. (2002); [f]Porsev et al. (2009); [g]Schade et al. (1988); [h]Lurio (1964); [i]Smith & Gallagher (1966); [j]Andersen et al. (1970); [k]Smith & Liszt (1971); [l]Lundin et al. (1973); [m]Marek & Richter (1973); [n]Kelly & Mathur (1978); [o]Liljeby et al. (1980); [p]Larsson et al. (1993); [q]Berengut et al. (2005); [r]Savukov & Dzuba (2008); [s]Salumbides et al. (2006); [t]Ansbacher et al. (1989); [u]Dzuba & Johnson (2007); [v]Batteiger et al. (2009); [w]Itano & Wineland (1981); [x]Sur et al. (2005). [y]Blackwell-Whitehead et al. (2005); [z]Kwiatkowski et al. (1982); [A]Pinnington et al. (1992); [B]Schnabel et al. (1995); [C]Kling & Griesmann (2000); [D]Berengut et al. (2004).

Appendix C

LFC-calibrated models of the HE0515−4414 $z_{abs} = 1.15$ system
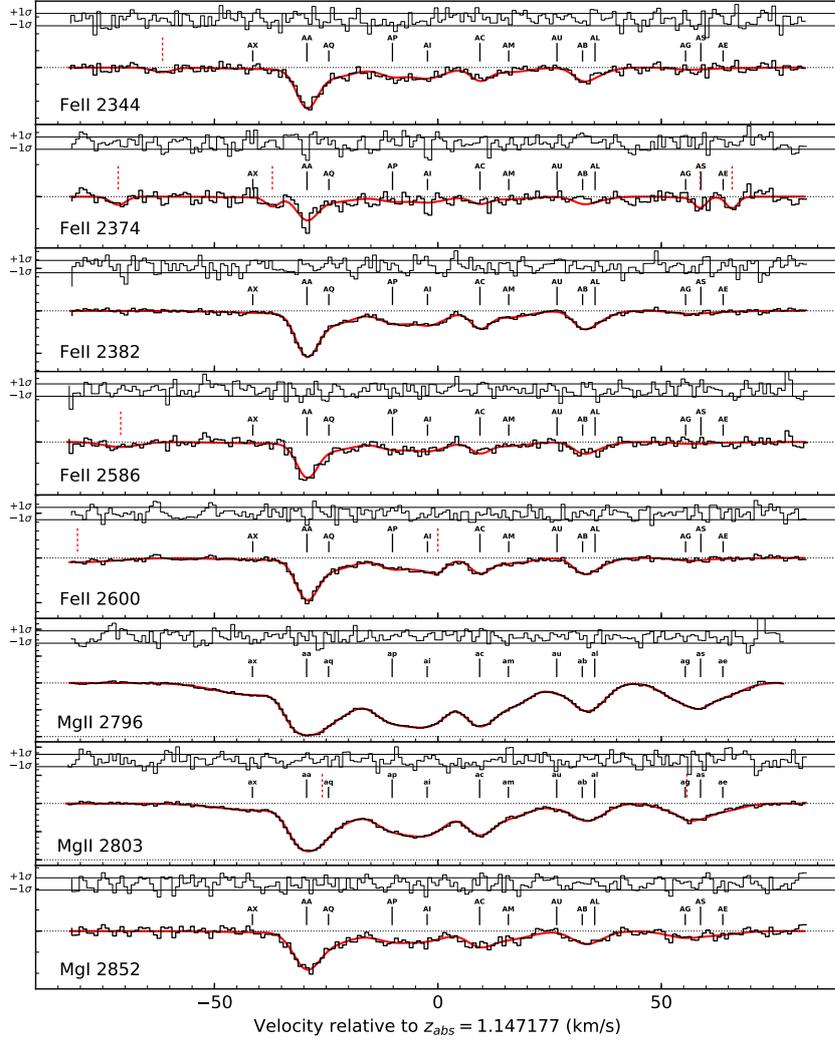
Figure C.1: The black histogram shows the continuum-normalised LFC-calibrated spectrum for spectral region I. Overplotted as a continuous red line is the lowest AICc `AI-VPFIT` model. Black labeled ticks mark the locations of absorption components in the model. The lowercase letters are associated with the transitions which provide the most information about the velocity structure (i.e. the "primary species", see Lee et al. 2020). Slightly longer, dotted red ticks mark the locations of blends from unidentified species (interlopers). The black histogram above the data and the model show the normalised residuals (data-model) and the horizontal lines show the $\pm 1\sigma$ levels. The dotted horizontal line corresponds to a normalised flux of unity. Major (minor) ticks on the y-axis label increments of 0.5 (0.1) in normalised flux.
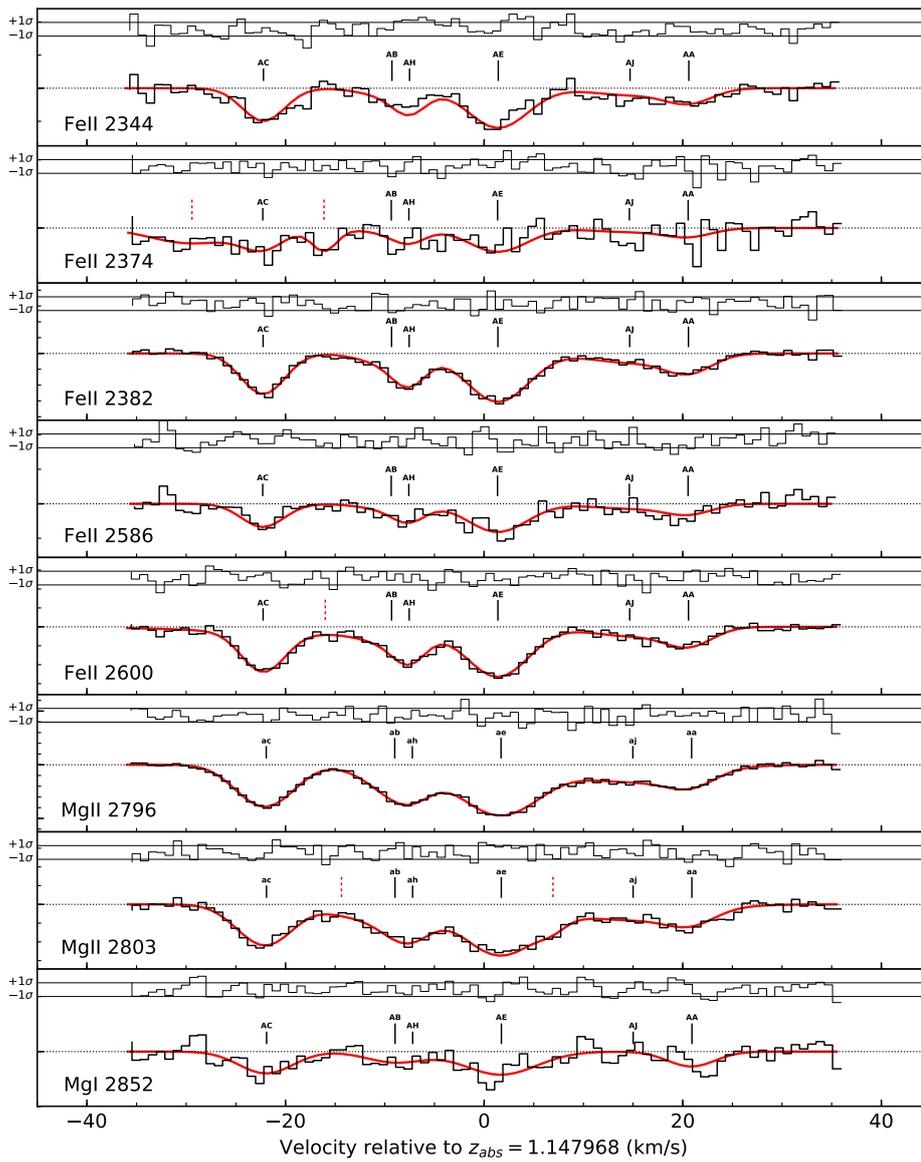
Figure C.2: The same as in Figure C.1, except for LFC-calibrated spectral region II.
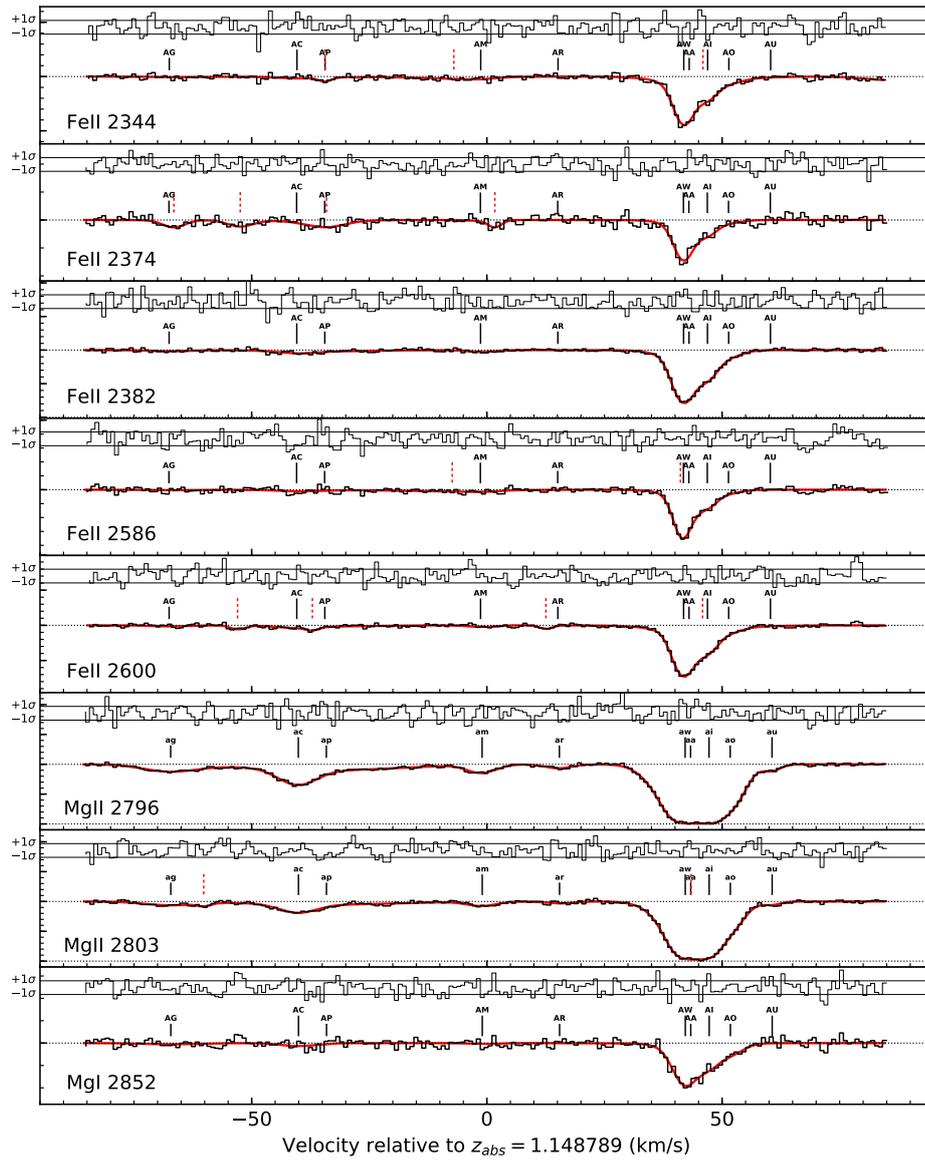
Figure C.3: The same as in Figure C.1, except for LFC-calibrated spectral region III.
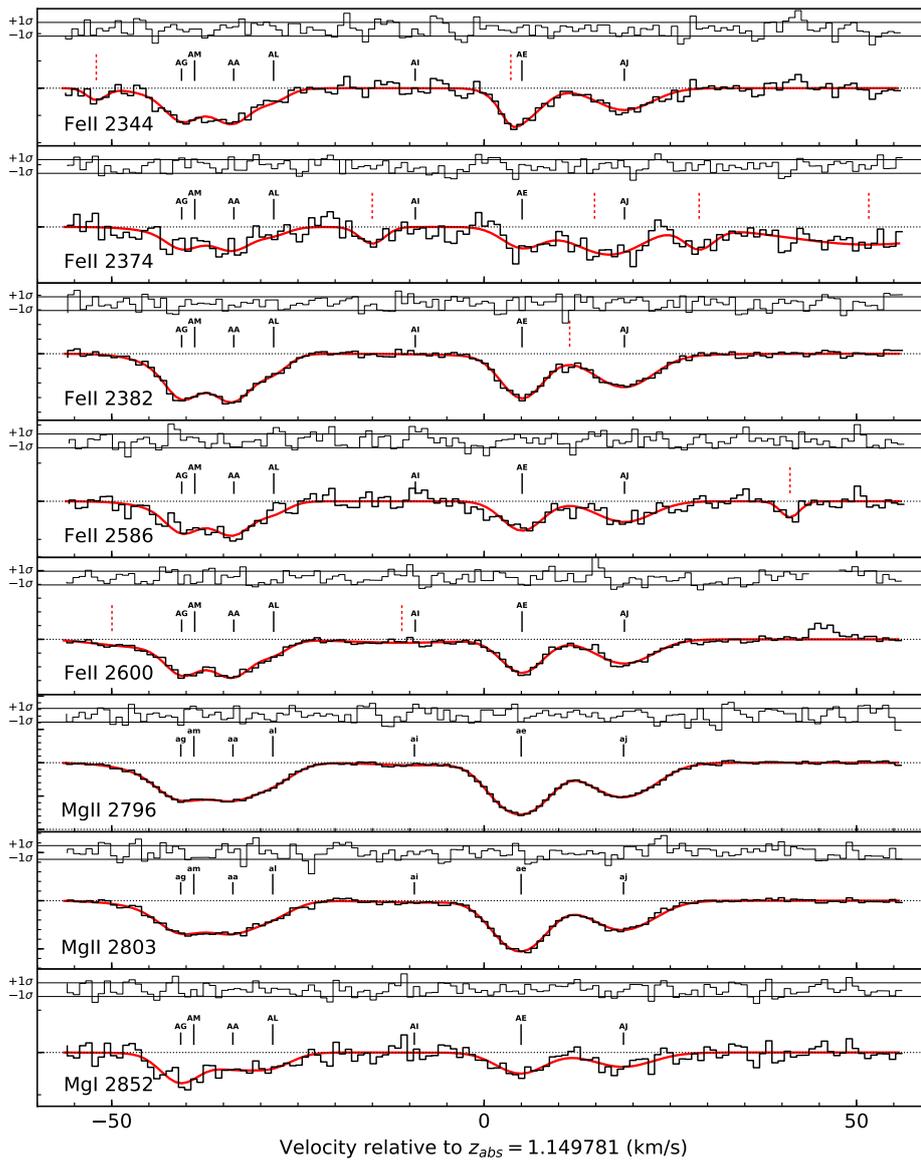
Figure C.4: The same as in Figure C.1, except for LFC-calibrated spectral region IV.
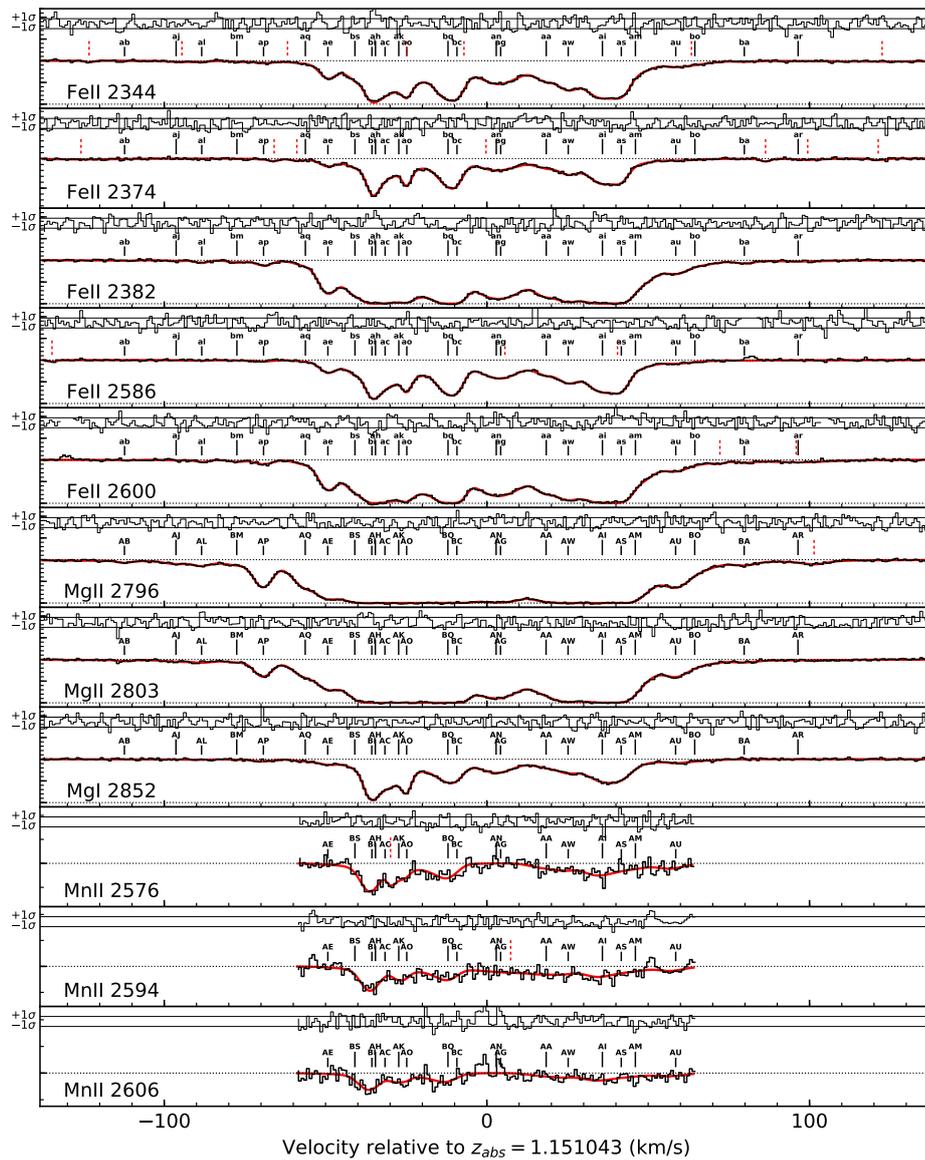
Figure C.5: The same as in Figure C.1, except for LFC-calibrated spectral region V.

Appendix D

ThAr-calibrated models of the HE0515$-$4414 $z_{abs} = 1.15$ system
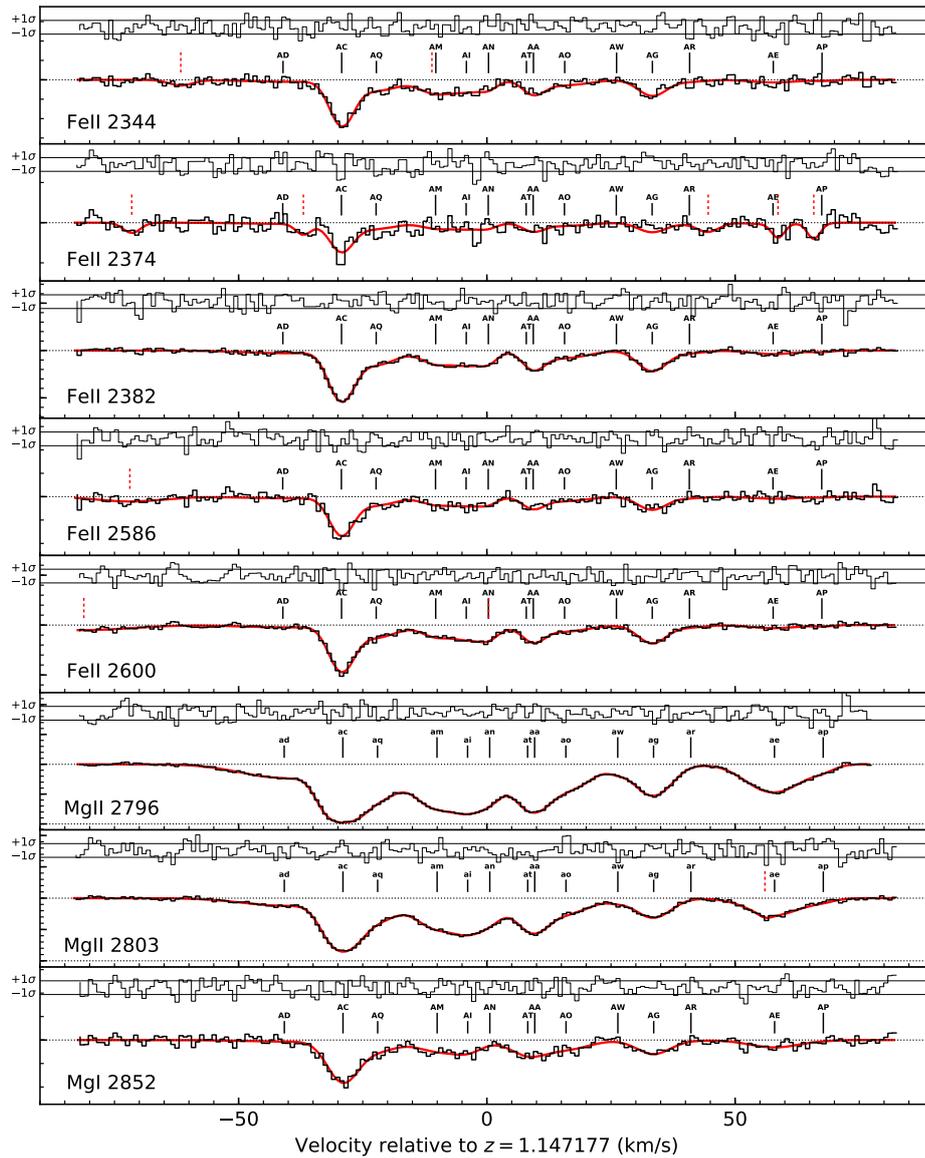
Figure D.1: The same as in Figure C.1, except for ThAr-calibrated spectral region I.
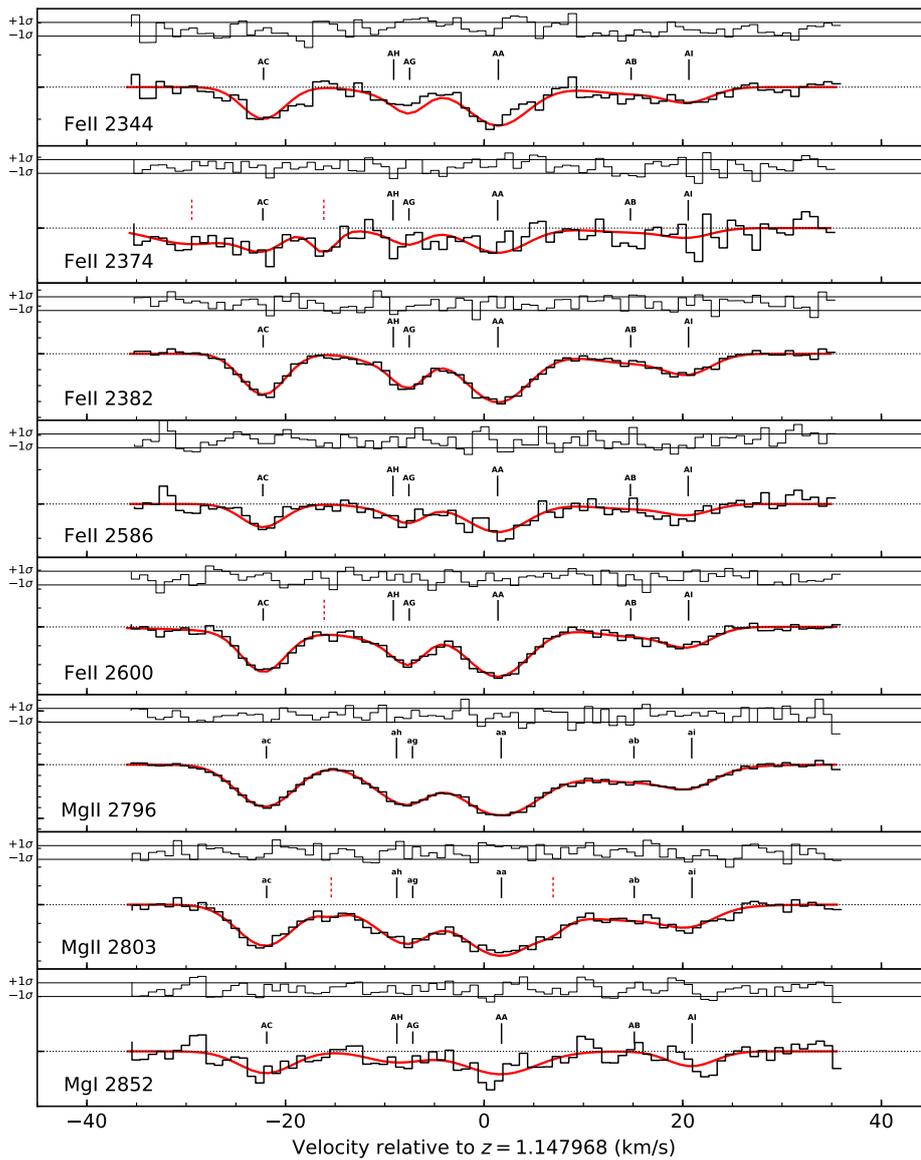
Figure D.2: The same as in Figure C.1, except for ThAr-calibrated spectral region II.
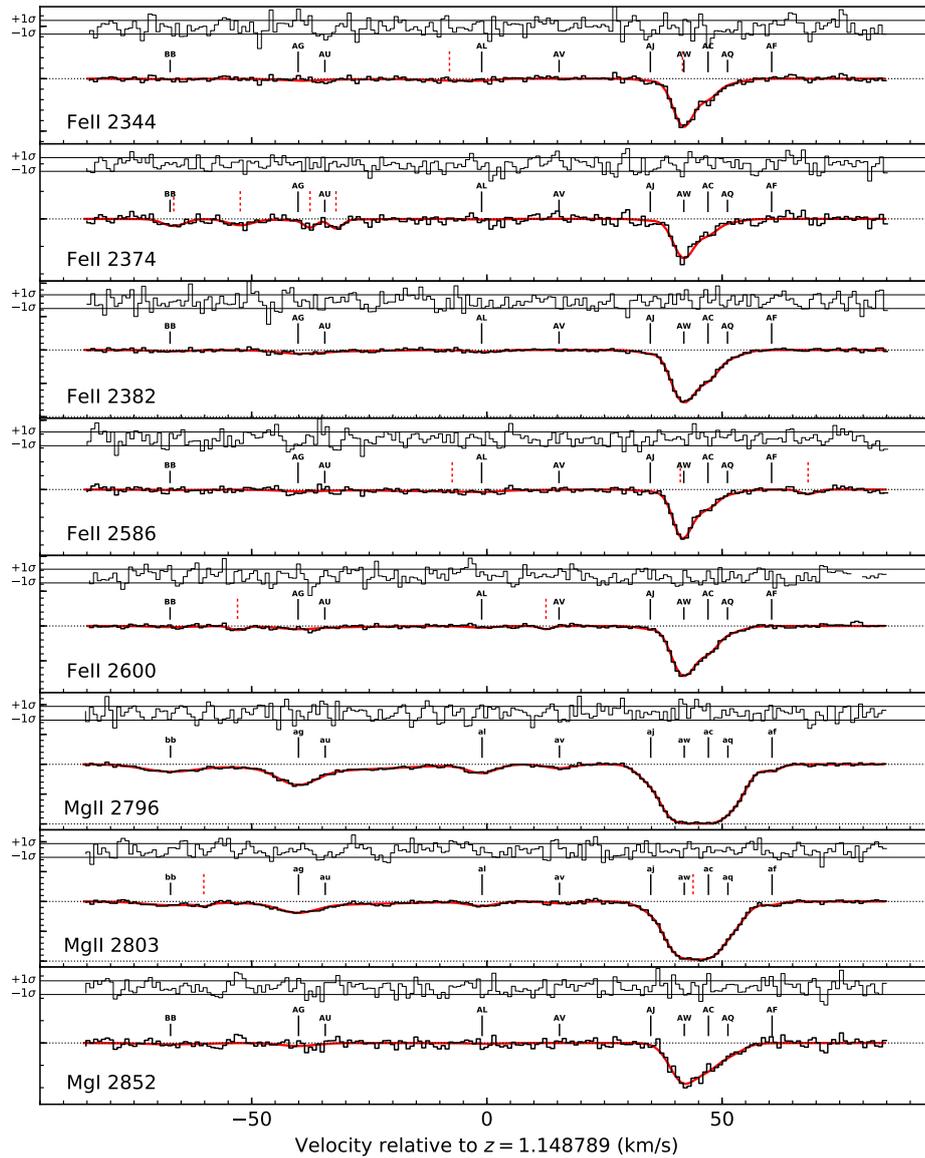
Figure D.3: The same as in Figure C.1, except for ThAr-calibrated spectral region III.
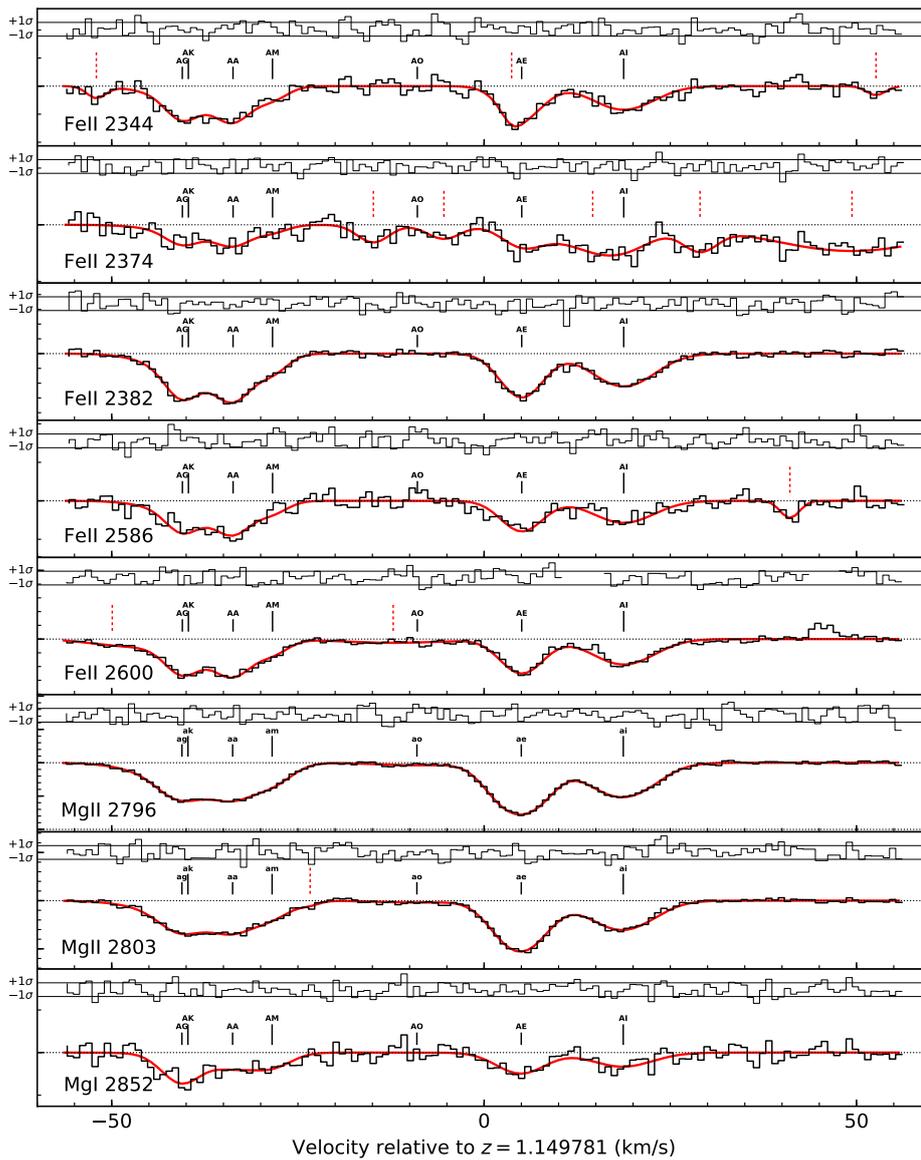
Figure D.4: The same as in Figure C.1, except for ThAr-calibrated spectral region IV.
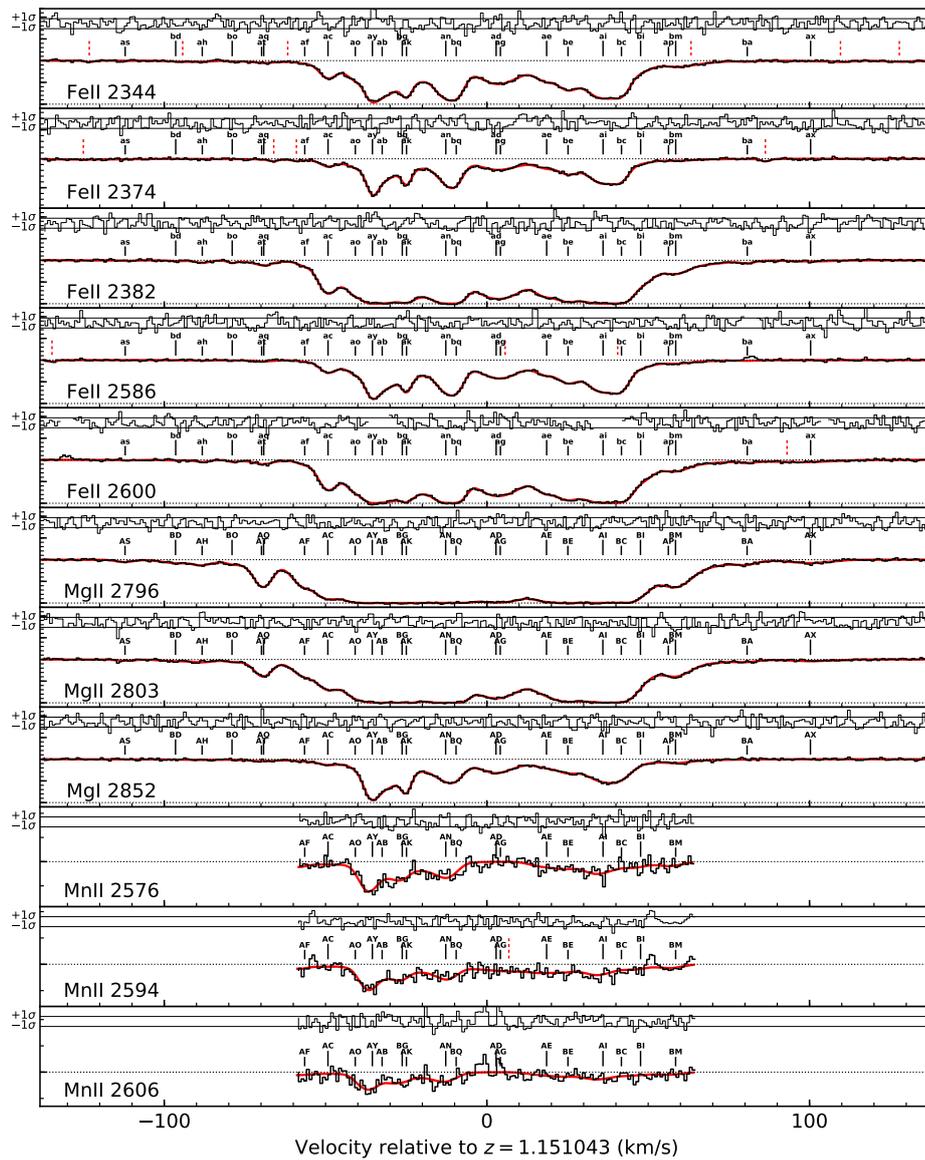
Figure D.5: The same as in Figure C.1, except for ThAr-calibrated spectral region V.

Bibliography

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 241103

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, Phys. Rev. Lett., 119, 161101

Ahmad, Q. R., Allen, R. C., Andersen, T. C., et al. 2002, Phys. Rev. Lett., 89, 011301

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

Alam, S., Ata, M., Bailey, S., et al. 2017, Mon. Not. Roy. Astron. Soc., 470, 2617

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv.org, 0609591

Aldenius, M. 2009, Physica Scripta Volume T, 134, 014008

Andersen, T., Desesquelles, J., Jessen, K. A., & Sørensen, G. 1970, Phys. Rev. A, 1, 1294

Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Nature, 536, 437

Ansbacher, W., Li, Y., & Pinnington, E. H. 1989, Physics Letters A, 139, 165

Antoniadis, J., Freire, P. C. C., Wex, N., et al. 2013, Science, 340, 448

Armendáriz-Picón, C., Damour, T., & Mukhanov, V. 1999, Physics Letters B, 458, 209

Armendariz-Picon, C., Mukhanov, V., & Steinhardt, P. J. 2000, Phys. Rev. Lett., 85, 4438

Armstrong, B. 1967, Journal of Quantitative Spectroscopy and Radiative Transfer, 7, 61

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

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

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, Annu. Rev. Astron. Astrophys., 47, 481

Astudillo-Defru, N., Díaz, R. F., Bonfils, X., et al. 2017, Astron. Astrophys., 605, L11

Avelino, P. P., Esposito, S., Mangano, G., et al. 2001, Phys. Rev. D, 64, 103505

Bahcall, J. N., Sargent, W. L. W., & Schmidt, M. 1967, Astrophys. J. Lett., 149, L11

Bahcall, J. N., Steinhardt, C. L., & Schlegel, D. 2004, Astrophys. J., 600, 520

Bainbridge, M., Barstow, M., Reindl, N., et al. 2017, Universe, 3, 32

Bainbridge, M. B. & Webb, J. K. 2017, Mon. Not. Roy. Astron. Soc., 468, 1639

Baranne, A., Queloz, D., Mayor, M., et al. 1996, Astron. Astrophys. Supp, 119, 373

Barrow, J. D. 1987, Phys. Rev. D, 35, 1805

Barrow, J. D. & Lip, S. Z. W. 2012, Phys. Rev. D, 85, 023514

Batteiger, V., Knünz, S., Herrmann, M., et al. 2009, Phys. Rev. A, 80, 022503

Battye, R. A., Crittenden, R., & Weller, J. 2001, Phys. Rev. D, 63, 043505

Bauer, F. F., Zechmeister, M., & Reiners, A. 2015, Astron. Astrophys., 581, A117

Bechtold, J. 2001, arXiv.org, 0112521

Bekenstein, J. D. 1982, Phys. Rev. D, 25, 1527

Bekenstein, J. D. 2004, Phys. Rev. D, 70, 083509

Berengut, J. C., Dzuba, V. A., Flambaum, V. V., & Marchenko, M. V. 2004, Phys. Rev. A, 70, 064101

Berengut, J. C., Flambaum, V. V., & Kozlov, M. G. 2005, Phys. Rev. A, 72, 044501

Berengut, J. C., Flambaum, V. V., Ong, A., et al. 2013, Phys. Rev. Lett., 111, 010801

Berengut, J. C., Kava, E. M., & Flambaum, V. V. 2012, Astron. Astrophys., 542, A118

Berti, E., Barausse, E., Cardoso, V., et al. 2015, Classical and Quantum Gravity, 32, 243001

Bertone, G., Hooper, D., & Silk, J. 2005, Physics Reports, 405, 279

Bielby, R., Crighton, N. H. M., Fumagalli, M., et al. 2017, Mon. Not. Roy. Astron. Soc., 468, 1373

Biemont, E., Baudoux, M., Kurucz, R. L., Ansbacher, W., & Pinnington, E. H. 1991, Astron. Astrophys., 249, 539

Blackman, R. T., Fischer, D. A., Jurgenson, C. A., et al. 2020, Astron. J., 159, 238

Blackwell-Whitehead, R. J., Toner, A., Hibbert, A., Webb, J., & Ivarsson, S. 2005, Mon. Not. Roy. Astron. Soc., 364, 705

Boggs, P. T., Byrd, R. H., Rogers, J. E., & Schnabel, R. B. 1992, Users Reference Guide for ODRPACK Version 2.01 Software for Weighted Orthogonal Distance Regression, Tech. rep., National Institute of Standards and Technology

Bosman, S. E. I., Becker, G. D., Haehnelt, M. G., et al. 2017, Mon. Not. Roy. Astron. Soc., 470, 1919

BOSS Collaboration, Aubourg, É., Bailey, S., et al. 2015, Phys. Rev. D, 92, 123516

Bouchy, F., Isambert, J., Lovis, C., et al. 2009, EAS Publications Series, 37, 247

Bouchy, F., Pepe, F., & Queloz, D. 2001, Astron. Astrophys., 374, 733

Braje, D. A., Kirchner, M. S., Osterman, S., Fortier, T., & Diddams, S. A. 2008, European Physical Journal D, 48, 57

Breckinridge, J. B., Pierce, A. K., & Stoll, C. P. 1975, Astrophys. J. Supp., 29, 97

Bretagnon, P. & Francou, G. 1988, Astron. Astrophys., 202, 309

Brucalassi, A., Grupp, F., Kellermann, H., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Proceedings of SPIE, 99085W

Burbidge, E. M., Lynds, C. R., & Burbidge, G. R. 1966, Astrophys. J., 144, 447

Campbell, B. A. & Olive, K. A. 1995, Physics Letters B, 345, 429

Carroll, S. M. 1998, Phys. Rev. Lett., 81, 3067

Carswell, R. F., Becker, G. D., Jorgenson, R. A., Murphy, M. T., & Wolfe, A. M. 2012, Mon. Not. Roy. Astron. Soc., 422, 1700

Carswell, R. F. & Webb, J. K. 2014, VPFIT: Voigt profile fitting program

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, Astrophys. J. Lett., 437, L9

Chance, K. & Kurucz, R. L. 2010, Journal of Quantitative Spectroscopy and Radiative Transfer, 111, 1289

Chand, H., Srianand, R., Petitjean, P., et al. 2006, Astron. Astrophys., 451, 45

Cingöz, A., Lapierre, A., Nguyen, A. T., et al. 2007, Phys. Rev. Lett., 98, 040801

Clara, M. T. & Martins, C. J. A. P. 2020, Astron. Astrophys., 633, L11

Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, Physics Reports, 513, 1

Coffinet, A., Lovis, C., Dumusque, X., & Pepe, F. 2019, Astron. Astrophys., 629, A27

Copeland, E. J., Nunes, N. J., & Pospelov, M. 2004, Phys. Rev. D, 69, 023501

Copeland, E. J., Sami, M., & Tsujikawa, S. 2006, International Journal of Modern Physics D, 15, 1753

Damour, T. & Dyson, F. 1996, Nuclear Physics B, 480, 37

Daprà, M., Niu, M. L., Salumbides, E. J., Murphy, M. T., & Ubachs, W. 2016, Astrophys. J., 826, 192

Darling, J. 2012, Astrophys. J. Lett., 761, L26

Davis, E. D. & Hamdan, L. 2015, Phys. Rev. B, 92, 014319

Dekker, H., D'Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4008, Proceedings of SPIE, ed. M. Iye & A. F. Moorwood, 534–545

Dicke, R. H. 1959, Nature, 183, 170

Doerr, H. P., Steinmetz, T., Holzwarth, R., Kentischer, T., & Schmidt, W. 2012, Solar Physics, 280, 663

Duff, M. J., Okun, L. B., & Veneziano, G. 2002, Journal of High Energy Physics, 2002, 023

Dumont, V. & Webb, J. K. 2017, Mon. Not. Roy. Astron. Soc., 468, 1568

Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015a, Astrophys. J. Lett., 814, L21

Dumusque, X., Pepe, F., Lovis, C., & Latham, D. W. 2015b, Astrophys. J., 808, 171

Dvali, G. & Zaldarriaga, M. 2002, Phys. Rev. Lett., 88, 091303

Dyson, F. J. 1972, in Aspects of Quantum Theory, ed. A. Salam & E. P. Wigner (Cambridge: Cambridge University Press), 213–236

Dzuba, V. A. & Flambaum, V. V. 2009, Canadian Journal of Physics, 87, 15

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

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. & Johnson, W. R. 2007, Phys. Rev. A, 76, 062510

eBOSS Collaboration, Alam, S., Aubert, M., et al. 2020, arXiv.org, 2007.08991

Einstein, A. 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), 142

ESO ELT team. 2010, The science case for the European Extremely Large Telescope: The next step in mankind's quest for the Universe (ESO online)

ESO ELT team. 2011, An Expanded View of the Universe Science with the European Extremely Large Telescope (ESO online)

Evans, T. M., Murphy, M. T., Whitmore, J. B., et al. 2014, Mon. Not. Roy. Astron. Soc., 445, 128

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, Astrophys. J. Lett., 875, L1

Ezquiaga, J. M. & Zumalacárregui, M. 2017, Phys. Rev. Lett., 119, 251304

Fanson, J., McCarthy, P. J., Bernstein, R., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10700, Proceedings of SPIE, 1070012

Fenner, Y., Murphy, M. T., & Gibson, B. K. 2005, Mon. Not. Roy. Astron. Soc., 358, 468

Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001

Frank, C., Kerber, F., Avila, G., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII, 107026P

Fujii, Y., Iwamoto, A., Fukahori, T., et al. 2000, Nuclear Physics B, 573, 377

Fukuda, Y., Hayakawa, T., Ichihara, E., et al. 1998, Phys. Rev. Lett., 81, 1562

Gamow, G. 1948a, Nature, 162, 680

Gamow, G. 1948b, Physical Review, 74, 505

Georgi, H. & Glashow, S. L. 1974, Phys. Rev. Lett., 32, 438

Glenday, A. G., Li, C.-H., Langellier, N., et al. 2015, Optica, 2, 250

Gravity Collaboration, Abuter, R., Amorim, A., et al. 2018, Astron. Astrophys., 615, L15

Gravity Collaboration, Abuter, R., Amorim, A., et al. 2020, Astron. Astrophys., 636, L5

Green, M. B. 1985, Nature, 314, 409

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

Guo, B., Ansbacher, W., Pinnington, E. H., Ji, Q., & Berends, R. W. 1992, Phys. Rev. A, 46, 641

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hannestad, S. 1999, Phys. Rev. D, 60, 023515

Hänsch, T. W. 2006, Reviews of Modern Physics, 78, 1297

Hao, Z., Ye, H., Han, J., et al. 2018, PASP, 130, 125001

Harris, Daniel L., I. 1948, Astrophys. J., 108, 112

Hazard, C., Jauncey, D., Goss, W. M., & Herald, D. 2018, PASA, 35, e006

Hees, A., Do, T., Roberts, B. M., et al. 2020, Phys. Rev. Lett., 124, 081101

Hook, I. 2009, in Science with the VLT in the ELT Era, ed. A. Moorwood (Dordrecht: Springer Netherlands), 225–232

Horne, K. 1986, PASP, 98, 609

Hu, J., Webb, J. K., Ayres, T. R., et al. 2020, arXiv.org, 2007.10905

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Hubble, E. & Humason, M. L. 1931, Astrophys. J., 74, 43

Hulse, R. A. & Taylor, J. H. 1975, Astrophys. J. Lett., 195, L51

Hunter, T. R. & Ramsey, L. W. 1992, PASP, 104, 1244

Itano, W. M. & Wineland, D. J. 1981, Phys. Rev. A, 24, 1364

Kaluza, T. 2018, International Journal of Modern Physics D, 27, 1870001

Kaplinghat, M., Scherrer, R. J., & Turner, M. S. 1999, Phys. Rev. D, 60, 023516

Kelly, F. M. & Mathur, M. S. 1978, Canadian Journal of Physics, 56, 1422

King, J., Mortlock, D., Webb, J., & Murphy, M. 2010, Highlights of Astronomy, 15, 318

King, J. A., Webb, J. K., Murphy, M. T., et al. 2012, Mon. Not. Roy. Astron. Soc., 422, 3370

Klein, O. 1926, Zeitschrift fur Physik, 37, 895

Kling, R. & Griesmann, U. 2000, Astrophys. J., 531, 1173

Kolb, E. W., Perry, M. J., & Walker, T. P. 1986, Phys. Rev. D, 33, 869

Kormendy, J. & Ho, L. C. 2013, Annu. Rev. Astron. Astrophys., 51, 511

Kotuš, S. M., Murphy, M. T., & Carswell, R. F. 2017, Mon. Not. Roy. Astron. Soc., 464, 3679

Kullback, S. & Leibler, R. A. 1951, Ann. Math. Statist., 22, 79

Kwiatkowski, M., Micali, G., Werner, K., & Zimmermann, P. 1982, Journal of Physics B Atomic Molecular Physics, 15, 4357

Landau, S. J., Mosquera, M. E., Scóccola, C. G., & Vucetich, H. 2008, Phys. Rev. D, 78, 083527

Lanza, A. F., Molaro, P., Monaco, L., & Haywood, R. D. 2016, Astron. Astrophys., 587, A103

Larsson, J., Zerne, R., Persson, A., Wahlström, C. G., & Svanberg, S. 1993, Zeitschrift fur Physik D Atoms Molecules Clusters, 27, 329

Leão, I. C., Pasquini, L., Ludwig, H. G., & de Medeiros, J. R. 2019, Mon. Not. Roy. Astron. Soc., 483, 5026

Lee, C.-C., Webb, J. K., Carswell, R. F., & Milaković, D. 2020, arXiv.org, 2008.02583

Leibundgut, B. 2017, History of Supernovae as Distance Indicators, ed. A. W. Alsabti & P. Murdin (Cham: Springer International Publishing), 2525–2542

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

Levshakov, S. A., Centurión, M., Molaro, P., et al. 2006, Astron. Astrophys., 449, 879

Lévy-Leblond, J.-M. 1979, in International school of physics on problems in the foundations of physics, ed. G. Toraldo di Francia

Liljeby, L., Lindgård, A., Mannervik, S., Veje, E., & Jelenkovic, B. 1980, Physica Scripta, 21, 805

Linde, A. D. 1982, Physics Letters B, 108, 389

Liske, J. 2014, https://www.eso.org/sci/facilities/eelt/docs/ESO-204697_1_Top_Level_Requirements_for_ELT-HIRES.pdf

Liske, J., Grazian, A., Vanzella, E., et al. 2008, Monthly Notices of the Royal Astronomical Society, 386, 1192

Lo Curto, G., Pepe, F., Avila, G., et al. 2015, The Messenger, 162, 9

Loeb, A. 1998, Astrophys. J. Lett., 499, L111

Löhner-Böttcher, J., Schmidt, W., Doerr, H. P., et al. 2017, Astron. Astrophys., 607, A12

Lopez-Honorez, L., Mena, O., Palomares-Ruiz, S., Villanueva-Domingo, P., & Witte, S. J. 2020, Journal of Cosmology and Astroparticle Physics, 2020, 026

Lovis, C., Mayor, M., Pepe, F., et al. 2006, Nature, 441, 305

Lundin, L., Engman, B., Hilke, J., & Martinson, I. 1973, Physica Scripta, 8, 274

Lurio, A. 1964, Physical Review, 136, 376

Lynden-Bell, D. 1969, Nature, 223, 690

Lynds, R. 1971, Astrophys. J. Lett., 164, L73

Lyth, D. H. D. H. & Riotto, A. A. 1999, Physics Reports, 314, 1

Maartens, R. 2004, Living Reviews in Relativity, 7, 7

Magueijo, J., Barrow, J. D., & Sandvik, H. B. 2002, Physics Letters B, 549, 284

Mahadevan, S., Halverson, S., Ramsey, L., & Venditti, N. 2014, Astrophys. J., 786, 18

Maiolino, R., Haehnelt, M., Murphy, M. T., et al. 2013, arXiv.org, 1310.3163

Marconi, A., Di Marcantonio, P., D'Odorico, V., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Proceedings of SPIE, 990823

Marek, J. & Richter, J. 1973, Astron. Astrophys., 26, 155

Marra, V. & Rosati, F. 2005, Journal of Cosmology and Astroparticle Physics, 2005, 011

Marrodán Undagoitia, T. & Rauch, L. 2016, Journal of Physics G Nuclear Physics, 43, 013001

Martins, C. J. A. P. 2015, General Relativity and Gravitation, 47, 1843

Martins, C. J. A. P. 2017, Reports on Progress in Physics, 80, 126902

Martins, J. H. C., Figueira, P., Santos, N. C., et al. 2018, Mon. Not. Roy. Astron. Soc., 478, 5240

Maurette, M. 1976, Annual Review of Nuclear and Particle Science, 26, 319

Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20

Mayor, M. & Queloz, D. 1995, Nature, 378, 355

McCracken, R. A., Charsley, J. M., & Reid, D. T. 2017a, Optics Express, 25, 15058

McCracken, R. A., Depagne, É., Kuhn, R. B., et al. 2017b, Optics Express, 25, 6450

Milaković, D., Pasquini, L., Lo Curto, G., et al. 2017, in ESO Calibration Workshop: The Second Generation VLT Instruments and Friends, 30

Milaković, D., Pasquini, L., Webb, J. K., & Lo Curto, G. 2020, Mon. Not. Roy. Astron. Soc., 493, 3997

Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, Astrophys. J., 471, 582

Moffat, J. W. 2006, Journal of Cosmology and Astroparticle Physics, 2006, 004

Molaro, P., Centurión, M., Whitmore, J. B., et al. 2013a, Astron. Astrophys., 555, A68

Molaro, P., Esposito, M., Monai, S., et al. 2013b, Astron. Astrophys., 560, A61

Molaro, P., Reimers, D., Agafonova, I. I., & Levshakov, S. A. 2008, European Physical Journal Special Topics, 163, 173

Mota, D. F. & Barrow, J. D. 2004a, Mon. Not. Roy. Astron. Soc., 349, 291

Mota, D. F. & Barrow, J. D. 2004b, Physics Letters B, 581, 141

Murphy, M. T. & Berengut, J. C. 2014, Mon. Not. Roy. Astron. Soc., 438, 388

Murphy, M. T., Flambaum, V. V., Webb, J. K., et al. 2004, Constraining Variations in the Fine-Structure Constant, Quark Masses and the Strong Interaction, ed. S. G. Karshenboim & E. Peik, Vol. 648, 131–150

Murphy, M. T., Udem, T., Holzwarth, R., et al. 2007, Mon. Not. Roy. Astron. Soc., 380, 839

Murphy, M. T., Webb, J. K., & Flambaum, V. V. 2003a, Mon. Not. Roy. Astron. Soc., 345, 609

Murphy, M. T., Webb, J. K., Flambaum, V. V., Churchill, C. W., & Prochaska, J. X. 2001a, Mon. Not. Roy. Astron. Soc., 327, 1223

Murphy, M. T., Webb, J. K., Flambaum, V. V., & Curran, S. J. 2003b, Astrophys. Space Sci., 283, 577

Murphy, M. T., Webb, J. K., Flambaum, V. V., Prochaska, J. X., & Wolfe, A. M. 2001b, Mon. Not. Roy. Astron. Soc., 327, 1237

Naudet, R. & Ronson, C. 1974, Bulletin dInformations Scientifiques et Techniques du Commissariats˜ lEnergie Atomique (France) No, 193, 36

Nave, G. 2012, Mon. Not. Roy. Astron. Soc., 420, 1570

O'Bryan, J., Smidt, J., De Bernardis, F., & Cooray, A. 2015, Astrophys. J., 798, 18

Obrzud, E., Brasch, V., Voumard, T., et al. 2019, Optics Letters, 44, 5290

Olive, K. A. & Particle Data Group. 2014, Chinese Physics C, 38, 090001

Olive, K. A. & Pospelov, M. 2002, Phys. Rev. D, 65, 085044

Olive, K. A., Pospelov, M., Qian, Y.-Z., et al. 2004, Phys. Rev. D, 69, 027701

Osterman, S., Diddams, S., Beasley, M., et al. 2007, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6693, Proceedings of SPIE, 66931G

Palmer, B. A. & Engleman, R. 1983, Atlas of the Thorium spectrum (LA, Los Alamos: National Laboratory)

Pati, J. C. & Salam, A. 1974, Phys. Rev. D, 10, 275

Pepe, F., Lovis, C., Ségransan, D., et al. 2011, Astron. Astrophys., 534, A58

Pepe, F., Molaro, P., Cristiani, S., et al. 2014, Astronomische Nachrichten, 335, 8

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, Astrophys. J., 517, 565

Petrov, Y. V. 1977, Soviet Physics Uspekhi, 20, 937

Pfeiffer, M., Frank, C., & Baumuller, D. 1992, Astronomische Gesellschaft Abstract Series, 7, 191

Pfeiffer, M. J., Frank, C., Baumueller, D., Fuhrmann, K., & Gehren, T. 1998, Astron. Astrophys. Supp, 130, 381

Phillips, D. F., Glenday, A. G., Li, C.-H., et al. 2012, Optics Express, 20, 13711

Picqué, N. & Hänsch, T. W. 2019, Nature Photonics, 13, 146

Pinho, A. M. M. & Martins, C. J. A. P. 2016, Physics Letters B, 756, 121

Pinnington, E. H., Guo, G., Ji, Q., et al. 1992, Journal of Physics B Atomic Molecular Physics, 25, L475

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, Astron. Astrophys., 594, A13

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, Astron. Astrophys., 580, A22

Porsev, S. G., Kozlov, M. G., & Reimers, D. 2009, Phys. Rev. A, 79, 032519

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in FORTRAN. The art of scientific computing (Cambridge University Press)

Probst, R. A. 2015, PhD thesis, Ludwig Maximillian Universität, Munich, Germany

Probst, R. A., Lo Curto, G., Ávila, G., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Proceedings of SPIE, 990864

Probst, R. A., Lo Curto, G., Avila, G., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91471C

Probst, R. A., Milaković, D., Toledo-Padrón, B., et al. 2020, Nature Astronomy, 4, 603

Probst, R. A., Steinmetz, T., Wilken, T., et al. 2013, Optics Express, 21, 11670

Probst, R. A., Steinmetz, T., Wilken, T., et al. 2013, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8864, Techniques and Instrumentation for Detection of Exoplanets VI, 88641Z

Probst, R. A., Steinmetz, T., Wu, Y., et al. 2017, Applied Physics B: Lasers and Optics, 123, 76

Probst, R. A., Wang, L., Doerr, H. P., et al. 2015a, New Journal of Physics, 17, 023048

Probst, R. A., Wu, Y., Steinmetz, T., et al. 2015b, in Conference on Lasers and Electro-Optics (CLEO), 1–2

Quast, R., Reimers, D., & Levshakov, S. A. 2004, Astron. Astrophys., 415, L7

Quinlan, F., Ycas, G., Osterman, S., & Diddams, S. A. 2010, Review of Scientific Instruments, 81, 063105

Rahmani, H., Wendt, M., Srianand, R., et al. 2013, Mon. Not. Roy. Astron. Soc., 435, 861

Ratra, B. & Peebles, P. J. E. 1988, Phys. Rev. D, 37, 3406

Rauch, M. 1998, Annu. Rev. Astron. Astrophys., 36, 267

Ravi, A., Langellier, N., Phillips, D. F., et al. 2019a, Phys. Rev. Lett., 123, 091101

Ravi, A., Phillips, D., Langellier, N., et al. 2019b, in APS Meeting Abstracts, Vol. 2019, APS Division of Atomic, Molecular and Optical Physics Meeting Abstracts, L01.029

Redman, S. L., Ycas, G. G., Terrien, R., et al. 2012, Astrophys. J. Supp., 199, 2

Reimers, D., Hagen, H. J., Rodriguez-Pascual, P., & Wisotzki, L. 1998, Astron. Astrophys., 334, 96

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, Astron. J., 116, 1009

Robertson, J. G. 1986, PASP, 98, 1220

Rodler, F. & Lo Curto, G. 2019, HARPS User Manual, 2nd edn., ESO

Rosenband, T., Hume, D. B., Schmidt, P. O., et al. 2008, Science, 319, 1808

Rosman, K. J. R. & Taylor, P. D. P. 1998, Journal of Physical and Chemical Reference Data, 27, 1275

Rubin, V. C., Ford, W. K., J., & Thonnard, N. 1980, Astrophys. J., 238, 471

Rupprecht, G., Pepe, F., Mayor, M., et al. 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5492, Ground-based Instrumentation for Astronomy, ed. A. F. M. Moorwood & M. Iye, 148–159

Safronova, M. S., Budker, D., DeMille, D., et al. 2018, Reviews of Modern Physics, 90, 025008

Salam, A. & Ward, J. C. 1964, Physics Letters, 13, 168

Salpeter, E. E. 1964, Astrophys. J., 140, 796

Salumbides, E. J., Hannemann, S., Eikema, K. S. E., & Ubachs, W. 2006, Mon. Not. Roy. Astron. Soc., 373, L41

Sandage, A. 1962, Astrophys. J., 136, 319

Sandvik, H. B., Barrow, J. D., & Magueijo, J. 2002, Phys. Rev. Lett., 88, 031302

Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D. 1980, Astrophys. J. Supp., 42, 41

Savedoff, M. P. 1956, Nature, 178, 688

Savukov, I. M. & Dzuba, V. A. 2008, Phys. Rev. A, 77, 042501

Schade, W., Mundt, B., & Helbig, V. 1988, Journal of Physics B Atomic Molecular Physics, 21, 2691

Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., et al. 1998, Astrophys. J., 507, 46

Schmidt, M. 1963, Nature, 197, 1040

Schnabel, R., Bard, A., & Kock, M. 1995, Zeitschrift fur Physik D Atoms Molecules Clusters, 34, 223

Seljak, U., Slosar, A., & McDonald, P. 2006, Journal of Cosmology and Astroparticle Physics, 2006, 014

Shaw, D. J. & Barrow, J. D. 2005, Phys. Rev. D, 71, 063525

Shlyakhter, A. I. 1976, Nature, 264, 340

Silva, M. F., Winther, H. A., Mota, D. F., & Martins, C. J. A. P. 2014, Phys. Rev. D, 89, 024025

Silverwood, H. & Easther, R. 2019, PASA, 36, e038

Simard, L., Ellerbroek, B., Bhatia, R., Radovan, M., & Chisholm, E. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9908, Proceedings of SPIE, 99081V

Skordis, C. & Złosnik, T. 2020, arXiv.org, 2007.00082

Smith, W. H. & Liszt, H. S. 1971, Journal of the Optical Society of America (1917-1983), 61, 938

Smith, W. W. & Gallagher, A. 1966, Physical Review, 145, 26

Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137

Stadnik, Y. V. & Flambaum, V. V. 2015, Phys. Rev. Lett., 115, 201301

Steinmetz, T., Wilken, T., Araujo-Hauck, C., et al. 2008, Science, 321, 1335

Stockton, A. N. & Lynds, C. R. 1966, Astrophys. J., 144, 451

Sugiura, N. 1978, Statistics-Theory and Methods, 7, 13

Suh, M.-G., Yi, X., Lai, Y.-H., et al. 2019, Nature Photonics, 13, 25

Sur, C., Sahoo, B. K., Chaudhuri, R. K., Das, B. P., & Mukherjee, D. 2005, European Physical Journal D, 32, 25

Tamai, R., Koehler, B., Cirasuolo, M., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10700, Proceedings of SPIE, 1070014

Tanabashi, M., Hagiwara, K., Hikasa, K., et al. 2018, Phys. Rev. D, 98, 030001

Thompson, R. I. 1975, Astrophys. Lett., 16, 3

Truppe, S., Hendricks, R. J., Tokunaga, S. K., et al. 2013, Nature Communications, 4, 2600

Ubachs, W. 2018, Space Sci. Rev., 214, 3

Udem, T., Holzwarth, R., & Hänsch, T. W. 2002, Nature, 416, 233

Uzan, J.-P. 2011, Living Reviews in Relativity, 14, 2

van de Hulst, H. C. & Reesinck, J. J. M. 1947, Astrophys. J., 106, 121

Varshalovich, D. A. & Levshakov, S. A. 1993, Soviet Journal of Experimental and Theoretical Physics Letters, 58, 237

Varshalovich, D. A., Panchuk, V. E., & Ivanchik, A. V. 1996, Astronomy Letters, 22, 6

Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, Astron. Astrophys., 536, A105

Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A. 2006, Phys. Rev. Lett., 97, 071301

Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2198, Proceedings of SPIE, ed. D. L. Crawford & E. R. Craine, 362

Webb, J. K., Flambaum, V. V., Churchill, C. W., Drinkwater, M. J., & Barrow, J. D. 1999, Phys. Rev. Lett., 82, 884

Webb, J. K., King, J. A., Murphy, M. T., et al. 2011, Phys. Rev. Lett., 107, 191101

Webb, J. K., Murphy, M. T., Flambaum, V. V., et al. 2001, Phys. Rev. Lett., 87, 091301

Webb, J. K., Wright, A., Koch, F. E., & Murphy, M. T. 2014, Memorie della Società Astronomica Italiana, 85, 57

Weinberg, S. 1967, Phys. Rev. Lett., 19, 1264

Whitmore, J. B. & Murphy, M. T. 2015, Mon. Not. Roy. Astron. Soc., 447, 446

Whitmore, J. B., Murphy, M. T., & Griest, K. 2010, Astrophys. J., 723, 89

Wilczynska, M. R., Webb, J. K., Bainbridge, M., et al. 2020, Science Advances, 6

Wilczynska, M. R., Webb, J. K., King, J. A., et al. 2015, Mon. Not. Roy. Astron. Soc., 454, 3082

Wilken, T., Curto, G. L., Probst, R. A., et al. 2012, Nature, 485, 611

Wilken, T., Lovis, C., Manescau, A., et al. 2010, Mon. Not. Roy. Astron. Soc., 405, L16

Wilkinson, D. H. 1958, Philosophical Magazine, 3, 582

Wu, Y. et al. 2020, expected to appear in Opt. Lett.

Ycas, G. G., Quinlan, F., Diddams, S. A., et al. 2012, Optics Express, 20, 6631

Yi, X., Vahala, K., Li, J., et al. 2016, Nature Communications, 7, 10436

Zel'dovich, Y. B. 1964, Soviet Physics Doklady, 9, 195

Zhao, F., Curto, G. L., Pasquini, L., & Zhao, G. 2014, in IAU Symposium, Vol. 293, Formation, Detection, and Characterization of Extrasolar Habitable Planets, ed. N. Haghighipour, 407–409

Zhao, F., Lo Curto, G., Pasquini, L., et al. in prep., Astronomy and Astrophysics

Zumalacárregui, M. & Seljak, U. 2018, Phys. Rev. Lett., 121, 141101

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

Acknowledgements

Although likely to be among the first complainers[1], I have nothing but the words of gratitude for the time I spent at ESO. These four years provided many opportunities which would not be possible elsewhere.

This thesis was written under the supervision of Luca Pasquini (ESO) and John K. Webb (UNSW). Both have been extremely supportive and helpful, providing guidance, insight, entertainment, and food when necessary. I am thankful to them both for everything they have done. To Luca, thank you for sending me to La Silla in my first month of studies, providing legendary pizza, your relaxed attitude, and allowing me to take a leave of absence when it was necessary. To John, thank you for the amusing stories, numerous long video calls, and the hospitality of your home in Cambridge. The phrases "do it right or do not do it at all" and "it should not take more than a few days" will most certainly resonate throughout my career.

I thank Stella, Nelma, Paola, Eric, Rob, Giacomo and Bruno from the Office for Science and Marina and Carlo from the User Support Department for their help along the way and always having their doors open for surprise visits. Many thanks to everyone from the Observing Programmes Office for the great fun at the legendary OPC 100. I thank Bruno and Jason for the enlightening, useful and entertaining conversations throughout the last four years. I thank Jason specifically for providing comments on the thesis, and occasionally informing me that what I have done is stupid, if not wrong. This was (partly) offset by Michele's invitation to join the ELT working group, which I found very encouraging.

Everything in Munich would have been very different without Anita. The time spent with you, travelling with no plans but also no rush, stopping to enjoy the good things in life, enriched my life in Munich immensely.

It is due to a large number of dinners with Anita, Federico, Alice, Chiara, Sara, Giuliana, and other Italians at ESO that my knowledge of Italian surpasses that of German. Caffè Intalleo provided a true experience of an Italian caffè right at my office doorstep, and a Christmas in Calabria set the basis to my future southern-italian accent. There are many other people who significantly enriched my stay in Munich by providing advice, entertainment, and encouragement. Among them are Prash, Remco, Audrey, Adam, Tereza, Richard, Paulo, Christian, other friends loosely grouped into "the British" and "the Croats", as well as many other ESO students, fellows, and staff members.

I hope this manuscript utterly disappoints Filip and Mislav, who expected me to produce a flying car or, at least, name a planet after them. The good company

---

[1] I was informed complaining is a very common trait among Croats. I suspect it is memetic.

★ ★ ★

⋆ ⋆ ⋆

None of this, of course, would be possible without the support from my Mother.

Jedne noći	**One night**
Te noći pisah sjedeć posve mirno	That night I wrote, sitting all calm,
Da ne bih majci u susjednoj sobi	So next door my mother's dream
Škripanjem stolca u san dirno.	Would suffer no alarm.
A kad mi koja ustrebala knjiga,	And if I had to get a book,
Sasvim sam tiho išao po sagu.	With every movement I would fear,
U svakoj kretnji bila mi je briga	With every single step I took,
Da staričicu ne probudim dragu.	Not to wake my mother dear.
I noć je tekla spokojna i nijema.	The peaceful night thus crept on.
A tad se sjetih da je više nema.	And then I remembered she was gone.

Dobriša Cesarić (translated by Pavle Ninković)