Probing the Hubble Constant: Time Delay Cosmography of SDSS J1433 with the Hubble Space Telescope and the 2.1-Meter Wendelstein Telescope

Giacomo Queirolo

München 2024

Probing the Hubble Constant: Time Delay Cosmography of SDSS J1433 with the Hubble Space Telescope and the 2.1-Meter Wendelstein Telescope

Giacomo Queirolo

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

vorgelegt von
Giacomo Queirolo
aus Turin, Italy

München, den 31/07/2024

Erstgutachter: Professor Dr. Ralf Bender
Zweitgutachter: Professor Dr. Sherry Suyu
Tag der mündlichen Prüfung: 20/09/2024

Contents

Zusammenfassung

Obwohl die ΛCDM-Kosmologie das erfolgreichste kosmologische Modell ist, das uns heute zur Verfügung steht, da es die meisten der beobachteten Phänomene erklären kann, wurde es durch immer mehr Probleme in Frage gestellt. Eine der größten Probleme, sowohl in Bezug auf die Werte als auch auf die Bedeutung des gemessenen Parameters, ist die berüchtigte "Hubble-Spannung". Sie bezeichnet die Unstimmigkeiten zwischen den Messungen der Hubble-Konstante, die die Expansionsrate des Universums beschreibt und einen wichtigen Bestandteil unseres kosmologischen Verständnisses darstellt. In den letzten Jahren sind die Methoden zur Messung von H_0 immer zahlreicher und ausgefeilter geworden, doch während die Unsicherheiten der Messungen abgenommen haben, wurde die Unstimmigkeit nicht gelöst, sondern hat sogar zugenommen. Solche Methoden lassen sich grob in "frühe" und "späte" Proben für H_0 einteilen, was sich ungefähr auf den Zeitpunkt der Entstehung des beobachteten Phänomens bezieht. Während "frühe" Proben, die z.B. auf dem kosmischen Mikrowellenhintergrund basieren, stark von dem angenommenen kosmologischen Modell abhängig sind, sind "späte" Proben im Allgemeinen modellunabhängig, aber anfälliger für systematische Fehler in den Messungen. In diesem Zusammen- hang gehört die "Time-delay"-Kosmographie zu den "späten" Proben, die H_0 messen kann, ohne dass eine Kalibrierung erforderlich ist. Diese Analyse basiert auf dem in der Allgemeinen Relativitätstheorie gut erprobten Phänomen des starken Gravitationslinseneffekts. Bei einer variablen Lichtquelle im Hintergrund und einer starken Gravitationslinse im Vordergrund kann die Zeitverzögerung zwischen den Mehrfachbildern der gelinsten Quelle durch Messung und Analyse ihrer Leuchtkraft über die Zeit gemessen werden. Eine Linsenmodell-Analyse des Systems kann dann das Massenprofil der Linse einschränken. Die Kombination aus beiden Informationen kann dann verwendet werden, um den Hubble-Parameter einzuschränken. In dieser Arbeit habe ich diese Analyse auf der Grundlage von *Hubble Space Telescope* Archivdaten und einer Beobachtungskampagne mit dem 2,1-Meter-Teleskop am Wendelstein durchgeführt. Ich habe die weltraumbasierten Daten verwendet und dabei die verfügbaren Filter sowie deren höhere Auflösung genutzt, um die Linsenmasse zu modellieren. Dadurch konnte ich das Fermat-Potential mit einer Präzision von 3 % bestimmen.

Ich habe die Daten der *Wendelstein*-Beobachtungskampagne verwendet, um die Lichtkurven der gelinsten Quellen zu erstellen, die die Zeitverzögerung einzuschränken, die je nach Bildpaar mit einer Genauigkeit von 8% bis 15% ermittelt wurde.

Anschließend habe ich die Ergebnisse nach einem Bayes'schen Ansatz kombiniert und so eine Beschränkung für H_0 von $71.3^{+5.0}_{-4.5}\,\frac{\mathrm{km}}{\mathrm{Mpc\,s}}$ mit einer statistischen Unsicherheit von $\sim 6,7\%$ erreicht. Diese Arbeit wurde größtenteils unabhängig von größeren Kollaborationen wie TDCOSMO durchgeführt, was eine unvoreingenommene Validierung der Methodik ermöglicht. Darüber hinaus ist das Ergebnis ein Beweis für die Stärke des Wendelstein-Observatoriums, das als zuverlässiges Hilfsmittel für die "Time-delay"-Kosmographie oder ähnliche Projekte, die hochqualitative Daten mit hoher Quantität erfordern, betrachtet werden sollte.

Abstract

While ΛCDM cosmology is the most successful cosmological model at our disposal today, being able to explain most of the observed phenomena, it has been challenged by more and more tensions. One of the greatest, both in terms of numerical tension and of the importance of the parameter measured, is the infamous Hubble tension. This refers to the disagreement between measurements of the Hubble constant, which describes the rate of expansion of the Universe, a cornerstone of our cosmological understanding. In recent years the methods for measuring H_0 have grown in number and sophistication, and yet, as the uncertainties of the measurements have decreased, the tension has not been solved; in fact, it has increased.
Such methods can be roughly divided between "early" and "late" probes of H_0, approximately referring to the time of origin of the phenomenon observed. While "early" probes, based for example on the cosmic microwave background, are strongly dependent on the assumed cosmology, "late" probes are generally model-independent but are more susceptible to systematic errors in the measurements. In this context, the time delay cosmographic method is a "late" time probe which can measure H_0 directly, without requiring any calibration. This analysis is based on the well-tested general relativity phenomenon of strong gravitational lensing. Given a background variable source and a foreground strong gravitational lens, the time delay between the multiple lensed images can be measured by monitoring and analysing their luminosity over time. A separate modelling analysis of the system can then constrain the mass profile of the lens. The two combined information can then be used to constrain the Hubble constant. In this work, I implemented this analysis based on *Hubble Space Telescope* archival data and a dedicated observational campaign from the 2.1-meter telescope at *Wendelstein*. I employed the space-based data by taking advantage of the multiple filters available and their higher resolution to model the lens mass, obtaining a result with 3% precision on the Fermat potential.
I instead used the data from the *Wendelstein* observational campaign to produce the lightcurves of the image and analyse them in order to constrain the time delay, which was obtained with a precision ranging from 8% to 15% depending on the image pair.
I then combined the results following a Bayesian approach, reaching a constraint on H_0 of $71.3^{+5.0}_{-4.5}\,\frac{\mathrm{km}}{\mathrm{Mpc\,s}}$ with a precision $\sim$ 6.7% considering random uncertainty. Notably, this work has been mostly independent of major collaborations, such as TDCOSMO, thus providing an unbiased validation of the methodology. Furthermore, the result is proof of the capabilities of the

Wendelstein observatory, which should be considered a reliable asset for time delay cosmography or similar projects that require high-sampling, high-quality data.

List of Figures

List of Tables

Chapter 1

Introduction

The idea of light interacting gravitationally with matter was entertained as early as 1704 by none other than Sir Isaac Newton Newton (1704) and was further studied by several great physicists: Cavendish (unpublished manuscript, 1784), Laplace Laplace marquis de (1813) and Soldner Soldner von (1804), among others. However, it was only with the publication of Einstein's theory of General Relativity (GR, Einstein, 1916) that this concept was revisited and correctly defined. Such a phenomenon was dubbed, somewhat inappropriately Lodge (1919), "gravitational lensing" (GL) and was first observed during the famous 1919 eclipses, being the second observational proof of GR. The discovery was so sensational that the New York Times reported it on the 10th of November with the headline "Lights All Askew in the Heavens" Times (1919). However, after such initial success, this field of research was met with little optimism, as even Einstein, as late as 1936, remarked that "there is no great chance of observing this phenomenon" Einstein (1936). Fortunately, Zwicky found one of the early hints of the existence of dark matter (DM) by applying the virial theorem to nearby clusters of galaxies, and thus considered the mass of galaxies, at the time referred to as "nebulae", to be underestimated. He therefore painted a much brighter picture of the possibility of observing GL, and, even more importantly, suggested some key points of interest for which this field of research was worthy of study. Quoting Zwicky (1937): "The discovery of images of nebulae which are formed through the gravitational fields of nearby nebulae would be of considerable interest for a number of reasons.

1. It would furnish an additional test for the general theory of relativity.
2. It would enable us to see nebulae at distances greater than those ordinarily reached by even the greatest telescopes. Any such extension of the known parts of the universe promises to shed very welcome new light on a number of cosmological problems.
3. The problem of determining nebular masses at present has arrived at a stalemate. [...] Observations on the deflection of light around nebulae may provide the most direct determination of nebular masses and clear up the above-mentioned discrepancy"

His vision proved to be accurate; today, not only is gravitational lensing a well-established and flourishing research field, but these points still encapsulate the focus of most of the GL studies.

Starting from the latter point, GL has been effectively utilized to investigate the mass distribution of galaxies and larger cosmic structures. Its primary aim has been to constrain Dark Matter (DM) candidates and properties, ranging from cold (Narayan, White, 1988; Kochanek, 1994), warm (Gilman et al., 2019) or a mixture of both, as well as self-interacting (SIDM, Gilman et al., 2021; Despali et al., 2019) and "fuzzy" DM (Laroche et al., 2022). Multiple studies also focused on DM small-scale substructures, such as angular structures in the lens systems (Vegetti, Koopmans, 2009; O'Riordan, Vegetti, 2024) and line-of-sight haloes (Dhanasingham et al., 2023; Hogg et al., 2023). Considering the second point of Zwicky's list, strong gravitational lensing has been used to study sources which would be unobservable directly, due to intrinsic brightness and distance. Recently, the farthest observed galaxies have been studied thanks to the combination of exquisite resolution and depth of the James Webb Space Telescope (Gardner et al., 2006, JWST) combined with the natural high magnification of cluster-scale lenses (e.g. Roberts-Borsani et al., 2023; Hsiao et al., 2023). The morphology and kinematics of galaxies at high redshifts would also not be observable if it were not for strong lensing (e.g. Rizzo, 2020; Amvrosiadis et al., 2024).

Finally, the first of Zwicky's points proved particularly fruitful, as multiple and different avenues of research are built upon it. For example, tests of GR have been developed by constraining the post-Newtonian parameter ($\gamma_{\rm PPN}$, Thorne, Will, 1971) in galaxy-galaxy strong lensing, indicating the ratio of the spatial curvature potential and the Newtonian potential. In standard GR, these are assumed to be identical, but strong GL can be used to verify such a hypothesis (e.g. Collett et al., 2018; Yang et al., 2020). Also tied to this first point are the constraints of cosmological parameters such as the matter density parameter $\Omega_{\rm m}$ and the dark energy equation of state w which are possibly constrained from "compound lensing", i.e. a single lens with multiple lensed sources, which is a rare occurrence for single galaxies (see e.g. Ballard et al., 2024), but is more common in galaxy clusters (e.g. Caminha et al., 2022), or H_0, the Hubble parameter. The latter will be the focus of this study, and its measurement is referred to as Time Delay Cosmography (TDC) or Time Delay Cosmographic analysis. I will describe the details of the theoretical framework of this method in detail in Chapter 2, while I present here a qualitative description of the method. The mass of a gravitational lens warps space-time around itself, altering the light-path. This warping causes the spatial length of the light-path to be elongated compared to the unperturbed path, resulting in a geometrical delay. Additionally, the gravitational lens induces time dilation in its surrounding space-time (gravitational or Shapiro's delay, see Shapiro, 1964). As a result, the light experiences a further delay due to this time dilation, causing its proper time to be dilated while traversing the lens environment.

The combination of these two effects results in the observed light being delayed with respect to the unperturbed light ray. The latter is unfortunately impossible to observe due to the presence of the lens. However, in the case of strong gravitational lenses (SGL or simply SL), the light of the source is deflected into multiple images, which are delayed differently from one to the other. Thus, a relative time delay between the arrival time of light is present for each pair of images. Such a time delay is dependent on the mass distribution of the lens, the geometrical configuration of the system (i.e. the relative positions of the source, the lens and the observer) and the cosmology. The time delay can be measured if the source presents a variability in luminosity over time, such as for Quasi-Stellar Objects (QSOs) or SuperNovae (SNae). In such cases, observational campaigns can be carried out to record the luminosity variation in each image with respect to time, indicated as lightcurves. The lightcurve variations are then correlated between the images, identifying their relative time delay and magnification shift (or, equivalently, their flux ratio). On the other hand, the mass distribution and geometrical configuration of the system can be obtained by modelling the system based on high-resolution observations. Ancillary information, such as integrated spectroscopy of the lens and characterisation of the lens environment, is furthermore required to break the Mass-sheet degeneracy (Suyu et al., 2014). These measurements are then used to constrain cosmological parameters, in particular, the Hubble parameter H_0.

This is a fundamental parameter for our understanding of the Universe, and measuring it accurately and precisely has become of great importance for the astrophysics community. Not only because its value affects many aspects of our cosmological understanding, from the size and age of the universe, to its critical density, but also due to the divergence in results from different measurements. This is commonly referred to as the "Hubble tension" Di Valentino et al. (2021a), one of the open and heavily debated crises in cosmology. It results from the tension, at present larger than 5σ Perivolaropoulos, Skara (2022), between its measurements. If these results are not to be corrected by unknown or underestimated systematics, they would indicate a significant failure of our current model of the Universe, and lead us to the discovery of a new, more exact understanding of it. A great effort within the scientific community is being made to try and solve this tension, either by improving the precision of the measurement or by defining a new cosmological model. In this context, this work is particularly timely as it provides an additional measurement of H_0 using the relatively new Time Delay Cosmography (TDC) method, which enables a direct determination of this parameter. Importantly, this research was conducted independently of larger collaborations, offering an external verification of the methodology employed.

Furthermore, this work will rely on the data obtained during a dedicated observational campaign carried out at the *Wendestein* observatory to measure the time delay. The results of this work, therefore, reinforce the argument that this is a state-of-the-art facility for cosmological observations and, more generally, a well-suited observatory for high-sampling, high-resolution campaigns. This will prove always more important with the advent of the "Big Data Era" of astrophysics, where large-scale surveys are expected to increase by orders of magnitude the number of suitable targets Collett (2015).

Chapter 2

Gravitational Lenses in Time Delay Cosmography

Cosmology is a fundamental branch of science that has been investigated from the early days of written history, but due to the lack of a sound theoretical framework and observational capabilities, it was rather limited to speculation, often on par with philosophy and religion. The birth of modern cosmology had to wait until a little more than a century ago, with the scientific revolution that Einstein's Relativity brought to physics, in 1916 (Einstein, 1916). Since then, our understanding of the cosmos has grown exponentially, in tandem with our observational capabilities: it took approximately 10 years to observe the expansion of the universe (Lemaître, 1927; Hubble, 1929), another four decades to observe the relic of the Big Bang as cosmic microwave background (CMB, Penzias, 1968), and then thirty years to discover that the universe's expansion rate is positive, meaning that we are experiencing an accelerated expansion (Riess et al., 1998).
Our understanding of the cosmos is now at a fairly mature stage, being based on the ΛCDM "standard model" (Peebles, 1984; Peebles, Ratra, 2003; Carroll, 2001). Its name refers to its two most critical energy components: cold dark matter (CDM, Zwicky, 1933; Freeman, 1970; Rubin et al., 1980), which dominates the matter component, and the cosmological constant (indicated by Λ, Carroll, 2001), which is associated with the dark energy (DE), which in turn consists of 68% of the total energy of the universe and is the leading cause for the accelerated expansion Peebles, Ratra (2003); Weinberg (1989). This has been a very successful model, as it is the simplest mathematical framework based on only 6 free parameters Aghanim et al. (2020): the cold dark matter density $\omega_c = \Omega_c h^2$, the baryon density $\omega_b = \Omega_b h^2$, the scalar spectral index n_s, the observed angular size of the sound horizon at recombination θ_{MC}, the initial super-horizon amplitude of curvature perturbations A_s, and the reionization optical depth τ. This model is able to describe most of the modern cosmological observations: the presence of the CMB, the accelerated expansion of the universe, the large-scale structure distribution of galaxies and the chemical abundance observed today.
However, ΛCDM is far from perfect, as it suffers from known theoretical problems, such as its "cosmic coincidence" (also referred to as the "Why now?" problem, i.e. the fact that matter and dark energy have a comparable energy density today, $\rho_m \sim \rho_\Lambda$, Arkani-Hamed et al., 2000;

Velten et al., 2014) and the disagreement between the vacuum energy density derived from observations and the one expected from Quantum Field Theory Weinberg (1989); Martin (2012); Burgess (2015). Moreover, while the model accounts for dark matter and dark energy, there is still no definitive theory describing the physical nature of either of them Bertone, Hooper (2018); Amendola, Tsujikawa (2010). Lastly, for a few decades, this model has been also facing numerous observational crises (see Abdalla et al., 2022, for a detailed review), first and foremost being the observed tension between different measurements of the expansion of the universe, the Hubble constant H_0, dubbed the "Hubble tension" (Riess et al., 2016; Verde et al., 2019, further discussed in section 2.1.2), followed by the "S8 tension" or "growth tension" Anchordoqui et al. (2021). The latter, however, might result to be an apparent tension (Sánchez, 2020) or might have been proved to be due to systematics (Ghirardini et al., 2024), although further independent measurements are required. Many other sources of tension are present, as reviewed in Perivolaropoulos, Skara (2022), but will not be discussed here: the CMB anisotropy anomalies, the cosmic dipoles, the Baryon Acoustic Oscillations (BAO) curiosities, the Cosmic Birefringence, and others. These tensions are to be considered great opportunities; indeed, if those can not be explained from systematics in the measurements or other unknowns or underestimated sources of error in the analysis, which all studies have thoroughly researched for, the only remaining explanation is to be found in an incoherent model. This would point the research to new physics and a deeper understanding of the universe. However, at present, no systematic appears to be able to explain all tensions, and no model has reached the capabilities of ΛCDM (e.g. Knox, Millea, 2020), as none can simultaneously explain the tension and all other observed cosmological features.
In this study, we will focus on the Hubble tension specifically, described in Section 2.1.2, and measure the Hubble parameter with the Time Delay Cosmographic method (TDC, see Section 2.3), first introduced by Refsdal as early as 1964 Refsdal (1964). In the following Sections, I will introduce a brief background of cosmology (Section 2.1). I will then give a general introduction to gravitational lensing (Section 2.2), and a more detailed discussion of how this phenomenon is taken advantage of in the TDC framework (Section 2.3). In this section, I will also discuss some details of the model peculiar to this study, such as the physical model of QSO (Section 2.3.4) and the effect of microlensing (Section 2.3.5).

2.1 Cosmological Background

The framework of modern cosmology lies in Einstein's field equations Einstein (1936). The following section is a general overview of the cosmological implications of these equations, which are now well-known and established in the field. For a more in-depth overview, refer to Peebles (1993). Einstein's field equations relate the geometric terms of spacetime with its energy content:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{2.1}$$

This is a tensor equation, where $R_{\mu\nu}$ is the Ricci tensor, $g_{\mu\nu}$ is the metric, $\mathcal{R}$ is the Ricci scalar or scalar curvature, which is the contraction of the Ricci tensor given the metric: $\mathcal{R}g^{\mu\nu}R_{\mu\nu}$ and Λ is the cosmological constant. For the right-hand side of the equation, encoding the energy and

mass content, G is the gravitational constant and $T_{\mu\nu}$ is the stress-energy tensor. Note that for different conventions, the constant factors and signs might change, but the general form of the equation will remain.
ΛCDM cosmology is then defined, given this equation, by assuming at cosmological scales the following:

- The cosmological principle: the universe is isotropic and homogeneous.
- The presence of four components of the energy content of the universe: matter (comprising both cold dark matter and baryonic matter), radiation, the cosmological constant, encoding the effect of dark energy, and the curvature.
- The metric being represented by the Friedmann-Lemaître-Robertson-Walker metric (FLRW) metric (this is indeed a consequence of the cosmological principle).
- The different components, matter, radiation and dark energy, are considered to behave as a perfect fluid.

The FLRW metric is described in general in the following form:

$$\mathrm{d}s^2 = -\mathrm{d}t^2c^2 + a^2(t)\cdot\left(\frac{1}{1-kr^2/R_0^2}\mathrm{d}r^2 + r^2\mathrm{d}\omega^2\right). \tag{2.2}$$

Here c is the speed of light, $\mathrm{d}\omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$ takes into consideration all angle dependence and k is the curvature sign and can take values between -1, 0 and +1, to which corresponds the geometry of space: open, flat or closed, respectively. Correspondingly, R_0 is the curvature scale or radius, related to the spatial curvature $K = \frac{k}{R_0^2}$. $a(t)$ is the scale factor, which depends only on cosmological time and encodes the scale of the universe at a given time (thus its expansion or contraction over time). We define for convenience $a(t_0) = a_0 = 1$, where t_0 indicates the cosmic time today; thus $a(t)$ serves as a relative scale with respect to the size of the Universe today. I furthermore follow the usual convention of indicating the given parameter of today with the “0” as subscript.
Considering the FLRW metric, the Equations 2.1 can be solved and rewritten as Friedmann’s equations :

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho - \frac{c^2k}{a^2R_0^2} + \frac{\Lambda c^2}{3} \tag{2.3}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3\frac{p}{c^2}\right) + \frac{\Lambda c^2}{3}, \tag{2.4}$$

where Equation 2.3 is obtained from the time-time component of Equation 2.1 (i.e. $\mu,\nu = 0,0$), while 2.4 is obtained from the trace of the tensor (i.e. $\mu,\nu = i,i$). The Hubble parameter is usually introduced here by defining $H = \frac{\dot{a}}{a}$, where $\dot{a} = \frac{\mathrm{d}a}{\mathrm{d}t}$. . Given the dependence on time of

both $\dot{a}(t)$ and $a(t)$, note that H is also a function of time. It is then a misnomer to define it as the Hubble *constant*, although this can be referred to as the value of the Hubble parameter today, $H(t_0) = H_0 = \frac{\dot{a}_0}{a_0} = \dot{a}_0$, as we should do in this work.
The Friedmann equations define the expansion rate of the universe and its acceleration as a function of its energy content. In order to solve them, we need to define the dependence of pressure and density of the universe's components (matter, radiation and dark energy) with respect to time. We then make use of the assumption that such components behave as perfect fluids. Thus, we can define their equation of state as $p = w\rho c^2$, where p indicates the pressure and w is assumed to be constant. It is worth mentioning that this is not an unchallenged assumption, and expansions of ΛCDM models have been considered where w is time-dependent, usually applied to DE and thus referred to as "dynamical DE" models, e.g. Linder (2003).

Given the cosmological principle, these components are independent of the positions (else the homogeneity would be broken) and are only dependent on cosmic time. Matter is considered to be pressure-less in cosmological scales, thus $w_m = 0$ (for this reason referred to as "cold dust"). For radiation, w_γ is instead 1/3. For dark energy, $w_\Lambda = -1$, although given the discussed effect of dark energy, this value is still under investigation. Note that the energy density relative to the cosmological constant is defined as $\rho_\Lambda = \frac{c^2\Lambda}{8\pi G}$.
Given this assumption, it can be derived that the solution $\rho(a) \propto a^{-3(1+w)}$. For each energy component and its corresponding factor w, this equation tells their cosmic evolution: cold matter dilutes as $\rho_m \propto a^{-3}$, as expected, while the density of the cosmological constant is indeed constant over time. The energy density of radiation instead dilutes as $\rho_\gamma \propto a^{-4}$, which is due to the additional loss of energy of the expanding space over which the electromagnetic waves are propagating. This is an observable effect, as the wavelength of light of a distant object is "stretched" by a factor $1/a$ from the moment it is emitted to now. This is the cosmological redshift z, which is defined as $z = 1/a - 1$. This is the actual observable of cosmological time, as it can be, in principle, directly observed from the spectra of the object, and is therefore often used in cosmology rather than the cosmic time or even the scale factor.
It is customary to redefine the first Friedmann equation 2.3 by defining the critical density $\rho_c = \frac{3H^2}{8\pi G}$, and introducing the non-dimensional density parameters $\Omega_i = \rho_i/\rho_c$. We then obtain $\sum_i \Omega_i = 1$, where i varies between matter, radiation, dark energy and curvature, for which we define $\Omega_k = -\frac{kc^2}{\dot{a}^2 R_0^2}$. If we compare that to today's density parameter, we obtain the following:

$$\frac{H^2}{H_0^2} = \Omega_{0,k} a^{-2} + \sum_i \Omega_{0,i} a^{-(1+w_i)}. \tag{2.5}$$

This latter equation, referred to as the dimensionless Friedmann equation, describes H as a function of a, and it's then straightforward to redefine it as a function of z:

$$H = H_0 \sqrt{\sum_i \Omega_{0,i} (1+z)^{1+w_i}}, \tag{2.6}$$

where we have incorporated the curvature in the summation over i, with $w_k = 1$. The square root on the left-hand side is sometimes defined as the function $E(z)$, thus having $H(z) = H_0 E(z)$.

2.1.1 Distances in Cosmology

I will conclude this brief introduction by discussing the topic of how to measure distances in an expanding and generally non-flat space. A more in-depth presentation can be found in Dodelson, Schmidt (2020). From the FLRW metric defined in Equation 2.2, the proper distance can be defined as $D(t) = a(t)\chi$. This is the physical distance between two objects at a given time, separated by the comoving distance χ. This depends on the sign curvature k of the universe, as

$$\chi = \begin{cases} R_0 \sin^{-1}(r/R_0) & \text{if } k = +1 \\ r & \text{if } k = 0 \\ R_0 \sinh^{-1}(r/R_0) & \text{if } k = -1, \end{cases} \tag{2.7}$$

where r is the radius obtained from the spherical coordinates as described for the FLRW metric in 2.2 and R_0 is the curvature scale. It can be seen that, given different geometries, the distances would be affected, with a closed universe ($k = +1$) having a larger proper distance to an object than its coordinate distance r. This is because the geometry is analogous to a sphere, where the distance along the surface (corresponding to χ) is longer than the straight-line distance (analogous to r). On the contrary, an open universe ($k = -1$), which is analogous to a hyperbolic space, would have a smaller proper distance than the coordinate distance. Finally, $k = 0$ would represent a flat universe, where no curvature could alter the relationship between the coordinate distance and the proper distance. The comoving distance can be related to the scale factor by considering the travel time of light and correcting for the expansion of the universe along the path. Thus considering two points at cosmic time t_0 and t_1, we obtain

$$\chi = c \int_{t_0}^{t_1} \frac{\mathrm{d}t'}{a(t')} = c \int_{a_0}^{a_1} \frac{\mathrm{d}a'}{\dot{a}'a'}. \tag{2.8}$$

Here we operated a simple change of variable by considering the time derivative of a, where $a(t_i) = a_i$ for $i = 0, 1$. It is usual to consider an observer at $t_0 = 0$, thus $a_0 = 1$, and a source at $a_1 = a$, thus their comoving distance becomes $\chi = c \int_1^a \frac{\mathrm{d}a'}{\dot{a}'a'}$. Considering this definition of χ, equation 2.5 and the corresponding definition of the function $E(a) = H(a)/H_0$, we can obtain:

$$\chi = \frac{c}{H_0} \int_1^a \frac{\mathrm{d}a'}{a'^2 E(a')} \tag{2.9}$$

$$= \frac{c}{H_0} \int_1^a \frac{\mathrm{d}a'}{a'^2} \frac{1}{\sqrt{\sum_i \Omega_{0,i} a'^{-(1+w_i)}}} \tag{2.10}$$

$$= \frac{c}{H_0} \int_0^z \frac{\mathrm{d}z'}{\sqrt{\sum_i \Omega_{0,i} (1+z')^{(1+w_i)}}}. \tag{2.11}$$

Here we take advantage of the shortened version of the density parameters as introduced in Equation2.5, and we apply a change of variable using the relation between redshift and scale factor $a = \frac{1}{1+z}$.

However, the proper distance is not observable due to the expansion of the universe. Instead, cosmologists have introduced two new definitions of distances: the angular diameter distance D_θ and the luminosity distance D_L. Both rely on the idea of a certain type of standard reference to measure the distance from an object: the first a "standard candle", while the second a "standard ruler". The luminosity distance D_L is defined by considering the relation between flux F and luminosity L, thus $D_\mathrm{L} = \sqrt{\frac{\mathrm{L}}{4\pi\mathrm{F}}}$. Considering the loss of energy due to redshift, it can then be demonstrated that $D_\mathrm{L} = (1+z)\chi$. The angular diameter distance D_θ, instead, relates the apparent size of an object in the sky $\delta\theta$ (measured as an angle, from which the name and the variable θ) with respect to its true, physical size, δs:

$$D_\theta = \frac{\delta s}{\delta\theta}. \tag{2.12}$$

Given the expansion of the universe, an observer at $a = 1$ would observe the object of physical size δs at scale factor a having a size of $\delta\theta = \frac{\delta s/a}{r}$, where r is the coordinate distance. Thus

$$D_\theta = ar. \tag{2.13}$$

Generally, cosmological observations are done using light, which follows a null geodesic: $ds = 0$. To avoid confusion, note that $ds = 0$ refers to the spacetime interval of the metric described in 2.2 along the path joining the observer and the source, which is different from δs, which is the size of the source itself. Following equation 2.2, $c\mathrm{d}t = a\sqrt{1/(1 - kr^2/R_0^2)}\mathrm{d}r = a\mathrm{d}\chi$ (where we have considered a frame of reference such that the angular component $\mathrm{d}\omega = 0$). We can then define $r = r(\chi)$ by inverting the given equation 2.7, depending on $k = +1, 0, -1$:

$$r = \begin{cases} R_0 \sin(\chi/R_0) & \text{if } k = +1 \\ \chi & \text{if } k = 0 \\ R_0 \sinh(\chi/R_0) & \text{if } k = -1. \end{cases} \tag{2.14}$$

Now, considering the definition of the density parameter for the curvature, we can rewrite the curvature scale R_0 as $R_0 = \frac{c}{H_0}\sqrt{\frac{k}{-\Omega_{k,0}}}$. Note that Ω_k has the opposite sign with respect to k, thus the factor under the square root is always positive and R_0 is always a real number. For simplicity, we can the write $R_0 = \frac{c}{H_0}\sqrt{|\frac{k}{\Omega_{k,0}}|}$. Then, the angular diameter distance can be rewritten as:

$$D_\theta = a \cdot \begin{cases} \frac{c}{H_0\sqrt{|\Omega_{k,0}|}} \sin\left(\chi \frac{H_0\sqrt{|\Omega_{k,0}|}}{c}\right) & \text{if } k = +1 \\ r & \text{if } k = 0 \\ \frac{c}{H_0\sqrt{|\Omega_{k,0}|}} \sinh\left(\chi \frac{H_0\sqrt{|\Omega_{k,0}|}}{c}\right) & \text{if } k = -1. \end{cases} \tag{2.15}$$

Here D_θ is defined between a source at scale factor a and the observer at $a_0 = 1$, but equation 2.15 can be generalised to any two scale factor a_1 and a_2 (a_2 being the source and a_1 being the observer) by substituting the prefactor a with a_2/a_1 in 2.15 and changing the integral limits of equation 2.9.

For the purpose of the analysis carried out in this paper, we shall only consider a flat universe, i.e. $k = 0$ and $r = \chi$, thus :

$$D_\theta = a \cdot \chi \tag{2.16}$$

$$= \frac{c}{(1+z)} \int_0^z \frac{dz'}{H(z')} \tag{2.17}$$

$$= \frac{c}{(1+z)H_0} \int_0^z \frac{dz'}{E(z')}. \tag{2.18}$$

Note how D_θ is therefore linearly dependent on H_0, which is a core component of the Time Delay Cosmography, as described in 2.3.

2.1.2 The Hubble Tension

As mentioned, multiple methods have been developed during the last century to measure the velocity of the expansion of the universe. These can, however, be divided into two main groups, the “late” and the “early” time probes. While intuitively those can be interpreted as the cosmic time at which the given observed phenomenon arises, this late/early dichotomy is more properly defined as the epoch of the ΛCDM model under consideration. These two subsets of measurements, when appropriately converted to the present cosmic time, converge to two ranges of results in tension with one another. This is the Hubble tension, one of the major crises in cosmology at the time of writing, ranging between 4σ to 6σ depending on the dataset under consideration.

I will here present a concise review of the topic, but the interested reader can find a thorough and updated review of the subject in Di Valentino et al. (2021a), from which Figure 2.1 is taken. This summarises the state of H_0 measurements as of 2021. While more measurements have been added to the set, the tension has remained unsolved, and therefore, this plot can be considered as a good starting point to understand the subject.

As can be seen from this Figure, the highest tensions arise with respect to the results obtained from CMB angular scale of fluctuations measurements, in particular from Planck data Aghanim et al. (2020), an “early” time probe. In practice, the most precise “early” time measurements for H_0 are based on the phenomenon of the Baryonic Acoustic Oscillations (BAO). Before recombination ($z > 1100$, Dodelson, Schmidt, 2020) the universe was permeated by plasma (i.e. a fluid of baryons and photons). This plasma oscillated due to spherical sound waves produced by perturbations of the baryon gas and driven by photon pressure. Due to the expansion of the universe and consequent cooling of the plasma, the universe eventually reaches the time of recombination, i.e. when the mean energy of the photons became too low to immediately ionise the baryon, which therefore became bound in atoms. The Universe became “transparent” to the light, and the photons began propagating freely in it. Thus, the sound waves shells of the baryon become frozen. This is visible on the CMB anisotropy spectrum as it defines the scale of the peaks, referred to as the BAO scale, as well as in the galaxy two-point correlation function Eisenstein et al. (2005). This scale can then be measured as an angle in the sky θ_s. This in turn can be converted into a constraint on H_0 by considering the angular diameter distance definition

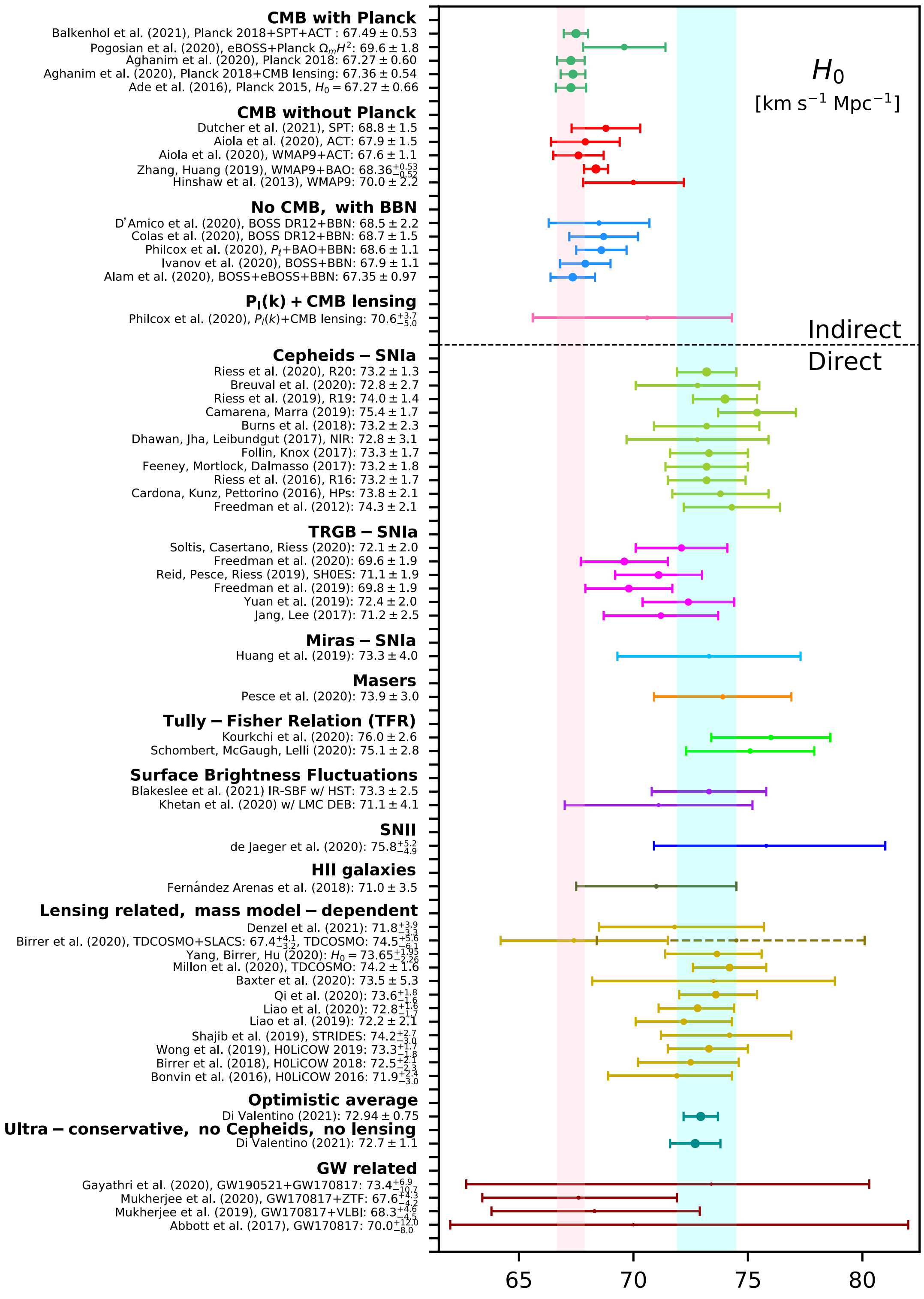

Figure 2.1: Figure 1 of Di Valentino et al. (2021a): "Whisker plot with 68% CL constraints of the Hubble constant H_0 through direct and indirect measurements by different astronomical missions and groups performed over the years. The cyan vertical band corresponds to the H_0 value from SH0ES Team (R20, $H_0 = 73.2 \pm 1.3 \mathrm{km s^{-1} Mpc^{-1}}$ at 68% CL, Riess et al., 2021) and the light pink vertical band corresponds to the H_0 value as reported by Planck 2018 team (N Aghanim et al., 2020) within a ΛCDM scenario."

described in equation 2.13 and treating it as a standard ruler Peebles (1980):

$$\theta_s = \frac{r_s}{d_A}, \tag{2.19}$$

where r_s is the radius of the sound horizon at the time of recombination and $d_A = \frac{D_A}{a} = (1+z)D_A = c\int_0^z \frac{dz'}{H(z')}$ is the comoving angular diameter distance to the last scattering, i.e. $z \sim 1100$. Assuming a cosmological model, d_A can then be written as $d_a \propto H_0^{-1}$. r_s can then be estimated as the distance that the sound could have travelled from the Big Bang, at $z \sim \infty$, to the moment of decoupling from the photons. Thus, measuring θ_s gives a joint constraint on H_0 and r_s, but does not constrain one without the other. Further information on r_s is needed to obtain a result on H_0, which can be obtained from the CMB power spectra Zarrouk et al. (2018) or by deuterium abundance (e.g. Addison et al., 2013). Similarly, θ_s can be constrained from independent data, such as from the imprint of BAO on the galaxy spatial distribution obtained from extensive surveys such as the Sloan Digital Sky Survey (SDSS, York et al., 2000), encompassing the Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al., 2012). More details can be found, e.g. in Perivolaropoulos, Skara (2022). It can be seen that r_s is heavily dependent on the model assumption, in particular on the assumption about the nature of DM and DE when using CMB to constrain it. Thus, this method, while obtaining exquisite precision, is model-dependent as well. Conversely, the "late" probes are mostly direct measurements of H_0, meaning that they are not dependent on the cosmological model. The first and arguably the most successful of these methods is based on the distance-redshift relation, i.e. improving upon the famous first Hubble diagram of 1929 Hubble (1929). This method is based on constructing a "cosmic distance ladder", for which several standard rulers and candles are used to calibrate the distance to the next, such that it becomes possible to measure the distance to objects in the Hubble flow, i.e. which velocity from us is dominated by the expansion of the universe. The most common approach is to use parallax to calibrate the luminosities of standard candles, i.e. objects with fixed luminosities: first, pulsating Cepheid variables Leavitt (1908) and then SuperNovae type Ia (SNIa)Colgate (1979). This method therefore relies in practice on one main theoretical assumption, which is that the standard candles behave consistently everywhere in the Universe, which is to say that the laws of physics are valid everywhere. This assumption is, in practice, a fundamental premise in most, if not all, cosmological studies, as it has its root in the Copernican principle. While sound from the theoretical perspective, this method suffers instead from a large number of possible systematic errors, as any error at any "step of the ladder" will inevitably affect the next and therefore the final constraints. Nevertheless, a great deal of effort has been made since the first modern measurements Freedman et al. (2001), based on the first decades of observations from the Hubble Space Telescope. In particular, the Supernova H_0 for the Equation of State (SH0ES) Project Riess et al. (2022) significantly improved the method by verifying several of the sources of errors (e.g. Riess et al., 2016, 2019, 2023). As shown in Figure 2.1, the multiple variations and improvements upon the model led to small changes in the result, which stayed consistent with each other and thus in strong tension with the "early" probes.

While there exist several other methods of measuring H_0, the most employed and precise of which is considering different types of standard candles. For example, using the Tip of the Red Giant Branch (TRGB), i.e. a subset of the Red Giant stars, as standard candles, combined with SNIa Scolnic et al. (2023). Another method was to take advantage of the Tully-Fisher relation Tully, Fisher (1977). This is a scaling law that correlates the luminosity and the rotational velocities of galaxies, thus rendering them "standardizable" candles Giovanelli et al. (1997).

Lastly, a fairly different and recently developed "late" type method is the Time Delay Cosmographic analysis. Given that this will be the focus of this work, I will describe in detail the methodology in Section 2.3. Here it will suffice to say that this approach provides fairly precise measurements, reaching a few $\frac{\mathrm{km}}{\mathrm{sMpc}}$ of precision, but also suffers from systematic that has to be accounted for, see Section 2.3.3. However, while the explicit analysis of such systematics increased the uncertainty and affected the mean of the result, the tension with the "early" type measurement remained. To conclude, despite the intense effort from multiple groups in various fields of the scientific community, the Hubble tension remains unsolved. While initially the most likely explanation was expected to be measurement systematics, this now seems unlikely (Di Valentino et al., 2021b), as the known sources of error have been considered and minimised, thus only multiple, unrelated errors would be able to explain the current tension. Multiple new cosmological models are under consideration to solve the tension, some of which are considered particularly promising, such as early (Karwal, Kamionkowski, 2016) or dynamical (Chevallier, Polarski, 2001) DE, dark neutrino interaction(Ghosh et al., 2018) or modified gravity (Capozziello, Francaviglia, 2008). However, while most of them solve or diminish the tension, none have yet reached the success of ΛCDM in explaining the vast majority of our observations. It is thus essential to further investigate the tension, adopting a comprehensive and unbiased approach to identify new sources of error and explore potential solutions. With this goal in mind, the current work was devised to follow a new methodology, such as TDC, from an independent perspective. In this context, this approach aimed to verify and replicate the procedures proposed by the major collaborations in the field, namely H0LiCOW-TDCOSMO.

2.2 Gravitational Lensing: Brief Theoretical Overview

The phenomenon of gravitational lensing is usually described by the common idea that a concentration of mass warps the surrounding spacetime, such that its geodesics are distorted and "curved" around the mass, which is then referred to as the lens. Thus, a background source of light, which would appear to a foreground observed at a given position in the sky $\vec{\beta}$, would instead appear at a new position $\vec{\theta}$. A typical representation of this phenomenon is shown in Figure 2.2.

If the distortion due to the mass is severe enough, several geodesics connecting the observer and the source would be present, thus the source would appear as multiple images in the sky. I will now offer a more detailed and quantitative explanation of this phenomenon by presenting a brief overview of the theory of gravitational lensing, specifically regarding its application to Time Delay Cosmography, which is described in Section 2.3. Given the importance of the topic, there exist many manuscripts far more comprehensive and detailed than the following. Interested readers are encouraged to consult sources such as the fundamental monograph Schneider et al. (1992), the lectures of Narayan Narayan, Bartelmann (1996), the recent lectures from Meneghetti Meneghetti (2021), which offer an extended set of `Python` exercises, or the even more recent "Essentials of Strong Gravitational Lensing" Saha et al. (2024). The following section is heavily

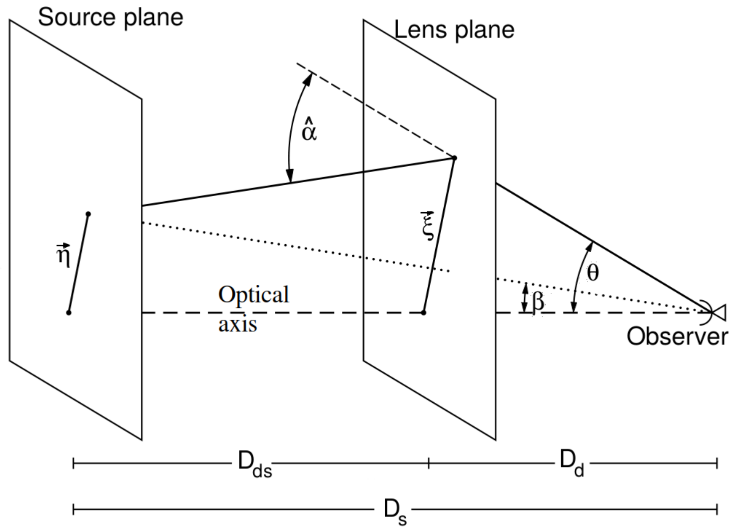

Figure 2.2: Sketch of the gravitational lens effect from Seitz (1998). Note that the distances D_s, D_l and D_ls (the distance between the observer and the source, the observer and the lens and between the lens and the source, respectively) indicated on the bottom are angular diameter distances (see Section 2.1.1).

inspired by this literature.

2.2.1 Light Deflection in General Relativity

In general, the study of the propagation of light in a curved spacetime is a complicated subject, which, however, can be simplified by some assumptions. Firstly, it is usually assumed that the overall geometry is well described by the FLRW metric (see equation 2.2), while the effect of the gravitational lens distortions is limited to a local perturbation. Secondly, the effect of the lens can be considered "weak", in the sense that its Newtonian gravitational potential Φ is assumed to be significantly smaller than c^2, $\Phi/c^2 \ll 1$. This is true for all objects of interest in this study, namely galaxy-scale lenses.
The first assumption, referred to as the "Born approximation", derived from scattering theory Born (1926), allows the lightpath to be considered in three sections: the path from the source to the lens, the path around the lens and the path from the lens to the observer. Given the assumption, light travels the first and last sections in an unperturbed spacetime, while the deflection happens only at the location near the lens. This allows us to compute the deflection along the unperturbed path, which greatly simplifies the approach.
The second assumption further simplifies the calculation of the deflection. First, it has been demonstrated that Fermat's principle, defined in geometrical optics, can be used to study the light deflection in a GR framework Perlick (1990). Specifically, deflection due to a GL can be seen as

a generalisation of the refraction phenomenon. Taking two fixed points A and B, let's consider the extremal travel time along a certain path as $t_{\text{travel}} = \int_A^B \frac{n}{c} \mathrm{d}l$, where n is the "refraction index" of the lens and l is the path. From Fermat's principle, the light will follow the path $\vec{x}(l)$ that minimises t_{travel}, thus

$$\delta \int_A^B n(\vec{x}(l)) \mathrm{d}l = 0. \tag{2.20}$$

Given the second assumption, there exists a local inertial frame of reference where spacetime is flat and described by a simple metric of Minkovsky:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.21}$$

This corresponds to the following line element:

$$\mathrm{d}s^2 = \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu = -(c\mathrm{d}t)^2 + (\mathrm{d}\vec{x})^2. \tag{2.22}$$

Note that this differs from the FLRW presented in equation 2.2 as we are now considering a small region of the universe, where the effect of the cosmic expansion and the curvature are negligible. However, this metric is now perturbed by the (weak) effect of the lens with Newtonian potential Φ, thus we obtain:

$$\eta_{\mu\nu} \longrightarrow g_{\mu\nu} \begin{pmatrix} -(1 + \frac{2\Phi}{c^2}) & 0 & 0 & 0 \\ 0 & 1 - \frac{2\Phi}{c^2} & 0 & 0 \\ 0 & 0 & 1 - \frac{2\Phi}{c^2} & 0 \\ 0 & 0 & 0 & 1 - \frac{2\Phi}{c^2} \end{pmatrix}. \tag{2.23}$$

The lens element is then:

$$\mathrm{d}s^2 = g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu = -(1 + \frac{2\Phi}{c^2})(c\mathrm{d}t)^2 + (1 - \frac{2\Phi}{c^2})(\mathrm{d}\vec{x})^2. \tag{2.24}$$

Light propagates along the null geodesics, thus $\mathrm{d}s = 0$. Therefore

$$(1 + \frac{2\Phi}{c^2})(c\mathrm{d}t)^2 = (1 - \frac{2\Phi}{c^2})(\mathrm{d}\vec{x})^2. \tag{2.25}$$

The speed of light in the section of the path affected by the lens is then

$$c' = \frac{|\mathrm{d}\vec{x}|}{\mathrm{d}t} = c \sqrt{\frac{1 + \frac{2\Phi}{c^2}}{1 - \frac{2\Phi}{c^2}}} \approx c\,(1 + \frac{2\Phi}{c^2}), \tag{2.26}$$

where the assumption $\frac{\Phi}{c^2} \ll 1$ was used to approximate $(1 - \frac{2\Phi}{c^2})^{-1} \approx 1 + \frac{2\Phi}{c^2}$. The "refractive index" of the lens is then the ratio of the speed of light c with c':

$$n = \frac{c}{c'} = \frac{1}{1 + \frac{2\Phi}{c^2}} \approx 1 - \frac{2\Phi}{c^2}, \tag{2.27}$$

where the same approximation is used.

The potential Φ is negative by definition, thus $n \geq 1$ and the speed of light c' in the proximity of the lens is smaller than the speed of light in the vacuum.

We can then derive the deflection angle of the lens from equation 2.20. This is a variational problem. Let's consider a reparametrisation of the path, given by $\mathrm{d}l = |\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}|\mathrm{d}\lambda$, where λ is an arbitrary parameter. Equation 2.20 can then be rewritten as

$$\delta \int_{\lambda_A}^{\lambda_B} \mathrm{d}\lambda n[\vec{x}(\lambda)]|\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}|= 0. \tag{2.28}$$

We can then consider λ as time, then we have $\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda} = \dot{\vec{x}}$ as a generalised velocity and we can rewrite the expression $n[\vec{x}(\lambda)]|\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}|$ as a Lagrangian $n[\vec{x}(\lambda)]|\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}| \equiv \mathcal{L}(\dot{\vec{x}}, \vec{x}, \lambda)$. The Euler equation is then

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial \mathcal{L}}{\partial \dot{\vec{x}}} - \frac{\partial \mathcal{L}}{\partial \vec{x}} = 0. \tag{2.29}$$

The partial derivative of $\mathcal{L}$ with respect to $\vec{x}$ is then

$$\frac{\partial \mathcal{L}}{\partial \vec{x}} = \frac{\partial n|\dot{\vec{x}}|}{\partial \vec{x}} = \frac{\partial n}{\partial \vec{x}}|\dot{\vec{x}}|= \vec{\nabla} n|\dot{\vec{x}}|, \tag{2.30}$$

where we introduced $\vec{\nabla} = \frac{\partial}{\partial \vec{x}}$. For the partial derivative of $\mathcal{L}$ with respect to $\dot{\vec{x}}$, considering $|\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}|= |\dot{\vec{x}}|= (\dot{\vec{x}}^2)^{1/2}$, we obtain

$$\frac{\partial \mathcal{L}}{\partial \dot{\vec{x}}} = \frac{\partial n|\dot{\vec{x}}|}{\partial \dot{\vec{x}}} = n\frac{\partial |\dot{\vec{x}}|}{\partial \dot{\vec{x}}} = n\frac{\dot{\vec{x}}}{|\dot{\vec{x}}|}. \tag{2.31}$$

Note that $\dot{\vec{x}}$ is by definition the tangent vector of $\vec{x}$; thus, it is a tangent vector of the light path. Given that λ is an arbitrary parameter, it can then be defined such that $\dot{\vec{x}}$ is normalised, i.e. $|\dot{\vec{x}}|= 1$. Thus the Euler equation becomes

$$\frac{\mathrm{d}n\dot{\vec{x}}}{\mathrm{d}\lambda} - \vec{\nabla} n = 0. \tag{2.32}$$

Considering the first derivative, we obtain $\frac{\mathrm{d}n\dot{\vec{x}}}{\mathrm{d}\lambda} = \dot{\vec{x}}\frac{\mathrm{d}n}{\mathrm{d}\lambda} + n\frac{\mathrm{d}\dot{\vec{x}}}{\mathrm{d}\lambda} = \dot{\vec{x}}(\vec{\nabla} n \cdot \dot{\vec{x}}) + n\ddot{\vec{x}}$. Thus equation 2.32 can be rewritten as :

$$n\ddot{\vec{x}} = \vec{\nabla} n - \dot{\vec{x}}(\vec{\nabla} n \cdot \dot{\vec{x}}), \tag{2.33}$$

where the first term of the right-hand side is the derivative of n with respect to $\vec{x}$ and the second term is the component of the derivative of n along the light path. Thus, the whole right-hand side is the derivative of *n perpendicular* to the light path. Therefore equation 2.33 becomes:

$$\ddot{\vec{x}} = \frac{1}{n}\vec{\nabla}_{\perp} n = \vec{\nabla}_{\perp} \ln n. \tag{2.34}$$

From equation 2.27, we have $n \approx 1 - \frac{2\Phi}{c^2}$ and considering the weak approximation $\frac{\Phi}{c^2} \ll 1$, $\ln n \approx \frac{-2\Phi}{c^2}$. We then obtain :

$$\ddot{\vec{x}} \approx -\frac{2}{c^2}\vec{\nabla}_\perp\Phi. \tag{2.35}$$

We can then obtain the deflection angle $\hat{\vec{\alpha}}$ (see Figure 2.2) as the difference between the initial and final $\dot{\vec{x}}$ (i.e. before and after the lens), thus by itegrating $-\ddot{\vec{x}}$ along the lightpath λ:

$$\hat{\vec{\alpha}} = \dot{\vec{x}}_{\text{in}} - \dot{\vec{x}}_{\text{out}} = \frac{2}{c^2}\int_{\lambda_A}^{\lambda_B} \vec{\nabla}_\perp\Phi \mathrm{d}\lambda. \tag{2.36}$$

Considering the previously mentioned Born approximation, the computation is simplified, as it is now integrated over the unperturbed path z, which can be defined such that the lens is at $z = 0$. A general example is given considering the deflection angle of a lightray passing at a distance ξ (referred to as the impact parameter) of a point mass of mass M, whose potential will then be $\Phi(\xi, z) = -GM\sqrt{\xi^2 + z^2}$. The deflection angle is then

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{2}{c^2}\int_{-\infty}^{+\infty} \vec{\nabla}_\perp\Phi(\xi, z)\mathrm{d}z = \frac{4GM}{c^2\xi^2}\vec{\xi}. \tag{2.37}$$

Note that, when accounting only for Newtonian relativity, the deflection angle obtained is smaller by a factor of 2; simplistically, this can be attributed to the fact that this approach does not take into account the whole spacetime distortion, but rather only the space distortion.

2.2.2 Deflection by Extended Mass Distributions

When observing the phenomenon of gravitational lensing, the lens is rarely a point mass; while black holes and other very compact objects might fit the description, their effect is usually very strong, thus breaking the assumption $\Phi/c^2 \ll 1$. These objects, therefore, need a more precise framework, which is beyond the interest of this work. Instead, I will be working with a galaxy-scale lens, thus requiring mass distributed in three dimensions. However, when considering the scale of the lightpath, the "depth" of the lens mass distribution, i.e. the length of the lens along the line of sight, is many orders of magnitude smaller. A further approximation can then be made, referred to as the "thin screen" (or "thin lens") approximation. In this optic, the lens is considered to be distributed only in two dimensions, the *lens plane*, ignoring its "depth". Instead of considering its 3-dimensional mass density distribution, $\rho(\vec{\xi}, z)$, the lens is then fully described by its 2-dimensional one, $\Sigma(\vec{\xi}) = \int \rho(\vec{\xi}, z)\mathrm{d}z$, where $\vec{\xi}$ is a vector on the lens plane, while z is the dimension along the line of sight. The distance to the lens is the constant for all lens components, D_{l}. Similarly, the source is approximated to be two-dimensional and to reside within the "source plane", which is located at D_{s} from the observer. The distance between the source and the lens plane is then referred to as D_{ls}. In lensing, such distances are usually expressed as angular diameter distances (see Section 2.1.1). This approach is advantageous because the calculations and measurements within this framework are inherently based on angular variables, as seen in the next Section 2.2.3.

Given the thin lens approximation, the deflection angle can be calculated by summing every contribution of the mass elements on the lens plane, $\Sigma(\vec{\xi})\mathrm{d}^2\vec{\xi}$:

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{4G}{c^2}\int \frac{(\vec{\xi}-\vec{\xi}')\Sigma(\vec{\xi}')}{|\vec{\xi}-\vec{\xi}'|^2}\mathrm{d}^2\vec{\xi}'. \tag{2.38}$$

The deflection angle is then generally a two-dimensional vector field. This can be reduced to one dimension in the case of a perfectly circular symmetric mass distribution, for which the frame of reference can be set to the centre of the mass. Thus the deflection angle is $\hat{\alpha}(\xi) = \frac{4GM(\xi)}{c^2\xi}$, where $M(\xi)$ is the enclosed mass within the radius ξ.

2.2.3 Lens Equation

Let us consider an optical axis, an arbitrary line over which the observer is located perpendicular to the lens and source plane. The angular positions of the components of the system are defined with respect to this axis. The optical axis is shown in Figure 2.2 as the dashed line. Consider a source located at the intrinsic angular position $\vec{\beta}$ in the sky (i.e. if the lens were not present, it would be observed at $\vec{\beta}$). Considering the angular diameter distance from the source plane D_s and given the definition 2.12, we can then obtain the position of the source in the source plane: $\vec{\eta} = \vec{\beta}D_\mathrm{s}$. Assuming a realistic source, such as a galaxy, for example, it will emit light isotropically, i.e. in all directions. Precisely half of this light will reach the lens plane; however, we are now only interested in the light which is finally deflected by such an amount that it reaches the observer. This light reaches the lens plane at the position $\vec{\xi}$ (the impact parameter), is deflected by an angle $\hat{\vec{\alpha}}$ and reaches the observer at an angle $\vec{\theta}$. Considering the distance from the lens plane, $\vec{\xi} = \vec{\theta}D_\mathrm{l}$. In the case of galaxy lensing, the angles $\vec{\theta}$, $\vec{\beta}$ and $\hat{\vec{\alpha}}$ are small and their relation is then described by the *lens equation*:

$$\vec{\theta}D_\mathrm{s} = \vec{\beta}D_\mathrm{s} + \hat{\vec{\alpha}}D_\mathrm{ls}. \tag{2.39}$$

This can be understood as the "projection" of the different angles on the source plane. $\vec{\beta}D_\mathrm{s}$ is the intrinsic position of the source, $\vec{\theta}D_\mathrm{s}$ is the apparent position of the source from the observer perspective, and $\hat{\vec{\alpha}}D_\mathrm{ls}$ is the difference vector between the two points.

The lens equation can be simplified by introducing the reduced angle:

$$\vec{\alpha}(\vec{\theta}) = \hat{\vec{\alpha}}(\vec{\theta})\frac{D_\mathrm{ls}}{D_\mathrm{s}}. \tag{2.40}$$

This is the deflection angle "rescaled" to the frame of reference of the observer. Thus the lens equation becomes

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}). \tag{2.41}$$

This equation can only be solved analytically for very simple mass distributions, while it is usually solved numerically for most concrete applications. In fact, this equation is usually highly non-linear, which leads to the possibility of multiple solutions for θ, i.e. multiple images. Ideally, when modelling a lens (as I will be discussing in Section 4) the positions of the images ($\vec{\theta}$) are measured and the mass distribution is modelled (from which $\vec{\alpha}$ is obtained), thus obtaining for each image the same source position $\vec{\beta}$ through equation 2.41.

2.2.4 Effective Lensing Potential

It is useful to introduce the *effective lensing potential* (or simply *lensing potential*) to characterise the mass distribution:

$$\psi(\vec{\theta}) = \frac{D_{\mathrm{ds}}}{D_{\mathrm{d}} D_{\mathrm{s}}} \frac{2}{c^2} \int \Phi(D_{\mathrm{d}}\vec{\theta}, z)\mathrm{d}z. \tag{2.42}$$

This is an integrated and rescaled version of the Newtonian potential of the lens, and is related to the reduced deflection angle by taking its gradient with respect to $\vec{\theta}$ and the equations 2.36 and 2.40, obtaining :

$$\vec{\nabla}_{\theta}\psi(\vec{\theta}) = D_{\mathrm{d}}\vec{\nabla}_{\perp}\psi(\vec{\theta}) = \frac{D_{\mathrm{ds}}}{D_{\mathrm{s}}} \frac{2}{c^2} \int \vec{\nabla}_{\perp}\Phi(D_{\mathrm{d}}\vec{\theta}, z)\mathrm{d}z = \vec{\alpha}(\vec{\theta}). \tag{2.43}$$

Taking the Laplacian of $\psi(\vec{\theta})$, we obtain a relation to the surface mass density $\Sigma(\vec{\theta})$. The Laplacian of $\psi(\vec{\theta})$ is:

$$\nabla^2_{\theta}\psi(\vec{\theta}) = \frac{D_{\mathrm{ds}} D_{\mathrm{d}}}{D_{\mathrm{s}}} \frac{2}{c^2} \int \nabla^2{}_{\perp}\Phi(D_{\mathrm{d}}\vec{\theta}, z)\mathrm{d}z. \tag{2.44}$$

We then introduce the *convergence* $\kappa(\vec{\theta})$, which is a dimensionless surface mass density:

$$\kappa(\vec{\theta}) \equiv \frac{\Sigma(\vec{\theta})}{\Sigma_{\mathrm{crit}}}, \tag{2.45}$$

where Σ_{crit} is the *surface mass density* and is defined as $\Sigma_{\mathrm{crit}} = \frac{c^2}{4\pi G} \frac{D_{\mathrm{s}}}{D_{\mathrm{l}} D_{\mathrm{ls}}}$. This is a function of the various distances of the system; thus, given a single system, it is constant and characterises it, as it gives a scale to the surface mass density. Considering the Poisson equation for the Newtonian potential, we obtain$\nabla^2\Phi = 4\pi G\rho$. The surface mass density is then

$$\Sigma(\vec{\theta}) = \frac{1}{4\pi G} \int \nabla^2\Phi \mathrm{d}z, \tag{2.46}$$

and the convergence is

$$\kappa(\vec{\theta}) = \frac{1}{c^2} \frac{D_{\mathrm{l}} D_{\mathrm{ls}}}{D_{\mathrm{s}}} \int \nabla^2\Phi \mathrm{d}z. \tag{2.47}$$

We consider then the two dimensional Laplacian ∇^2_{θ} :

$$\nabla^2_{\theta} = \frac{\partial^2}{\partial\theta_1^2} + \frac{\partial^2}{\partial\theta_2^2} = D_{\mathrm{l}}^2(\frac{\partial^2}{\partial\xi_1^2} + \frac{\partial^2}{\partial\xi_2^2}) = D_{\mathrm{l}}^2(\nabla^2 - \frac{\partial^2}{\partial z^2}), \tag{2.48}$$

thus

$$\nabla^2\Phi = \frac{1}{D_{\mathrm{l}}^2}\nabla^2_{\theta}\Phi + \frac{\partial^2\Phi}{\partial z^2}. \tag{2.49}$$

Combining this equation and equation 2.47 yields:

$$\kappa(\vec{\theta}) = \frac{D_{\mathrm{ls}}}{D_{\mathrm{s}} D_{\mathrm{l}}} \frac{1}{c^2} \left(\int_{-\infty}^{+\infty} \nabla^2_{\theta}\Phi \mathrm{d}z + D_{\mathrm{l}}^2 \int_{-\infty}^{+\infty} \frac{\partial^2\Phi}{\partial z^2} \mathrm{d}z \right). \tag{2.50}$$

The second term of the parenthesis vanishes, as the integral results in $\frac{\partial \Phi}{\partial z}$ and, assuming that the lens is gravitationally bound, this value is zero for $z \longrightarrow \pm\infty$. This is then related to the Laplacian of the effective lensing potential through equation 2.44:

$$\kappa(\vec{\theta}) = \frac{1}{2}\nabla_{\theta}\psi(\vec{\theta}). \tag{2.51}$$

Integrating this equation, we obtain

$$\psi(\vec{\theta}) = \frac{1}{\pi}\int \kappa(\vec{\theta}')\ln|\vec{\theta} - \vec{\theta}'|\mathrm{d}^2\theta'. \tag{2.52}$$

From this equation and 2.43 which the reduced deflection angle can be related to the convergence:

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi}\int \kappa(\vec{\theta}')\frac{\vec{\theta} - \vec{\theta}'}{|\vec{\theta} - \vec{\theta}'|^2}\mathrm{d}^2\theta'. \tag{2.53}$$

Note that this equation is consistent with equation 2.38 when considering Σ_{crit}.

Point-Mass Lenses and Einstein Angle

Considering a point-mass lens greatly simplifies the lens equation. Let us set the optical axis such that it crosses the lens plane at the position of the lens. Then, taking the equation 2.47, the deflection angle can be rewritten as:

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{4GM}{c^2}\frac{\vec{\xi}}{|\vec{\xi}|^2} = \frac{4GM}{c^2 D_{\mathrm{l}}}\frac{\vec{\theta}}{|\vec{\theta}|^2} = \hat{\vec{\alpha}}(\vec{\theta}). \tag{2.54}$$

The direction of $\hat{\vec{\alpha}}$ is outward with respect to the lens, and due to the symmetry of the system, we can reduce the problem to one dimension:

$$\hat{\alpha} = \frac{4GM}{c^2 D_{\mathrm{l}}\theta}. \tag{2.55}$$

The lens equation, now one-dimensional as well, can then be written as

$$\beta = \theta - \frac{4GM}{c^2 D_{\mathrm{l}}\theta}\frac{D_{\mathrm{ls}}}{D_{\mathrm{s}}}. \tag{2.56}$$

This equation is quadratic in θ, thus, for every β, there exist two solutions θ, thus two images. Moreover, we can introduce θ_E, the *Einstein radius*, as

$$\theta_E \equiv \sqrt{\frac{4GM}{c^2}\frac{D_{\mathrm{ls}}}{D_{\mathrm{l}}D_{\mathrm{s}}}}. \tag{2.57}$$

The equation 2.56 is then simplified as $\beta = \theta - \frac{\theta_E^2}{\theta}$. If we then consider $\beta = 0$, i.e. the source is located exactly behind the centre of the lens, the solutions for θ are $\theta_{\pm} = \pm\theta_E$, which corresponds to a perfect circle of radius θ_E around the centre of the lens. This phenomenon, while surprising and rare, has been observed and is called Einstein's ring (e.g. see Figure 2.3).

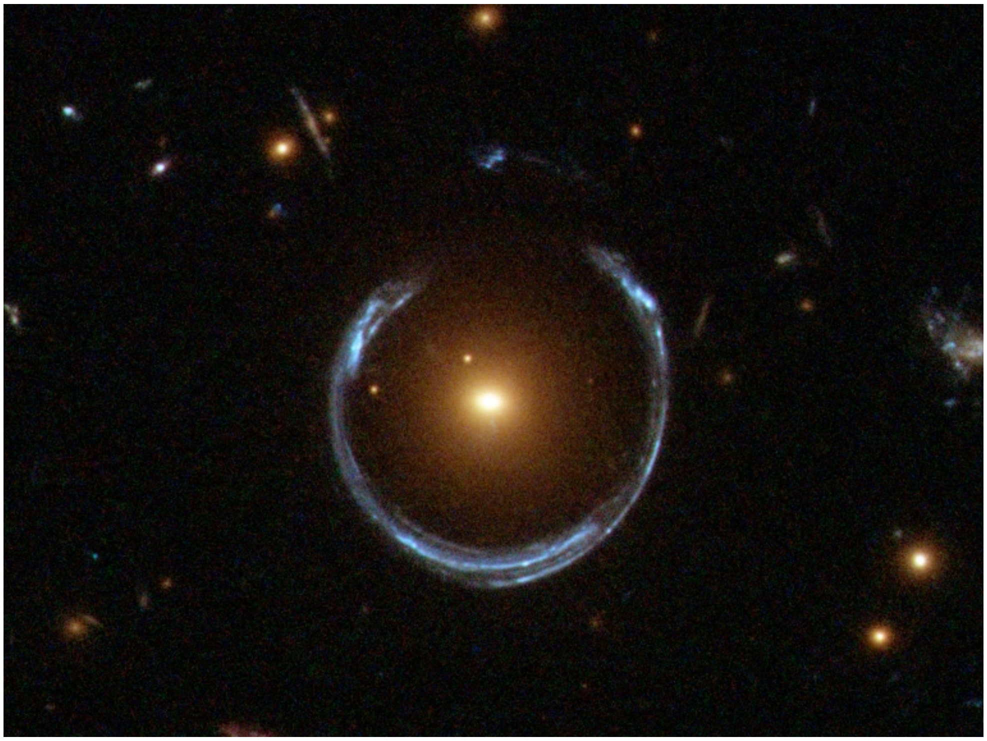

Figure 2.3: LRG 3-757 (nicknamed "Cosmic Horseshoe") is a famous example of an almost complete optical Einstein ring around a luminous red galaxy (LRG) which acts as a lens. The image is a combination of multifilter exposures taken with the Hubble Space Telescope's Wide Field Camera 3.

Moreover, while introduced in the context of point-masses, θ_E can be generalised to any kind of lens and is a useful parameter as it gives a general idea of the size of the lens.

2.2.5 Jacobian Matrix

The lens effect on the background image is effectively a distortion. This can be studied by taking the mapping from the source plane to the lens plane, $\frac{\partial\vec{\beta}}{\partial\vec{\theta}}$, which is the Jacobian matrix A:

$$A = \frac{\partial\vec{\beta}}{\partial\vec{\theta}} = \left(\delta_{ij} - \frac{\partial\alpha_i(\vec{\theta})}{\partial\theta_j}\right) = \left(\delta_{ij} - \frac{\partial^2\psi(\vec{\theta})}{\partial\theta_i\partial\theta_j}\right), \tag{2.58}$$

where the letters i and j indicate the components of $\vec{\theta}$ on the lens plane and the steps obtained through the equation 2.41 and the equation 2.43. Thus, the Jacobian components can be written as second partial derivatives of the effective lensing potential, which we abbreviate as $\frac{\partial^2\psi(\vec{\theta})}{\partial\theta_i\partial\theta_j} = \psi_{ij}$.

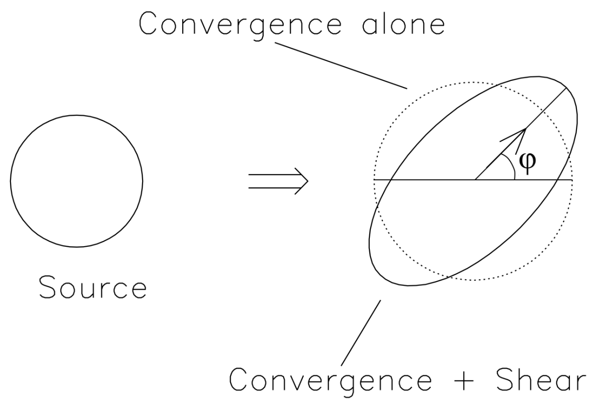

Figure 2.4: Sketch describing the effects of shear and convergence on a spherical source (image taken from Narayan, Bartelmann, 1996, , Figure 13)

From equation 2.51, we can rewrite the convergence as $\kappa = 1/2(\psi_{11} + \psi_{22}) = 1/2\,\mathrm{tr}\,\psi_{ij}$. We then introduce the *shear tensor*, which is defined by its two components $\gamma_1(\vec{\theta})$ and $\gamma_2(\vec{\theta})$:

$$\gamma_1(\vec{\theta}) = \frac{1}{2}(\psi_{11} - \psi_{22}) \equiv \gamma(\vec{\theta})\cos[2\phi(\vec{\theta})] \tag{2.59}$$

$$\gamma_2(\vec{\theta}) = \psi_{12} = \psi_{21} \equiv \gamma(\vec{\theta})\sin[2\phi(\vec{\theta})], \tag{2.60}$$

where the magnitude γ and the orientation ϕ are other common reparametrisation of the shear tensor.
The Jacobian matrix can then be written as

$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}. \tag{2.61}$$

This explains the effects of convergence and shear. κ does not affect the shape, as it is isotropic, but rather scales the image by a constant factor $1/(1-\kappa)$. Instead, the shear deforms the shape of the image along the axis ϕ by elongating it along this direction and shrinking it along the perpendicular direction. A sketch of the effect is shown in Figure 2.4.

2.2.6 Magnification

A fundamental aspect of gravitational lensing is that, according to Liouville's theorem (Carroll, 2019), the surface brightness of the source is preserved during the lens effect, as the gravitational effect of the lens does not absorb or emit photons. This is not true for physical lenses, which both absorb and emit light; however, in most cases of interest, the gravitational effect dominates over such "baryonic" effects. Thus, considering a solid angle element of the source plane $\delta\beta^2$, mapped by the lens into $\delta\theta^2$, will change in shape but not in surface brightness. This effectively means that the image is magnified (or demagnified), as shown in Figure 2.5.

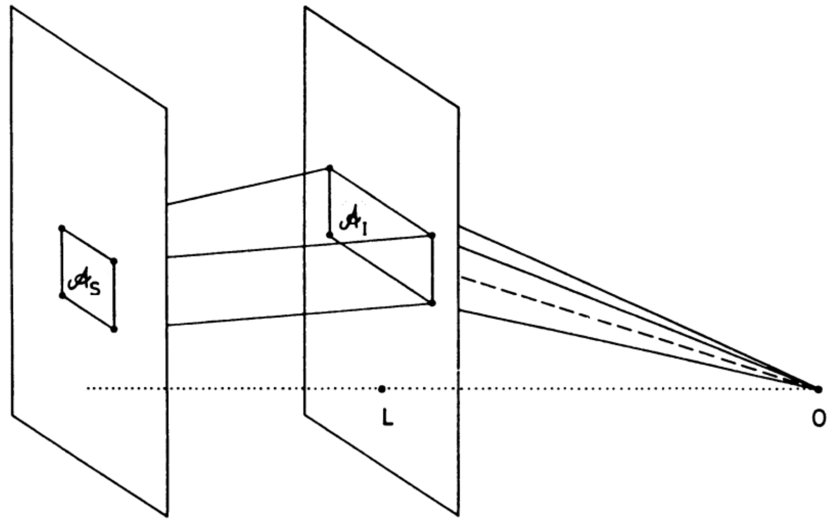

Figure 2.5: Sketch of the magnification effect due to the lens L: The source area $\mathcal{A}_S$ is mapped into $\mathcal{A}_I$. The ratio between the solid angle of the unperturbed image ($\Delta\omega_0 = \mathcal{A}_S/D_s^2$) and the observed one ($\Delta\omega = \mathcal{A}_I/D_d^2$) at the observer position O gives the magnification factor. This image is adapted from Figure 2.3 of Schneider et al. (1992).

The magnification factor due to the lens, μ, is then given by $\frac{\partial\theta^2}{\partial\beta^2}$, which is the inverse of the determinant of the Jacobian matrix A defined in equation 2.58:

$$\mu(\vec{\theta}) = \frac{\partial\theta^2}{\partial\beta^2} = \frac{1}{\det A(\vec{\theta})} = \frac{1}{(1-\kappa(\vec{\theta}))^2 - \gamma(\vec{\theta})^2}. \tag{2.62}$$

Note that there could be points $\vec{\theta}$ such that $|1-\kappa(\vec{\theta})| = |\gamma(\vec{\theta})|$, thus where the magnification is formally infinite. These points form the *critical curves* on the lens plane, which can be mapped to the source plane by the lens equation, in which case they form the *caustics* of the lens. However, an infinite magnification would correspond to an infinite increase in flux, which is clearly unphysical. This does not happen due to the fact that the source is not infinitely small; thus, the magnification is more precisely the weighted mean of equation 2.62 over the source. Secondly, we treated here the light propagation in the framework of the approximation to geometrical optics. When reaching the region of the critical curves, this approximation fails, and we should be employing the formally more precise wave-optic framework. This is, however, beyond the scope of this work, for which it will be enough to consider the critical curves as regions around which the lensed images are strongly magnified.

If we take the inverse matrix of the Jacobian $M = A^{-1}$, which is referred to as the *magnification tensor*, the magnification can be written as $\mu = \det M$. We then have two eigenvectors of this matrix:

$$\mu_t = \frac{1}{1-\kappa-\gamma} \tag{2.63}$$

$$\mu_r = \frac{1}{1-\kappa+\gamma}, \tag{2.64}$$

which are referred to as the *tangential* and *radial* magnification factors. This naming comes from the fact that for axis-symmetric lenses, these eigenvectors are oriented tangentially and radially with respect to the lens iso-density contours. Moreover, this corresponds to two critical curves, which are then referred to as tangential and radial.

Another important aspect of critical curves and caustics is related to the number of images. We have seen that the lens equation can have multiple solutions, and this depends on the position of the source. It can be demonstrated (e.g. Schneider et al., 1992) that the number of images changes by two when the source crosses a caustic. Finally, it is interesting to note that the magnification has a sign, which is referred to as the parity of the image. This, in turn, reflects the chirality of the lensed image, such that a positive parity image will preserve the chirality of the source, while a negative parity one will have an inverted chirality.

2.2.7 External Shear

In most practical scenarios, the lens system is not isolated, and there are significant matter distributions in the surroundings of the system which the lensing effect has to be accounted for, without explicitly modelling every component. This is usually faced by considering an external source of shear. This accounts for observable distortions in the data, as opposed to external convergence, whose effects would not be visible. Hence, the external convergence is set to 0 (I will discuss this further in Section 2.3.3). To model the shear, an ulterior potential ψ_γ can be considered, which has to satisfy the following conditions, such that $\gamma_1 = 1/2(\psi_{11} - \psi_{22}) = \text{const}$, $\gamma_2 = \psi_{12} = \psi_{21} = \text{const}$ and $\kappa = 1/2(\psi_{11} + \psi_{22}) = 0$. Given that $\psi_{11} \pm \psi_{22} = \text{const}$, both ψ_{11} and ψ_{22} must be constant. The effective potential can then be written in polar coordinates as $\psi_\gamma(\gamma, \phi) = \frac{\gamma}{2}\theta^2 \cos 2(\phi - \phi_\gamma)$.

Note that the external shear breaks the circular symmetry of the lens; thus, it presents degeneracies with the ellipticity components of the lens.

2.3 Gravitational Lensing: Application to Time Delay Cosmography

Given the general overview of gravitational lensing described in the previous Section 2.2, I will now focus on the lensing time delay effect and how this phenomenon can be employed in cosmology. Firstly, I will discuss the theory behind the time delay phenomenon (Section 2.3.1) and then how this is used in time delay cosmography (in Section 2.3.2). I will then focus on specifics of the Time Delay Cosmography, such as the mass-sheet degeneracy in Section 2.3.3, the variable sources used in this study in Section 2.3.4 and the effects of microlensing in such analysis 2.3.5.

2.3.1 Theory of Gravitational Time Delay

Let us consider the time it takes a photon to traverse the path from the source to the observer, t_0, which I will refer to as “flight time”. Due to the presence of the lens, the photons will be delayed by δt, where such delay is due to two separate effects, a gravitational delay δt_{grav} and a geometrical one, δt_{geom}. The first is also known as Shapiro time delay (from Shapiro, 1964), which is due to the relativistic effect of the lens. This can be computed by considering the difference in flight time of a photon traversing the *same* path, with and without a gravitational potential Φ which affects the space-time. Considering the refractive index obtained from equation 2.27 as the effective refractive index, we obtain:

$$\delta t_{\mathrm{grav}} = \int \frac{\mathrm{d}z}{c'} - \int \frac{\mathrm{d}z}{c} = \frac{1}{c}\int (n-1)\mathrm{d}z = -\frac{2}{c^2}\int \Phi \mathrm{d}z. \tag{2.65}$$

Considering the definition of the effective lensing potential from equation 2.42, we obtain:

$$\delta t_{\mathrm{grav}} = -\frac{D_{\mathrm{l}} D_{\mathrm{s}}}{D_{\mathrm{ls}}} \frac{1}{c}\psi. \tag{2.66}$$

The geometrical time delay can be instead intuitively considered as the excess of time the light takes due to the deflection compared to the “straight line”. Formally, this can either by measured by considering the metric or it can be approximated from a simpler geometrical construction, which is sketched in Figure 2.6. The dashed line represents the lightpath of a photon emitted by the source S, which reaches the lens plane in H’ and is deflected due to the lens by an angle $\hat{\alpha}$. It then arrives at the observer positioned at O from the direction θ. The unperturbed path is instead the straight solid line connecting S and O, $\overline{\mathrm{SO}}$. Let us then consider two circles centred on S and O, which are tangent to each other at the point H, which is located on the unperturbed path $\overline{\mathrm{SO}}$. Note, however, that H is not necessarily on the lens plane. These two circles meet the perturbed lightpath at K and K’. The difference in length of the lightpath is therefore $\Delta l = \overline{\mathrm{KHK'}}$. Following the notation of Figure 2.6, this corresponds to $\Delta l \approx \xi \hat{c}$.

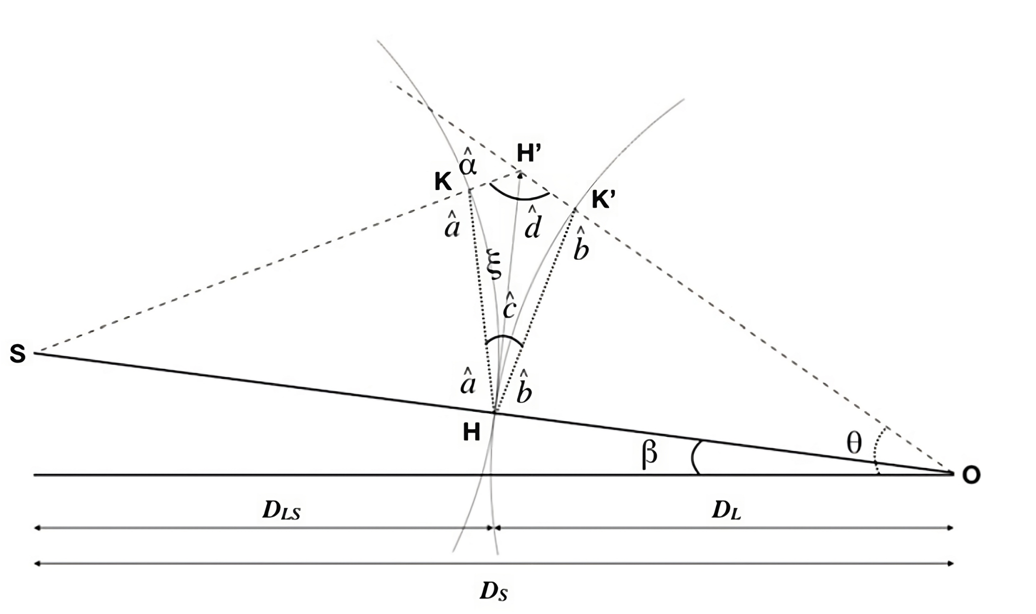

Figure 2.6: Scheme representing the geometrical time delay δt_{geom}. This image is taken from Figure 3.6 of Meneghetti (2021).

Moreover, given that the triangles SHK and OHK' are isosceles, we can obtain the following relations:

$$\hat{d} = \pi - \hat{\alpha} \tag{2.67}$$

$$\hat{a} + \hat{b} + \hat{c} = \pi \tag{2.68}$$

$$\hat{a} + \hat{b} = \hat{c} + \hat{d}. \tag{2.69}$$

Therefore angle $\hat{c}$ is $\hat{c} = \hat{\alpha}/2$. Therefore $\Delta l \approx \xi \frac{\hat{\alpha}}{2} = (\vec{\theta} - \vec{\beta}) \frac{D_{\text{l}} D_{\text{s}}}{D_{\text{ls}}} \frac{\vec{\alpha}}{2} = \frac{1}{2} (\vec{\theta} - \vec{\beta})^2 \frac{D_{\text{l}} D_{\text{s}}}{D_{\text{ls}}}$. Thus the geometrical time delay is $\delta t_{\text{geom}} = \frac{\Delta l}{c}$.
Both components of the time delay occur at the position of the lens; thus, they must be scaled by a factor $1 + z_l$ to account for the cosmic expansion. The total time delay is then

$$\delta t(\vec{\theta}) = \frac{1 + z_l}{c} \frac{D_{\text{l}} D_{\text{s}}}{D_{\text{ls}}} \left(\frac{1}{2} (\vec{\theta}^2 - \vec{\beta}^2) - \psi(\vec{\theta}) \right). \tag{2.70}$$

This is usually rewritten as

$$\delta t(\vec{\theta}) = \frac{D_{\Delta t}}{c} \phi(\vec{\theta}), \tag{2.71}$$

where

$$D_{\Delta t} = (1 + z_l) \frac{D_{\text{l}} D_{\text{s}}}{D_{\text{ls}}} \tag{2.72}$$

is referred to as the *time delay distance* and

$$\phi(\vec{\theta}) = \frac{(\vec{\theta} - \vec{\beta})^2}{2} - \psi(\vec{\theta}) \tag{2.73}$$

is called the *Fermat potential*.
In all practical cases, the time delay with respect to the unperturbed path is not observable. However, when the source is lensed into multiple images, the relative time delay between said images is observable:

$$\Delta t(\vec{\theta}_A, \vec{\theta}_B) = \delta t(\vec{\theta}_B) - \delta t(\vec{\theta}_A) = \frac{D_{\Delta t}}{c}\Delta\phi(\vec{\theta}_A, \vec{\theta}_B), \tag{2.74}$$

where

$$\Delta\phi(\vec{\theta}_A, \vec{\theta}_B) = \phi(\vec{\theta}_B) - \phi(\vec{\theta}_A) = \frac{1}{2}[(\vec{\theta}_B - \vec{\beta})^2 - (\vec{\theta}_A - \vec{\beta})^2] - [\psi(\vec{\theta}_B) - \psi(\vec{\theta}_A)], \tag{2.75}$$

and $\vec{\theta}_A$ and $\vec{\theta}_B$ are the positions of the two lensed images A and B.

2.3.2 Time Delay Cosmography

From Section 2.3.1 I have shown that the time delay between two images A and B, Δt_{AB} is a function of the difference of Fermat potential at the images position $\Delta\phi_{AB}$ and the time delay distance $D_{\Delta t}$, which is constant given a lens system. The latter is defined in equation 2.72 as a ratio of angular diameter distances; thus, following equation 2.16, it can be written as:

$$D_{\Delta t} = (1+z_l)\frac{D_{\mathrm{l}}D_{\mathrm{s}}}{D_{\mathrm{ls}}} \tag{2.76}$$

$$= (1+z_l)\frac{\frac{c}{(1+z_l)H_0}\int_0^{z_l}\frac{\mathrm{d}z'}{E^{-1}(z')}\frac{c}{(1+z_s)H_0}\int_0^{z_s}\frac{\mathrm{d}z'}{E^{-1}(z')}}{\frac{c}{(1+z_s)H_0}\int_{z_l}^{z_s}\frac{\mathrm{d}z'}{E^{-1}(z')}} \tag{2.77}$$

$$= \frac{c}{H_0}\frac{\int_0^{z_l}E^{-1}(z')\mathrm{d}z'\int_0^{z_s}E^{-1}(z')\mathrm{d}z'}{\int_{z_l}^{z_s}E^{-1}(z')\mathrm{d}z'}, \tag{2.78}$$

where $E(z) = \sqrt{\sum_i \Omega_{0,i}(1+z)^{1+w_i}}$ and assuming a flat universe. Interestingly, the fraction of integral on the RHS of this equation is only weakly dependent on cosmology, specifically on the different density components. $D_{\Delta t}$ is instead far more sensible to H_0, as it is inversely dependent on it. Thus, H_0 can be measured directly from the relation

$$H_0 = k\frac{\Delta\phi_{AB}}{\Delta t_{AB}}, \tag{2.79}$$

where k encodes all remaining cosmological parameters. This implies that we can measure directly H_0 from a strongly lensed system, independently of all other methods described in Section 2.1.2.
This methodology, referred to as "time delay cosmography" (TDC) was first introduced by Refsdal Refsdal (1964), but it remained beyond the observational capabilities until recently. In order to accurately constrain H_0, a TDC analysis requires four components:

- Constraints on the time delay between the images.

- Constraints on the mass distribution of the lens, from which the Fermat potential model can be computed.
- Information on the positions of the lens and source, i.e. their redshifts.
- Constraints on the effect of the environment of the lens on the measurement.

The time delay can be observed in the presence of a variable source such as SNae (as theorised by Refsdal and recently becoming a reality, e.g. in Kelly et al., 2016; Suyu et al., 2020) or QSOs (e.g. Wong et al., 2020). The first type of source has the advantage of being standard candles (for SNIa) or standardizable (for SNII De Jaeger et al., 2020), thus providing additional information as the absolute magnification can be used to constrain the lens model. SNae also have a known lightcurve (Hamuy et al., 1996), thus simplifying the time delay measurement. Finally, the SNae will eventually vanish after the explosion, thus leaving the lensed system unperturbed by its light, allowing the host galaxies of the SNae to be visible as lensed images, which can be used to greatly constrain the lens modelling Cañameras et al. (2021). The drawback of SNae is their rarity, such that only recently a handful have been observed (one of the most recent being SN "H0pe", Frye et al., 2024). Comparatively, strongly lensed QSOs are more common, ranging now on the order of hundreds Treu et al. (2018); Lemon et al. (2023). I will focus on the QSO sources given that this work will be based on a strongly lensed QSO, and will further discuss them in Section 2.3.4. The lensed images of the source must then be observed over time, recording their luminosity in so-called "lightcurves". These will then present the same variability structures shifted in time (due to the relative time delay between the images, see equation 2.75) and brightness (due to the different magnification of the images, see equation 2.62). The time delay can then be measured by correlating the lightcurves variabilities (as it has been successfully done in, e.g. Courbin et al., 2018; Millon et al., 2020b).

The constraints on the Fermat potential have increased in accuracy recently due to breakthroughs in the methodology of lens modelling using high-resolution images from space-based telescopes such as the Hubble Space Telescope (Suyu et al., 2010; Chen et al., 2019; Birrer et al., 2015; Shajib et al., 2020). Spectroscopy is then required to constrain the redshift of both lens and source, as the method is strongly affected by their positions. Finally, the effects of the environment have been recently stressed as one of the major sources of possible systematic, namely in the form of mass-sheet degeneracy Falco et al. (1985) and its generalised form of source position transformation Schneider, Sluse (2014). However, this pitfall has been recently studied in depth Suyu et al. (2010); Greene et al. (2013); McCully et al. (2017) and has been considered in recent analyses (e.g. Wells et al., 2023). I will further the topic in Section 2.3.3.

The combined improvement in the quality of the data and sophistication of the models yielded a precision on the order of 2% on H_0 in 2020 based on only a small sample of 7 lenses Wong et al. (2020). The focus now is shifted toward understanding and minimising systematics.

The analysis presented here offers a uniquely advantageous perspective on the problem, having been conducted entirely independently of any major current collaboration. While many analytical choices were influenced by previous works and established literature, the work was carried out autonomously. Consequently, this analysis serves as proof of reproducibility for the TDC methodology.

2.3.3 Mass-Sheet Degeneracy

A critical point of lens modelling is the so-called mass-sheet degeneracy (MSD), first introduced by Falco et al. (1985): there would be almost no observables to distinguish the mass distribution described by the convergence $\kappa(\vec{\theta})$ and (Schneider, Sluse (2013)):

$$\kappa_\lambda(\vec{\theta}) = \lambda\kappa(\vec{\theta}) + (1 - \lambda), \tag{2.80}$$

along with a rescaling of the source position $\vec{\beta} \to \lambda\vec{\beta}$, as the lens model would reproduce exactly the same dimensionless observables: image shape, position, magnification ratio, etc. However, this degeneracy would linearly affect the Fermat potential differences: $\Delta\phi \to \lambda\Delta\phi$, thus biasing the final constraint on H_0 by a factor λ.

In practice, choosing a model for the mass profile formally breaks the degeneracy by arbitrarily selecting one of the available $\kappa(\vec{\theta})_\lambda$.

This degeneracy can be divided into two separate sources of uncertainty: the inner shape of the mass profile λ_{int} and the external mass sheet degeneracy κ_{ext}. In the first case, it is assumed to be due to the freedom of the internal profile: for a given lens system, there exist multiple different profiles that fit equally well the data but return different constraints on $\Delta\phi$. This degeneracy can be broken through complementary data to the internal structure of the lens, such as the spatially resolved stellar kinematics of the lens Shajib et al. (2023). Similarly, external convergence is correlated to the environment of the lens, such as over-/under-densities along the line of sight. In this case, deep photometric observations with a large field of view around the system are required to identify possible targets Suyu et al. (2010, 2013). These must then be spectroscopically observed in order to determine their redshift. Once this catalogue is obtained, it can be compared in a statistical sense to cosmological simulations, such as the Millennium simulation Springel et al. (2005), where the κ_{ext} can be measured numerically. Other approaches instead explicitly model the most luminous components near the line of sight, such as Millon et al. (2020b).

Due to time constraints and lack of complementary data (most notably, spectroscopic data) in this analysis, I will not constrain the MSD, which is then left as a future improvement.

2.3.4 Variable Sources: QSO

In this work, I will be considering a quadruply lensed QSO, SDSSJ1433, presented in detail in Section 3.2. I will discuss here some of the properties of QSOs most relevant for this analysis.

Starting from the name, QSO stands for Quasi-Stellar Object, which is a remnant of its history Burbidge (1967), as they appeared as point-sources, but were significantly brighter, with very large redshift and with significantly different spectra with respect to stars. The current understanding (see e.g. Beckmann, Shrader, 2012) explains this finding by describing the QSO as a very bright active galactic nucleus (or AGN). These are systems powered by super-massive ($M \geq 10^6 M_\odot$) black holes residing at the centre of galaxies, which accrete matter from their surroundings. As the infalling material forms an accretion disk around the black hole, its gravitational energy is converted into electromagnetic radiation through friction Frank et al. (2002). This results in the emission of significant amounts of energy, on the order of $L_{\mathrm{bol}} \approx 10^{48}\mathrm{erg/s}$ across the electromagnetic spectrum Padovani et al. (2017).

A schematic reproduction of an AGN is shown in Figure 2.7.

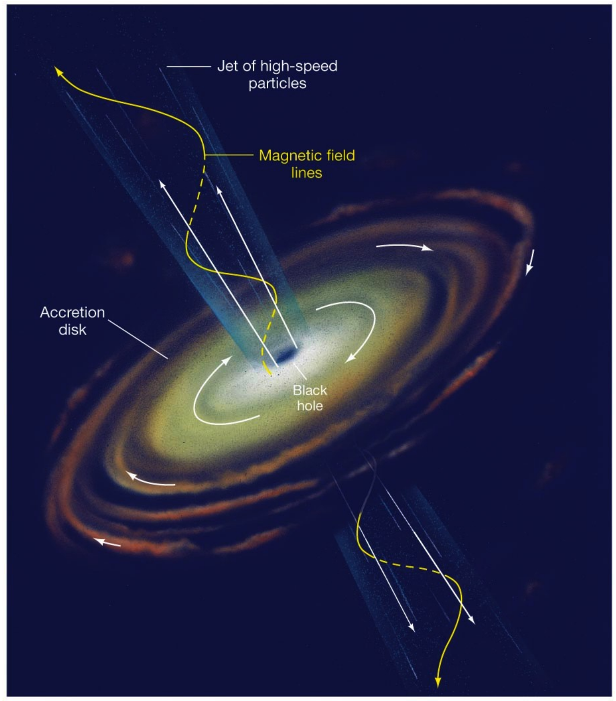

Figure 2.7: Artist rendition of the structure of an active galactic nucleus, Figure taken from Addison-Weasley (2019).

The common understanding of these AGN considers two ultra-relativistic polar jets. The physics of such jets is still discussed (Blandford et al., 2019), but is expected to be related to the magnetic fields produced by the infalling matter.
For what concerns TDC, however, the most important aspect to consider is that QSOs are (from Peterson, 1997):

1. high-redshift objects, spanning a large redshift range: $0.5 < z < 7$
2. bright in most wavelengths
3. time-variable
4. point-like

The first point makes them likely to be strongly lensed. The second point means that they can be observed even at this high redshift, especially if magnified by lensing. The time-variable luminosity makes them optimal targets for TDC analysis. The last point is however, a minor drawback, as point-like sources are more affected by microlensing, which is discussed in the following Section 2.3.5.
Due to the correlation between the infalling material and the corresponding luminosity of the

QSO, the variability would appear to be effectively stochastic and can vary significantly even on short timescales: on the order of a fraction of a magnitude, in the timescale of days Peterson (1997). An example of lightcurve of a QSO is shown in Figure 2.8. However, there appears to be observational evidence Kawaguchi et al. (2000) of a non-flat power spectrum of the lightcurves. This seems to indicate that small variations are more likely to occur over short timescales and large variations on longer timescales.

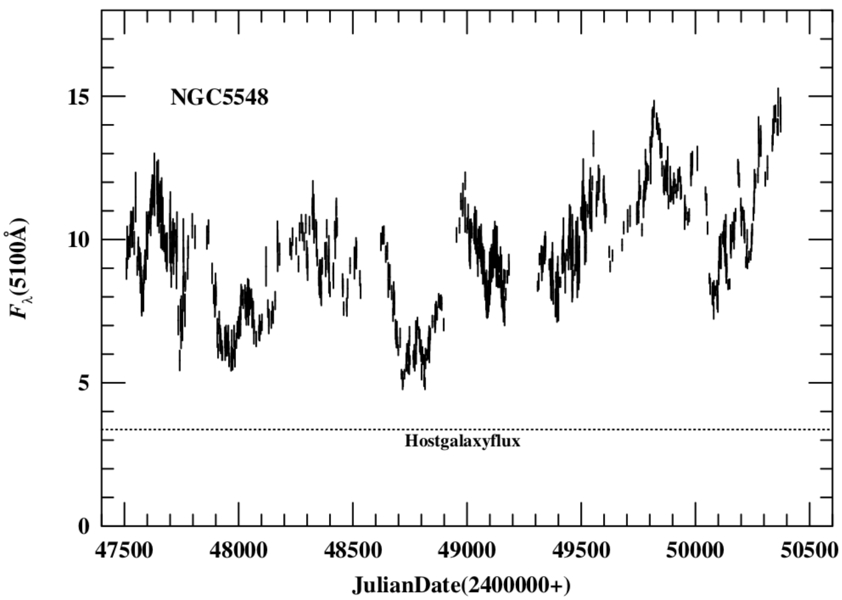

Figure 2.8: Flux variability in the optical for of NGC 5548 between 1988 to 1996 Peterson (2001). The horizontal line indicates the constant contribution of the host galaxy.

In practice, the lightcurves appear to be well-fitted by splines, the number of knots depending on the scale of variability of the single system. This has been empirically demonstrated by the success of the COSMOGRAIL collaboration Millon et al. (2020a) in fitting lightcurves of lensed QSOs. I will then follow a similar approach in my time delay analysis in Chapter 5.

Another aspect to consider is that brighter QSOs present lower variability amplitude Wills et al. (1993). This poses an interesting optimisation problem when selecting an optimal target for TDC: brighter QSO would have a higher photometric signal and therefore less noise, but the reduced amplitude in variability would hinder the time delay measurement, for which the variability is the signal, and vice-versa. For this study, this limitation had to be ignored due to the limited number of objects observable at the time (see Chapter 3), but it will appear that luckily the chosen system presented high enough variability for the time delay to be observable, see Chapter 5.

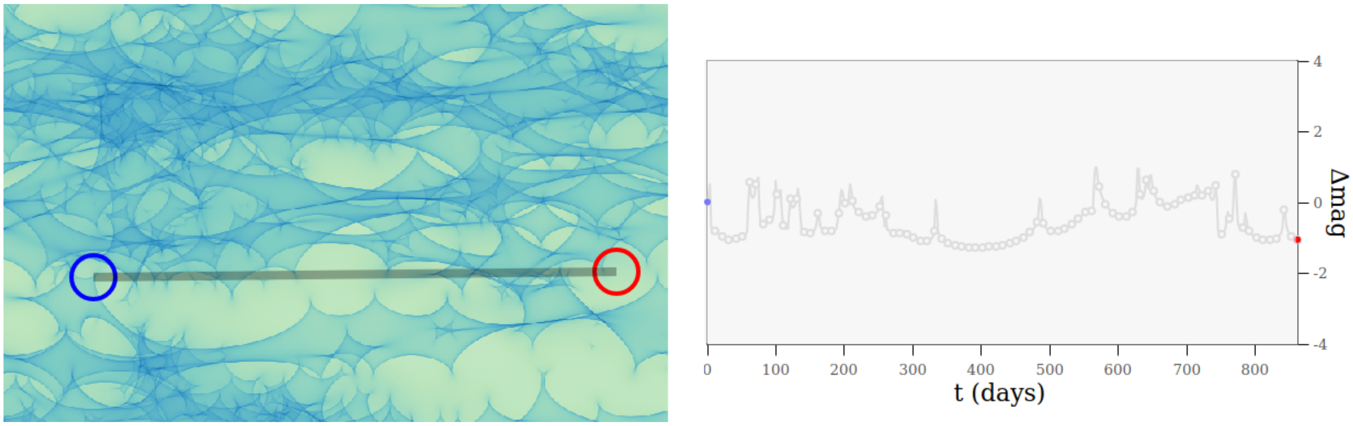

Figure 2.9: Simulation of a microlens map with a background source obtained from GIMLET Astrophysics for, Technology of (2020).

2.3.5 Microlensing

Until now, I have discussed galaxy-scale lenses, as this will be the main focus of the lens modelling. However, any mass, if concentrated enough, can be a GL. We then refer to them as microlensing if their mass is in the range of $10^{-6} \leq \mathrm{m}/M_{\odot} \leq 10^{6}$ (Schneider et al., 2006). Let us then focus on such microlenses on the lens plane of the macro-lens. Being that the microlens can be approximated to a point mass, the corresponding $\theta_{E,\,\mathrm{micro}}$ is obtained as

$$\theta_{E,\,\mathrm{micro}} \approx 10^{-6}\sqrt{m/M_{\odot}}\mathrm{arcsec} \tag{2.81}$$

for microlenses in a lensing galaxy at $z \approx 0.5$, acting on a background source $z \approx 2$. Due to the scale of the phenomenon, the multiple images due to the microlens are not observable; however, the magnification of the object can be very significant, on the order of magnitudes (e.g. Schmidt, Wambsganss, 2010). Moreover, the configuration of the microlenses is by nature highly unstable, i.e. the alignment of the source, the microlens and the observer is not constant over time. The timescale of such a phenomenon is, however, very variable as well and can range from days to months and even years Schmidt, Wambsganss (2010). Finally, the density of microlenses at a given position on the lens plane is not easily constrained a priori. Nonetheless, in the case of strongly lensed quasars, the convergence is assumed to be on the order of $\kappa_{\mathrm{mircolens}} \sim 1$, meaning that on average several microlenses are affecting the quasar image at any given time Schmidt, Wambsganss (2010).

The final effect of a pattern of microlenses on a moving source can be seen in Figure 2.9. This is a simulation obtained from GIMLET Astrophysics for, Technology of (2020), where a constant source moves on the background of a fixed map of microlenses. It is important to notice that the strongest variation appears at the crossing of the caustics, not within the caustics. Moreover, the scale of this phenomenon is dependent on the size of the source. Larger sources will take longer to cross the lightcurves, thus "diluting" the event over time instead of having a sharp peak. Let us then consider what this phenomenon implies for our study. Given that the lens is not observable directly, the effect of the microlenses on the observed lightcurves would correspond

to a brightness variation of an image independently of the "intrinsic" variability of the source. This effectively introduces an ulterior source of error in the time delay analysis, which is based on the principle of "matching" the lightcurves by shifting them; this is hindered if the lightcurves do not match. In extreme cases, if the microlensing effect distortion is too severe, the time delay is not recoverable. However, most of the time, the effect is not so dramatic, and it only contributes to increasing the noise. As it will be shown in Chapter 5, this will be the case for this analysis.
Interestingly, while the size of the QSO is well approximated to a point source when compared to the scale of the macro-lens, this is not valid when compared to the size of a microlens, which is expected to be significantly smaller than the QSO. Considering then the QSO accretion disk, due to its structure briefly described in Section 2.3.4, it will present higher temperatures in the centre and lower temperatures outwards. Correspondingly, the spectra emitted will be different. Given the localised magnification of the microlens, the magnified light would be colour-dependent, and it would therefore be possible to identify microlensing events by observing the lightcurves in multiple filters. There is, however, a further step to consider, which is the "lamp post" model introduced by Tie, Kochanek (2017). Here, it is considered that the variability of the QSO also originates from the central black hole, and it then propagates outward in the disk within the light-crossing time of the disk, which is on the order of days. Considering this lag in combination with the microlensing effect, it is possible to observe a microlensing-induced time delay.
In this analysis, however, I do not consider this source of systematic error due to several reasons, such as the lack of observation in multiple filters (as I will discuss in Chapter 5), the good results obtained when considering minimal microlensing and, most importantly, the limitation in time. A future improvement of the analysis might take into account this possible source of error.

Chapter 3

Time Delay Cosmographic Analysis at Wendelstein: Project Outline

In this project, I aimed to use one or possibly several multiply-lensed QSOs for time delay cosmography (TDC). The work aimed to accomplish three main goals:

- Firstly, add a novel system to the array of TDC systems analysed till now, which is limited to 7 lenses as of 2020, see Birrer et al. (2020);

- Offer an independent validation of the TDC method implemented by larger collaborations such as H0LiCOW and TDCOSMO. While the general outline of the project is based on the same approach, several paths were redefined independently. In particular, this study faces the lens modelling with a new approach by analysing each exposure independently from each other, which proves to be of paramount importance in this study (see Section 4.4.1);

- Finally, we aim to introduce the Wendelstein 2.1-meter telescope as a state-of-the-art follow-up facility for observational cosmology, particularly TDC. The success in measuring accurate and precise time delays between the images further consolidates the argument that 2-meter class ground-based telescopes are well suited for multi-year, high cadence, high SNR observational campaigns for time-domain surveys. These facilities will be necessary assets in the "Big Data Era" of astrophysics, where stage-IV surveys are expected to deliver a wealth of objects of interest. Only for strong lensing, LSST and Euclid are expected to detect $O(10^5)$ objects of interest each Collett (2015). Facilities such as Wendelstein are well suited to carry the follow-up observations required to take advantage of such a dataset fully.

3.1 Method

The study is divided into two, mostly independent studies.

Following the equation 2.79, in order to constrain H_0 I had first to constrain $\Delta\phi$ and Δt, given that k is taken as a fixed constant assuming a flat ΛCDM cosmology with $\Omega_{m,0} = 0.3$. To measure $\Delta\phi$, I needed to measure the Fermat potential ϕ at the image position, which in turn required me to model the mass distribution of the lens. I addressed this by employing the public Python library `lenstronomy` on archival Hubble Space Telescope (*HST*) images across multiple filters. The redshifts of the source and lens were known from prior studies (Agnello et al., 2018).

This analysis is explained in detail in chapter 4.
The second part of the analysis revolves around the time delay estimation, described in chapter 5. Given the variability of the source, here a QSO, and the time delay between the images, the luminosity of each image over time, referred to as a "lightcurve", should share a common shape, although shifted in time and brightness (due to the different magnification of the images). Thus, the time delay can be measured by observing such variations over time. For that, I made use of the dedicated observational campaign carried out from the 2.1 meter Fraunhofer telescope at the Wendelstein Observatory using the Wendelstein Wide Field Imager (WWFI, Hopp et al., 2014). I reduced the daily observations and obtained the lightcurves, which I then analysed using the publicly available Python library `PyCS3` (Tewes et al., 2013). I describe the details of this analysis in Chapter 5.

Finally, the two posteriors on $\Delta\phi$ and Δt are used to constrain H_0 with a Bayesian approach in chapter 6.

3.2 Object Selection and Description

The first step of the project was to select the objects of interest, i.e. known multiply-lensed QSOs. The sample of potential targets was based on the public Gravitationally Lensed Quasar (GLQ) Database Lemon et al. (2019), which presented a pool of 220 lensed QSOs known in 2019. This had to abide by several criteria to allow for the observations from *WST* and optimise the chances of high-precision measurements. Here they are listed loosely in order of importance:

- The QSO system has to be visible from the Wendelstein Observatory for at least half of the year (possibly the whole year). Given the location of the *WST* , this constrains that the pool of eligible objects has to have a high declination, in the range of DEC$\geq$ 0. Moreover, the requirement on photometric precision restricted the observations to a relative airmass of 2, further limiting the elevation to be higher than 30 degrees.

- The QSO system had to have high luminosity in the visible ($g' \lesssim 23$ mag) to be observable with high photometric accuracy. This meant that most, if not all images had to be bright in the visible.

- The QSO images have to be lensed with a large enough separation to be deblended in the observations. The median seeing from *WST* is $\sim 1''$, thus dictating the minimum separation between the observed images.

- The QSO had to be multiply imaged, and the presence of four images highly improved the constraints on both the time delay measurement and the lens model reconstruction. The search was therefore limited to the quadruply lensed QSOs, referred to as "quads".

- The mass model required exquisite spatial resolution to be effective; thus, the search was focused on objects that had archival *HST* observations or with similar resolution. Multiple filters, especially infrared ones, were strongly favoured as they allowed increased modelling precision due to higher information content (e.g. due to the presence of lensed extended sources).
- The simple systems were preferred, i.e. galaxy-scale lenses, isolated from other larger structures (e.g. galaxy clusters). This simplified the lens modelling and therefore increased the measurement precision and accuracy.
- The availability of spectroscopic constrain on the redshift of the QSO and its lens was a requirement, as it is a necessary component of a cosmological analysis such as TDC.
- The presence of additional extended lensed features, such as arclets or arcs produced by the host galaxy of the QSO or other lensed sources, was favoured as it significantly improved the lens modelling constraints, but was not considered a strict requirement.

Previous analyses present in the literature of the obtained sample were taken into account, especially concerning the estimates or forecasts of the time delay between the images. The observational campaign was to be limited to a few years, possibly less, and the sampling was expected to be on the order of a day. Any time delay larger than 1/3 year or smaller than a few days could hardly be detected.
This search resulted in two optimum candidates, SDSSJ1433+6007 and PSJ1721+8842. While promising, PSJ1721+8842 proved to be highly complex in its lens modelling and presented little variability in luminosity during the observation period, while simultaneously being plagued by intense microlensing. Therefore, the analysis of this system is yet to be finalised and is not presented here. I instead present here the study carried out on SDSSJ1433+6007, referred to hereafter as J1433.

J1433

Firstly discovered by (Agnello et al., 2018) (henceforth A18), this system is composed of a QSO quadruply lensed by a foreground bright, early-type galaxy, henceforth referred to as the Main Lens. The system also presents a lensing "perturber", which is believed to be a satellite galaxy of the Main Lens (see Section B for an estimation of its photometric redshift). Both the lens's and QSO's spectra were observed, the first from the Keck II telescope using the Echellette Spectrograph and Imager (ESI) and the second from the 2.5 m Nordic Optical Telescope (La Palma) using the Andalucia Faint Object Spectrograph and Camera (ALFOSC). This resulted in $z_{\mathrm{QSO}} = 2.737 \pm 0.003$ and $z_{\mathrm{lens}} = 0.407 \pm 0.002$, reported by A18.

HST observations in 5 filters from the Wide Field Camera 3 (WFC3) are available (F475W, F814W, F105W, F140W and F160W), in which the host galaxy of the QSO is visible in the two

QSO Lensed Image Pairs	Separation ["]
A-B	3.757
A-C	2.262
A-D	2.984
C-D	2.807
B-C	1.795
B-D	2.588

Table 3.1: Separation in arcseconds between the various pairs of lensed images. It can be seen that they are all well separated ($\Delta\vec{x} > 1.6''$). The separation between images and lens components (the main lens and the perturber) can be smaller, but this can be accounted for (see Chapter 5). These values are obtained by measuring the images position with **SExtractor** (Bertin, Arnouts, 1996) in filter F814W.

most infrared ones (F140W and F160W). This is expected due to its higher brightness in such wavelength caused by its redshift: due to the Balmer break at $\approx$ 3645Å and the redshift of the QSO, it is expected to be brightest at $\lambda \gtrapprox 1360$ nm.
J1433 is optimally located high enough in the sky to be observable year-round from the *WST*. The coordinates of image A, the lensed QSO image taken as the reference point, are RA=14:33:22.7786 and DEC=+60:07:16.928.
It also presents a large enough separation between lensed images (min($\Delta\vec{x}$)~1.8", see Table 3.1), which permits the observation of the images as unblended from *WST*. Furthermore, these images are bright enough to be observable with high photometric accuracy in *g'* filter. This was deduced a priori from the available literature at the time, specifically from (Shajib et al., 2019), Table 5, where the luminosity of the QSO images is reported (note that the image labelling is different with respect to this analysis). Considering F475X as being the most similar filter to our *g'*, we can see that the dimmest image, D, has a magnitude of 21.93±0.04. While this is close to the minimum threshold imposed for the luminosity, all other images are significantly brighter and thus should provide a high enough photometric accuracy.
As mentioned, J1433 was previously studied in the literature, although never with the accuracy required for cosmographic constraints. Nevertheless, this provided spectroscopic constraints for the main lens and the QSO, as well as initial lens models useful to estimate the range of expected time delays. From A18, the expected time-delays range from ~ 15 days to ~ 110 days, which is the optimal range of observable delays for this type of analysis. Finally, the fact that this system had not yet been studied within a time-delay cosmographic framework made it appealing to produce new science.

Chapter 4

TDC@W: Lens Modelling of SDSSJ1433

I here describe how the data was collected, preprocessed and analysed to constrain the mass profile of J1433. Note that the focus of this analysis will be to constrain the Fermat potential posterior at the positions of the images.
The main idea of this study, in contrast with most of the standard work done in similar analyses, is to model the lens independently for the various filters. This would allow for a comparison of the independent posterior distributions and therefore stress the potential tension between them. Indeed, a choice of model might fit well every filter, but if doing so the posterior appears to be in tension, it would indicate a failure to converge to a single result. Given the achromaticity of lensing effects, excluding time-varying systems such as microlenses (see Section 2.3.5), each filter should converge to the same mass model. Note that, since the spectra of the different components (QSO, satellite, early- and late-type galaxies) are not constant over the observed wavelengths, the amount of information will vary. Moreover, the filter differed for exposure time and pixel resolution, as discussed in Section 4.1. However, this should only mean a different degree of constraint of certain parameters - e.g., the position of QSO images is badly recovered in lower-resolution exposures. On the other hand, explicit tension between the observed posterior would then indicate that the model is converging to different results for different filters. This would then hint that the modelling approach chosen is not able to fit the data, either due to over- or underfitting. Analysing the filters jointly, usually referred to as multifilter modelling, could instead lead to convergence to some intermediate solution. This could be, for example, a local minimum that would fit reasonably well each filter, while not raising any suspicion to a superficial investigation. Finally, for the independent modelling of each filter, once the convergence is reached, the posterior can be multiplied together. I then show in Section 4.5 that this procedure is mathematically equivalent to sampling the posterior given a joint likelihood informed by all filters.
When required, the default cosmology is assumed to be a flat ΛCDM with $\Omega_{m,0} = 0.3$, and $H_0 = 70\frac{\text{km}}{\text{Mpc s}}$, taking care not to be biased by such an assumption.

I introduce the details of the data used in the modelling in Section 4.1, and I detail their preprocessing before the modelling in Section 4.2. I explain the modelling choices in Section 4.3, and

Filter	Central Wavelength [Å]	Exposure Date	Exposure Time [s]	Resolution (original/drizzled) [$\frac{''}{\mathrm{pix}}$]
F475X	4772.2	2018-05-04	1504.00	0.040/0.040
F814W	8048.1	2018-05-04	1428.00	0.040/0.040
F105W	10551.0	2019-02-14	124.23	0.128/0.128
F140W	13922.9	2019-02-14	446.93	0.128/0.080
F160W	15369.2	2018-05-04	2196.93	0.128/0.080

Table 4.1: Specifics for the *HST* exposures. The resolution is indicated before and after drizzling.

the details of the modelling run in Section 4.4. I then combine the results in Section 4.5 and discuss the result in Section 4.6. Here I also present the additional outcomes yielded by this analysis, such as the total mass and mass-to-light ratio (Section 4.6.2) and the colour profile of the main lens (Section 4.6.3). Lastly, I compare the obtained results with the available literature (Section 4.6.4).

4.1 HST Data

The system was observed from the Wide Field Camera 3 (WFC3) of *HST* on two separate occasions, first using the optical filters F475X and F814W and the near-infrared (NIR) filter F160W in May 2018, then using the two NIR filters F105W and F140W in February 2019. The details of the observations are reported in Table 4.2, while a colour image is reported in Figure 4.1. In this image, I also define the naming convention which is followed in this work. The naming of the image follows the naming convention of A18; this is defined by the time-delay predicted by the lens configuration: image A being the one with minimal time delay, followed by B and so on. In Figure 4.1, I also indicate the main lens galaxy as G and the nearby perturber as Pert.

While a possible approach to the model would be to model every filter fully independent of each other, preliminary tests indicated that the constraints carried by the optical filters were insufficient to constrain the full lens model. This becomes clear if we consider that these filters, while featuring the highest pixel resolution, only present the QSO image positions as the constraint for the lens model, as explained in Section 3.2. As a rough approximation, it can be seen that the images correspond to only 4x2=8 constraints, while the free parameters are on the order of 30 parameters (see Table 4.5). Thus, the model would be underconstrained.
Moreover, a similar discourse is valid for F105W, which I discarded from the lens modelling due to the low exposure time and lack of a lensed host galaxy. The details of this choice will be discussed in Section 4.2.1.
Therefore, I developed a different approach and used the optical filters only to measure and tightly

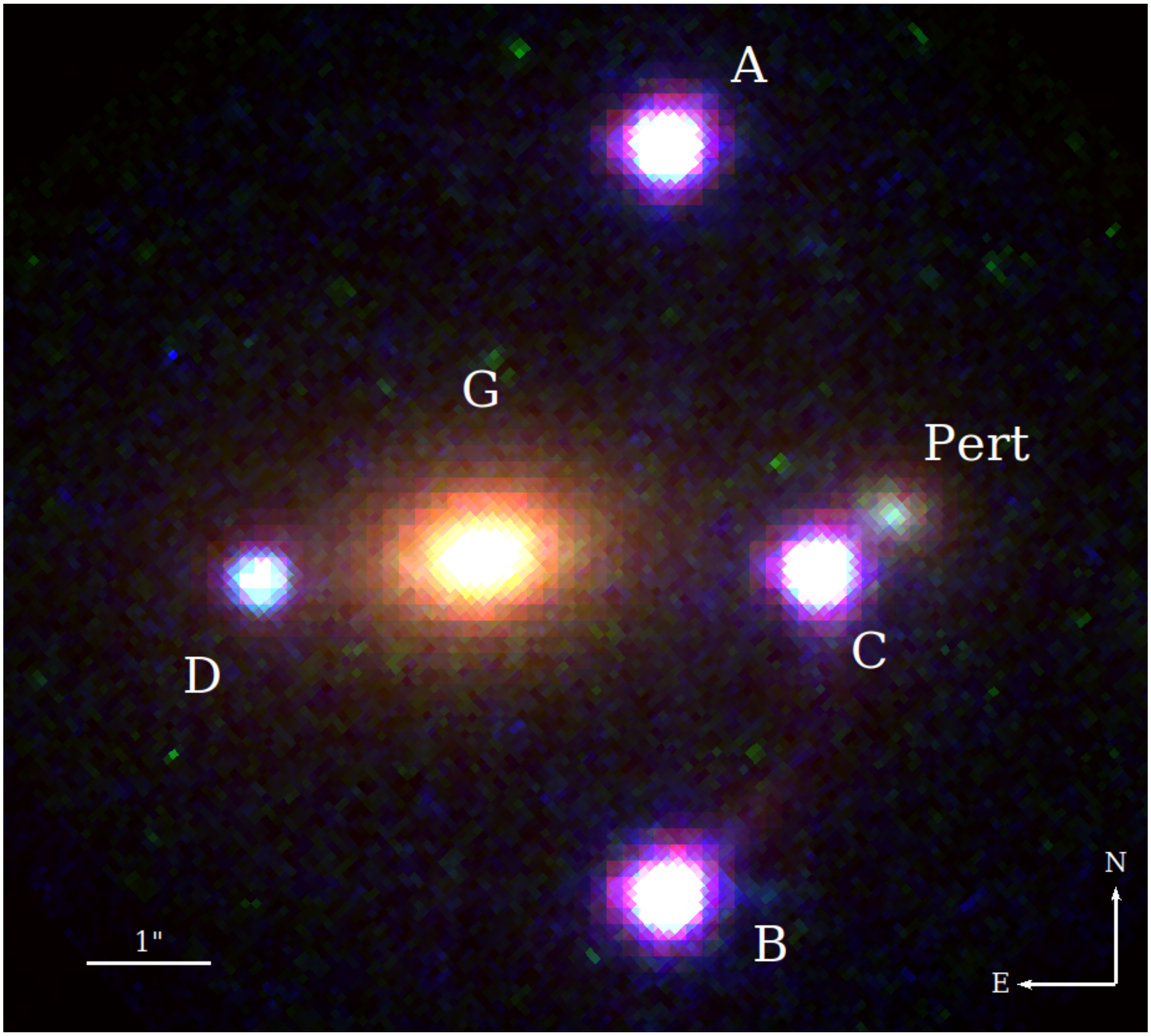

Figure 4.1: Colour image from *HST* of J1433 using three of the available filters: F475X (blue), F814W (green) and F160W (red). Following the convention of Agnello et al. (2018), the blue point sources are the four lensed images of the QSO and are indicated with capital letters from A to D. G Indicates the central red galaxy, which acts as the main lens, and "Pert" indicates a smaller galaxy, northwest of image C, acting as a perturber for the main lens (likely its satellite galaxy, as later discussed in Appendix B). The North and East directions and the 1-arcsecond scale are also shown.

constrain the QSO position for the modelling prior. I then consider only the model of the NIR filters to constrain the posterior (see Section 4.5. These filters had lower resolution, however, they featured extended source features, which constrained far better the lens parameters, as I will show in Section 4.4.

4.2 Preprocessing of *HST* Data for Lens Modelling

The *HST* data obtained from the online archive MAST (09/2023) has to be reduced and preprocessed to be input in the lens modelling pipeline. This preprocessing is composed of the "drizzling" of the images (see Section 4.2.1), the sky correction and error computation (see Section 4.2.2), the Point Spread Function (PSF) modelling (see Section 4.2.3), the lens light modelling of the main lens (see Section 4.2.4) and the masking (see Section 4.2.5).

4.2.1 Drizzling of the Filters

The *HST* online archive provides all required images ranging from the raw single exposures to the combined and reduced science frames, which have been "drizzled".

"Drizzling", first introduced by Fruchter, Hook (1997), refers to the method used to improve the resolution of multiple dithered images. This is done by mapping each low-resolution pixel into a higher-resolution, rotated and shifted grid. To limit the convolution effect of the original pixel grid into the new image, the pixels are rescaled by a factor p before the computation. The resized pixels can be seen as "drops" that "drizzle" (hence the name) into the output subsampled image.

The *HST* images are processed by the `DrizzlePac` pipeline, specifically `astrodrizzle` (Fruchter, et al., 2010). This pipeline masks cosmic rays, resamples and combines multiple dithered exposures by drizzling. However, this method does not provide a corresponding error frame, which must be computed separately as described in Section 4.2.2. Moreover, the drizzled images obtained from the archive are drizzled to the same resolution as the original observations. This maximises the SNR while not increasing the image resolution.

Given the requirements of the analysis, I decided to consider the standard drizzled images for the optical, as their original resolution is sufficiently high, reaching $0.04''$/pixel. This approach maximises the SNR. Instead, I opted to drizzle the infrared images in order to increase their resolution, which is significantly poorer than the optical, originally being $0.128''$/pixel. However, F105W had only one exposure (see Section 4.1) and therefore could not be drizzled to higher resolution, as well as having a very poor SNR. Moreover, the extended source is still not visible in this filter, either due to low SNR or the wavelength observed, and therefore suffers from the same lack of constraining power as the optical filters. This led me to discard this filter from the lens modelling.

The two near-infrared exposures available, F140W and F160W, were drizzled following the standard implementation, using a pixel fraction $p = 0.8$ (indicated by `final_pixfrac` in the code) and a final pixel scale of $0.08\frac{''}{\mathrm{pix}}$. For consistency, I employed `astrodrizzle` to resample

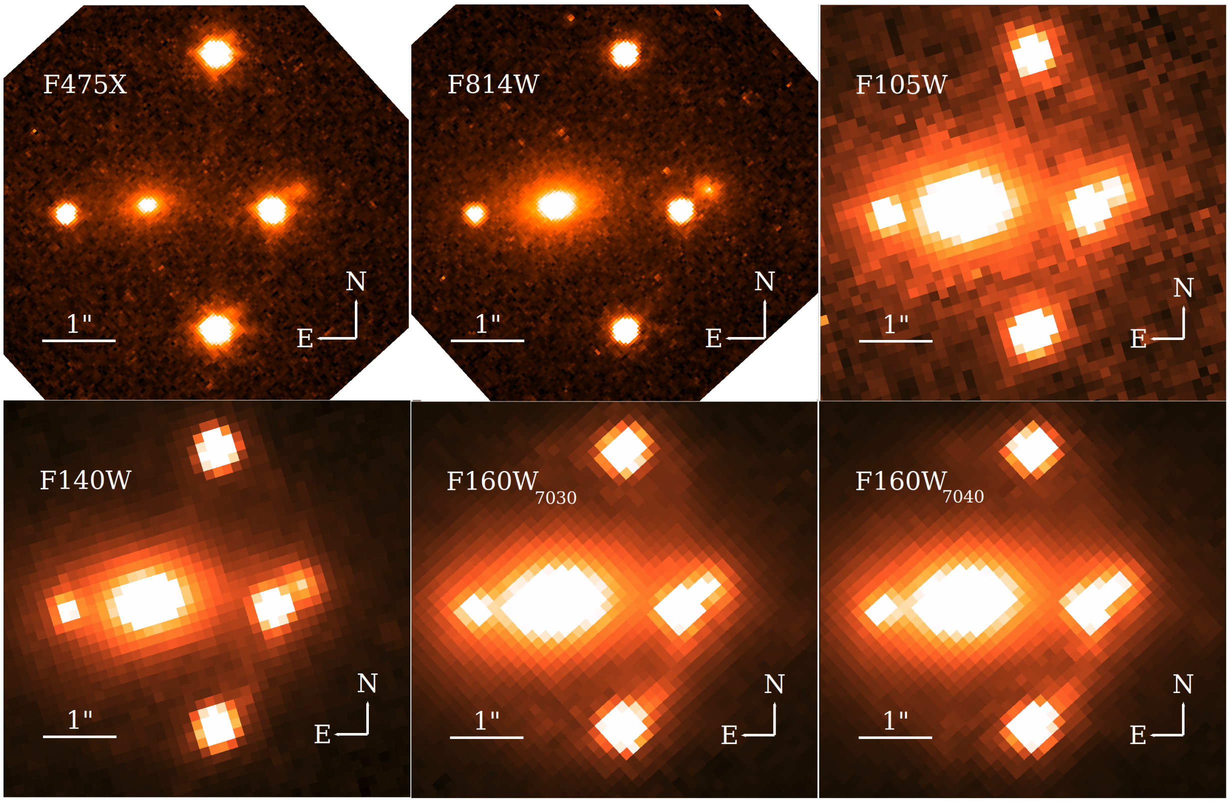

Figure 4.2: Science frames as recovered from the *HST* archive. Notice that the drizzling is performed without increasing the resolution. For more details refer to Figure 4.1.

the exposure F105W, without increasing its resolution (considering thus `final_pixfrac`=1 and `final_scale`=$0.128\frac{''}{\text{pix}}$). As previously mentioned, the optical images were maintained on their original frame, which is equivalent to running `astrodrizzle` with `final_pixfrac`=1 and `final_scale`=$0.04\frac{''}{\text{pix}}$.
The original frames are shown in Figure 4.2, while the drizzled results are shown in Figure 4.3 (which are only present for F105W, F140W and F160W, as the optical filters are not drizzled further). Moreover, the multiple exposures of F160W were drizzled automatically into two separated combined frames in the *HST* archive, referred to as $\mathrm{F160W}_{7030}$ and $\mathrm{F160W}_{7040}$, as shown in Figure 4.2. In this analysis, these exposures are instead combined into a single frame by the drizzling of the original exposures. I also report in Table 4.2 an overview of the various filters, with their respective wavelength, exposure date, length of exposure and resolution.

Note that while only two of the filters are actively used in the lens modelling, I will present in this Section the preprocessing of the observed data (namely the error frame estimation and PSF modelling) and corresponding results for all filters. This has multiple reasons; firstly, I tested various approaches to the modelling during the analysis, and most of them required the full set of preprocessed data for each filter. Secondly, the successful repetition of the process for different datasets solidifies the claim that such a process is well-defined. Lastly, most of the data will be eventually used for other purposes outside the lens modelling proper, such as the isophote lens

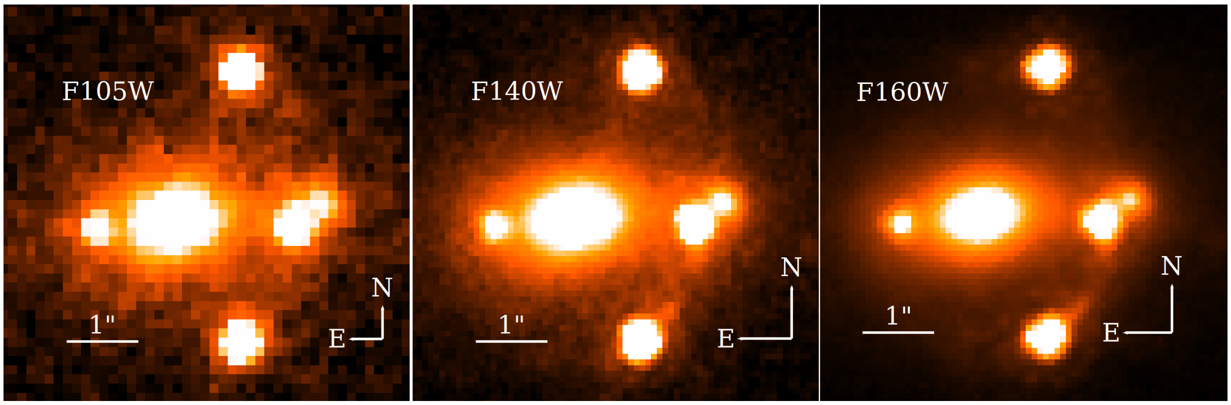

Figure 4.3: Science frames after drizzling, which was performed only on the NIR filters F105W, F140W and F160W. Note how for F105W the resolution can not be improved, while for the two other filters, it reduces to 0.08 $\frac{''}{\text{pix}}$

Filter	Central Wavelength [Å]	Exposure Date	Exposure Time [s]	Resolution (original/drizzled) [$\frac{''}{\text{pix}}$]
F475X	4772.2	2018-05-04	1504.00	0.040/0.040
F814W	8048.1	2018-05-04	1428.00	0.040/0.040
F105W	10551.0	2019-02-14	124.23	0.128/0.128
F140W	13922.9	2019-02-14	446.93	0.128/0.080
F160W	15369.2	2018-05-04	2196.93	0.128/0.080

Table 4.2: Specifics for the *HST* exposures. Note the comparison between the original sampling and the result after the drizzling procedure, which was applied to the NIR frames to increase their resolution. This was not possible for F105W due to the lack of multiple exposures.

light modelling (see Section 4.2.4), the colour profile of the main lens (see Section 4.6.3) and the photometric redshift estimation of the perturber (see Section B).

4.2.2 Error Frames and "Sky" Correction

As previously mentioned, the drizzling pipeline does not provide a "drizzled error image" by default. Quoting from Hoffmann, Mack (2021), " At present, there's no easy way for the user to do this. Perhaps this could be done by swapping in the `flc.fits/flt.fits` ERR array for the science [...] array, then running `astrodrizzle` on the modified images.", where ERR indicate the original error frame for the single exposure. This approach, however, does not seem valid. Firstly, as drizzling combines various pixels together, their uncertainty is now correlated, as explained in Section 7 of Fruchter, Hook (2002) (henceforth F02, from which all following quotes in this Section are taken). Even when ignoring this effect,the suggested approach appears incorrect. The drizzling pipeline, as described in F02, should follow Equation 5:

$$I_{x_0y_0} = \frac{d_{x_iy_i}a_{x_0y_0x_iy_i}w_{x_iy_i}s^2}{W_{x_0y_0}}, \tag{4.1}$$

where i indicate an input image with pixels (x_i, y_i) with corresponding data $d_{x_iy_i}$ and weight $w_{x_iy_i}$. s^2 is "introduced to conserve surface density" while the output image is defined by the pixels (x_0, y_0) " with value $I_{x_0y_0}$, weight $W_{x_0y_0}$, and fractional pixel overlap $0 < a_{x_0y_0x_iy_i} < 1$". $W_{x_0y_0}$ is computed as described in Equation 4 of F02: $Wx_0y_0 = a_{x_0y_0x_iy_i}w_{x_iy_i}$. Here we have implicitly made use of the Einstein notation, thus rewriting explicitly equation 4.1 we obtain:

$$I_{x_0y_0} = \frac{\sum_i d_{x_iy_i}a_{x_0y_0x_iy_i}w_{x_iy_i}s^2}{\sum_i a_{x_0y_0x_iy_i}w_{x_iy_i}}. \tag{4.2}$$

Following F02, the uncertainty of the pixel for the drizzled image is described in Equation 7:

$$\sigma_p^2 = \frac{\sum_{d_{xy}\in\mathcal{P}} a_{xy}^2 w_{xy}^2 s^4 \sigma_{xy}^2}{(\sum_{d_{xy}\in\mathcal{P}} a_{xy}w_{xy})^2}, \tag{4.3}$$

where $\mathcal{P}$ refers to the "set of all input pixels whose drops overlap with a given output pixel " and σ_{xy} "is the standard deviation of the noise distribution of the input pixel d_{xy}". Note the slight change in notation. Also note that the correlated noise only enters the computation when considering a selection of pixels, rather than a single one, and thus σ_p^2 is, for now, a valid estimation of the variance of the noise per pixel.

What is clear from these two equations, however, is that giving as input the error frame, i.e. σ_{xy} to `astrodrizzle` does not provide the proper "drizzled error frame". It would be possible to correct for this effect and still use the `astrodrizzle` pipeline, but to avoid complications, I instead followed a simpler approach, based on standard Gaussian error propagation:

$$\sigma_{xy}\left[\frac{\text{e}-}{\text{sec}}\right] = \frac{\sqrt{(d_{xy} + \text{sky})[\frac{\text{e}-}{\text{sec}}] \times \text{TEM[sec]} + \text{RN}^2[\text{e}-]}}{\text{TEM [sec]}}, \tag{4.4}$$

where TEM is the Time Exposure Map and RN is the Read-Out Noise . The sky is the median of the σ-clipped drizzled image and is subtracted from the image, obtaining $d_{xy} = d'_{xy} - \text{sky}$, where d'_{xy} is the default output science frame obtained after drizzling. d_{xy} and σ_{xy} are then the value and uncertainty of pixel (x, y) in units of electrons per second.

The images later used for the analysis are then cropped, taking a square of $\sim 6''x6''$ around the main galaxy. This, along with the masking later described in Section 4.2.5, ensures that the modelling only takes into account the most informative pixels and reduces the memory requirements.

4.2.3 PSF Modelling

During the lens modelling process, a parametric model is defined which must be convolved with a PSF kernel to be properly compared with the observation. In `lenstronomy`, the PSF kernel can be given as a pixellated model, which is the optimal approach compared to analytic profiles. Moreover, due to the high precision required, the PSF should be supersampled to maximise astrometric precision and best fit the model. This step is of great relevance for the accuracy of the model, as seen in (Shajib et al., 2019). By construction, the PSF model mediates the comparison between a given parametric model and the observations and, therefore, is crucial for high-precision results.
In this analysis, I therefore opted to obtain supersampled PSF kernels. Note that the procedure described here is applied to all *HST* images as well as the ground-based observations later analysed (see Chapter 5). The PSF kernel is obtained using the software `psf` (Riffeser, 2006), which outputs the resulting pixelated profile by supersampling the input point source (e.g. stars or QSO) in a flux-conserving manner. If multiple inputs are given, those are then stacked to increase the SNR. Therefore, such a method is most effective when more such sources are available in the field close to the observed target. In the case under analysis, the feature most affected by the PSF precision is the QSO image position. In fact, this parameter informs the lens mass model and indicates where to measure the Fermat potential. Therefore, it has to be constrained with the highest accuracy and precision. The best option is to model the PSF kernel over the QSO images themselves. This is not always possible in practice due to several reasons, such as low brightness of the image and blending with other light profiles, both originating from the source and the lens plane. In particular, image D is too dim and too blended with the lens light to be considered in any filter. Similarly, image C is, in turn, too blended with the perturber light profile. Images A and B can be used for the optical filters and F105W, although only after modelling and subtracting the lens light, as described in Section 4.3.2. However, this is not applicable to the remaining filters, F140W and F160W. In these filters, the resolution is lower (see Table 4.2), the lens light appears brighter, and the host galaxy of the QSO appears lensed in arclets at the image position. Thus, for the PSF modelling in these filters, I considered nearby stars (four for F140w and three for F160W), selecting the closest and brightest available in the field. The vicinity in the image allows for minimising the variability of the PSF due to optical effects (see e.g. Howell et al., 1996, , Figure 3), while the high brightness increases the SNR. Note, however, that the stars considered must not be saturated. The use of stars as the model for the PSF is, in

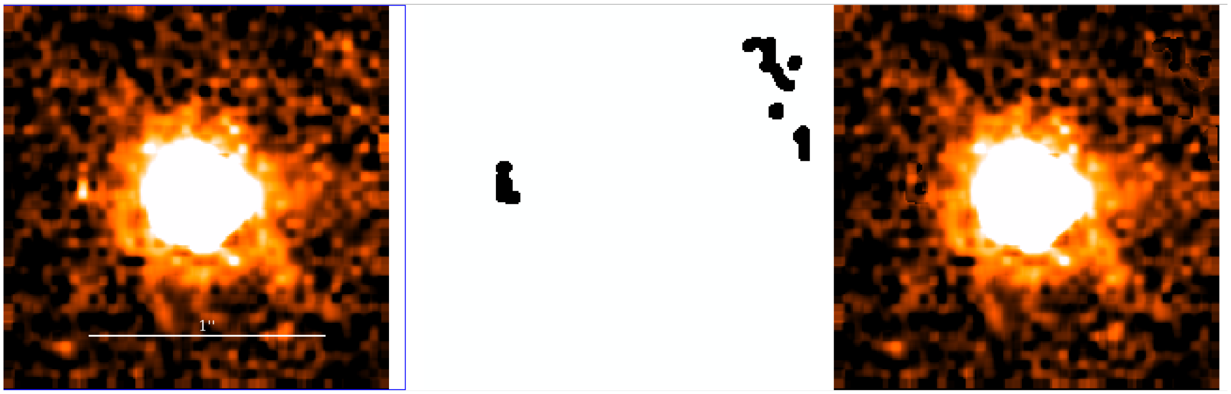

Figure 4.4: The correction of contaminants in the PSF model for F814W. From left to right: the original PSF model, the mask considered and the final PSF model.

principle, a sub-optimal choice, as their colours differ significantly from the QSO images and thus the obtained PSF profile would also be different, as the PSF is colour-dependent. However, this effect is mitigated by the low resolution of these filters, in which the mismatch of the PSF profile is negligible. Moreover, the modelling of the QSO position will be prior-dominated by the constraints obtained from the optical filters. Thus, the effect of the PSF on the model would be mostly present in the convolution of the arclets and the perturber, which share a redder colour. The difference in colour between the modelled objects and the stars used for the PSF model is therefore mitigated.
Due to the different approaches, the size of the PSF kernel differed mildly between the filters. For most filters, this amounted to 41 pixels, while for F105W it was 31 pixels. This was limited by the presence of straylight (originating from the perturber, other QSO images and residuals of the main lens light subtraction) at larger distances from the QSO images. This is exacerbated in F105W due to its lower pixel resolution. In general, a larger PSF kernel allows for more precise modelling of the outer wings and diffraction spikes, but also increases the computational cost.
To further improve the quality of such PSF models, I implemented further steps. Firstly, I verify by visual inspection that there are no further light contaminations, such as small secondary light sources or straylight. Such contamination has to be small in amplitude and in size, i.e. $\lessapprox 1/100$ of the maximum of the PSF and only affecting a few tens of pixels. I then approximate the PSF to be circular symmetric and replace the affected pixels with their corresponding pixels at a given angle of rotation with respect to the centre of the PSF kernel. Usually, the angle should be taken to be 180, so as to select the circularly symmetrical opposite pixels, but this is sometimes not possible due to the presence of overlapping contaminants on both pixels; thus, also rotations of 90 degrees are used. The same substitution is carried out in the corresponding pixels of the error frame. This correction only affects a few pixels per PSF kernel, and therefore does not significantly alter the model, while marginally improving the residual. An example of such a process is shown for the PSF of F814W in Figure 4.4.

Secondly, the PSF model of F105W is significantly affected by the presence of external light coming from the residual of the main lens light. I thus subtract a constant value in order to force

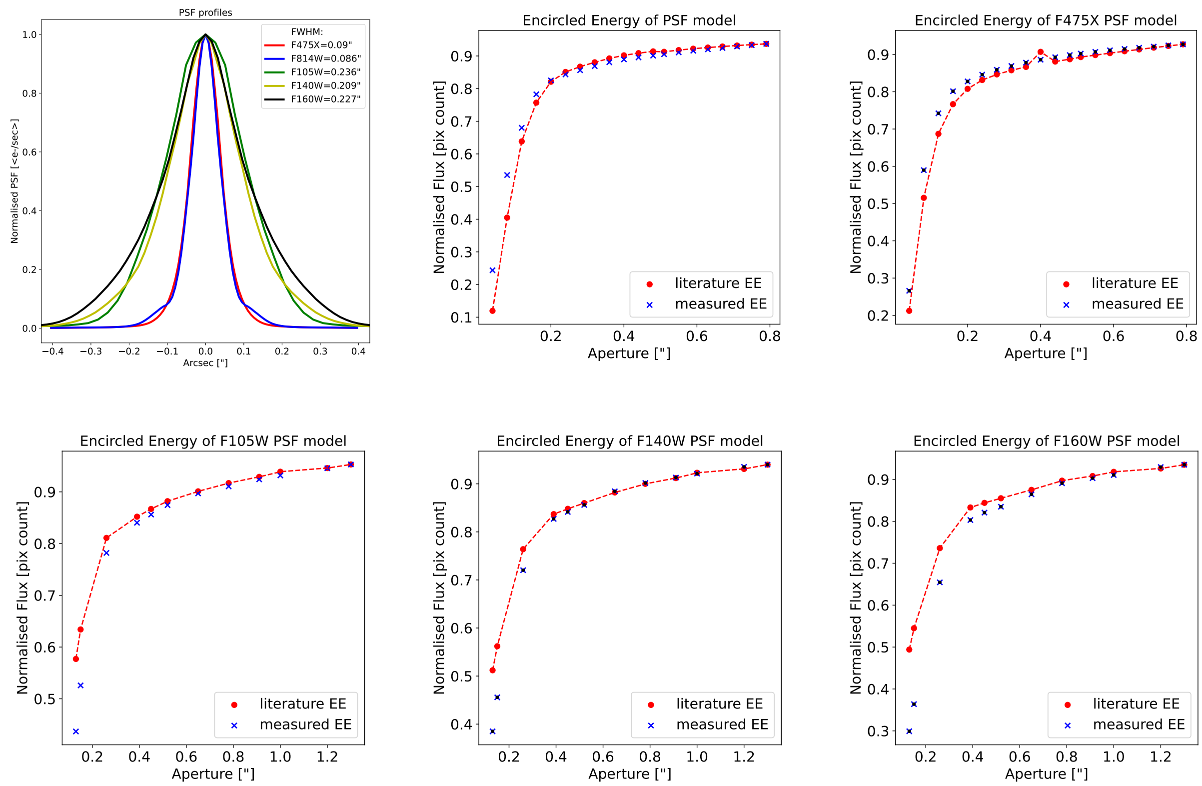

Figure 4.5: PSF models for the various filters. Top Left: PSF profiles normalised to their maximum value. Top Center, Top Right Center and Bottom Row: Encircled Energy (EE) of the PSF model at increasing radii compared to the literature values (STSI, 07/10/2023) for filter F814W, F475X, F105W, F140W and F160W. Note that the literature values of the EE for F475X present a clear outlier at aperture ≈ 4”. This derives directly from the data reported in the literature. While it does not affect the current analysis, it is nonphysical and should be corrected.

the wings of the PSF to converge to zero. The numerical value of this constant is the average of the kernel’s edges. In principle, this approach was considered for the optical filters as well, as the PSF was modelled on the QSO images. However, this correction was not necessary for these filters.

The final test before the modelling itself is carried out is the comparison of the PSF with the reported Encircled Energy (EE, STSI, 07/10/2023) of the given filter. All PSF models appear to agree with the expected profiles, as seen in Figure 4.5, with marginal deviation at the centre for the infrared due to their lower resolution.

In Table 4.3, the Full Width at Half Maximum (FWHM) and the radius at which the Encircled Energy is 50 %, R_{EE}, are reported. These are fundamental parameters to characterise the PSF shape. It is interesting to note that the FWHM does not always increase with the central wavelength λ, which would be expected from a diffraction-limited telescope as *HST*, following $\theta \propto \lambda/d$ (considering the small angle approximation and d as the telescope aperture, constant

Filter	FWHM ["]	R_{EE} ["]
F475X	0.095	0.068
F814W	0.086	0.075
F105W	0.236	0.145
F140W	0.209	0.169
F160W	0.228	0.201

Table 4.3: Table with FWHM of the various PSF models for the different filters and their radius at which the EE=50 %, referred to as R_{EE}.

for all filters). This can be explained by the fact that the FWHM, consistently smaller than $2R_{EE}$, probes the core region of the PSF kernel, which deviates from spherical symmetry. On the other hand, $2R_{EE}$ is larger and thus is mostly affected by the wings, which dominate the energy density of the PSF. This region is less affected by the core asymmetry and recovers the expected dependency on λ. Moreover, this further consolidates our previous assumption of circular symmetry at large radii. It can be further noted that F105W is the greatest outlier, as seen in FWHM and in the profile in Figure 4.5, which is likely due to the aforementioned low exposure time and undersampled image. While all such tests are valid and in general cheap verifications of the PSF model, it is important to remember that the main test for the goodness of the PSF model is the lens modelling itself. Most of the steps discussed here to optimise the PSF had been developed iteratively by improving upon imperfect lens models. A possible future expansion of the analysis could take into account the automated PSF reconstruction available in `lenstronomy`, which iteratively optimises the PSF by minimising the residuals. Initial tests in this direction showed little improvement in the final result. It is yet to be probed if that is due to a suboptimal implementation of the PSF reconstruction algorithm or to an already optimised PSF.

4.2.4 Lens Light Modelling and Subtraction

Before the proper mass modelling, I fitted an isophotal model to the main lens light independently for each filter. There are two main reasons for this. Firstly, the lens light model allows for the subtraction of the light profile of the main lens, which is then discarded from the modelling. This reduces the complexity of the analysis and allows for faster convergence and overall better-constrained parameters. Secondly, the lens light models obtained from all filters can be used to inform the prior of the mass distribution of the galaxy. I will present the details of how this was implemented in the later Section 4.3.5. For now, I will describe how I modelled the lens light and how the results for the various filters can be combined in a common prior.

Isophotal modelling is a well-established technique to fit elliptical galaxies (Jedrzejewski, 1987). In particular, the light modelling of this analysis takes advantage of the fitting procedure presented in Kluge, Bender (2023) and Kluge et al. (2023). The procedure iteratively fits elliptical isophotes at increasing radii to the image. Each isophote is defined by its centre coordinates (x, y), pointing angle ϕ, ellipticity ϵ (related to the axis ratio q as $\epsilon = 1 - q$), major axis a and diskyness/boxiness indicator a_4. Note that given the major axis and the ellipticity, the minor axis b is fixed by the relation $b = aq = a(1 - \epsilon)$. The latter defines the deviations from exact elliptical profiles. The fit is performed from the inner radii outwards, and its shape parameters $(x, y, \phi, \epsilon, a_4)$ are fixed after a certain major axis threshold $a_{\text{fix}}^{1/4} = 1.013\,\text{arcsec}^{1/4}$. This is set as the obtained parametric model is stable enough to fit the extended light beyond a_{fix} while avoiding being biased by unmasked neighbours.

In order to accurately fit the light of the lens, all contaminant sources of light must be eliminated, either by subtraction or masking. Furthermore, to improve the fit, the isophotes models are performed over the whole exposure, thus ranging between $\sim$ 20 to $\sim$ 100 squared arcseconds, depending on the filter. I will, however, concentrate on the region of the lens system.

The most critical light components to eliminate are the closest ones to the centre, in particular the lensed QSO images, the perturber and, when visible, the lensed arclets of the host galaxy. The first can be partially subtracted by employing a PSF model, obtained from neighbouring stars using the aforementioned `psf` program (see Section 4.2.3). The model is then rescaled to match the image brightness and resampled to match the position of the QSO with sub-pixel precision. This subtraction aims to subtract the outer wings of the QSO's PSF. These are more extended and therefore would affect more importantly the isophotal model, while still being subdominant with respect to the luminosity of the lens itself. Given this approach, the fit will present residual at the core, although limited to a few pixels. These are later masked along any other source which can not be subtracted by a point source, i.e. the perturber, the extended host galaxy (if present) and other luminous objects nearby. To mask all external sources, a σ-clipping (Zhang, 1995, see e.g.) with $2 \lessapprox \sigma \lessapprox 5$ is applied to all external regions. Assuming the profile to be elliptically symmetrical, the masked pixels are then "mirrored", i.e. substituted by the symmetrical opposite pixels (unless these are also masked). This limits the effect of masking, which might otherwise significantly impact the modelling. The steps are shown in Figure 4.6 for F160W. For this filter in particular, it was necessary to iterate the process in order to accurately mask the spurious light of the arclets and perturber, as in the first iteration, masking was hindered by the presence of the lens, which was strongly blended with these components. In practice, this is done by doing a first lens light model. The lens light is then subtracted from the image, where the QSO images have been subtracted. Here, every pixel which shows significant residuals due to the presence of external light is masked. The resulting mask is then used for the second and final lens model, which is the one shown in Figure 4.6.

The other filters were less problematic, mostly thanks to the lower wavelengths and therefore lower brightness of the arcs and of the lens galaxy itself. The whole set of the resulting models is shown in Figure 4.7

Now it is interesting to study further the resulting distribution of parameters obtained from such fit, in order to define an informative prior for the lens mass model. This analysis provides for each isophote the best fit of the parameters, along with the estimation of their uncertainty. Plotting the following for all filters with respect to the major axis a, we obtain Figure 4.8.

The threshold for fixing the parameter is $a_{\text{fix}}^{1/4} = 1.013\,\text{arcsec}^{1/4}$, or $a_{\text{fix}} = 1.053''$, depending on the filter and corresponding brightness of the lens. The lensed QSO images are located in the

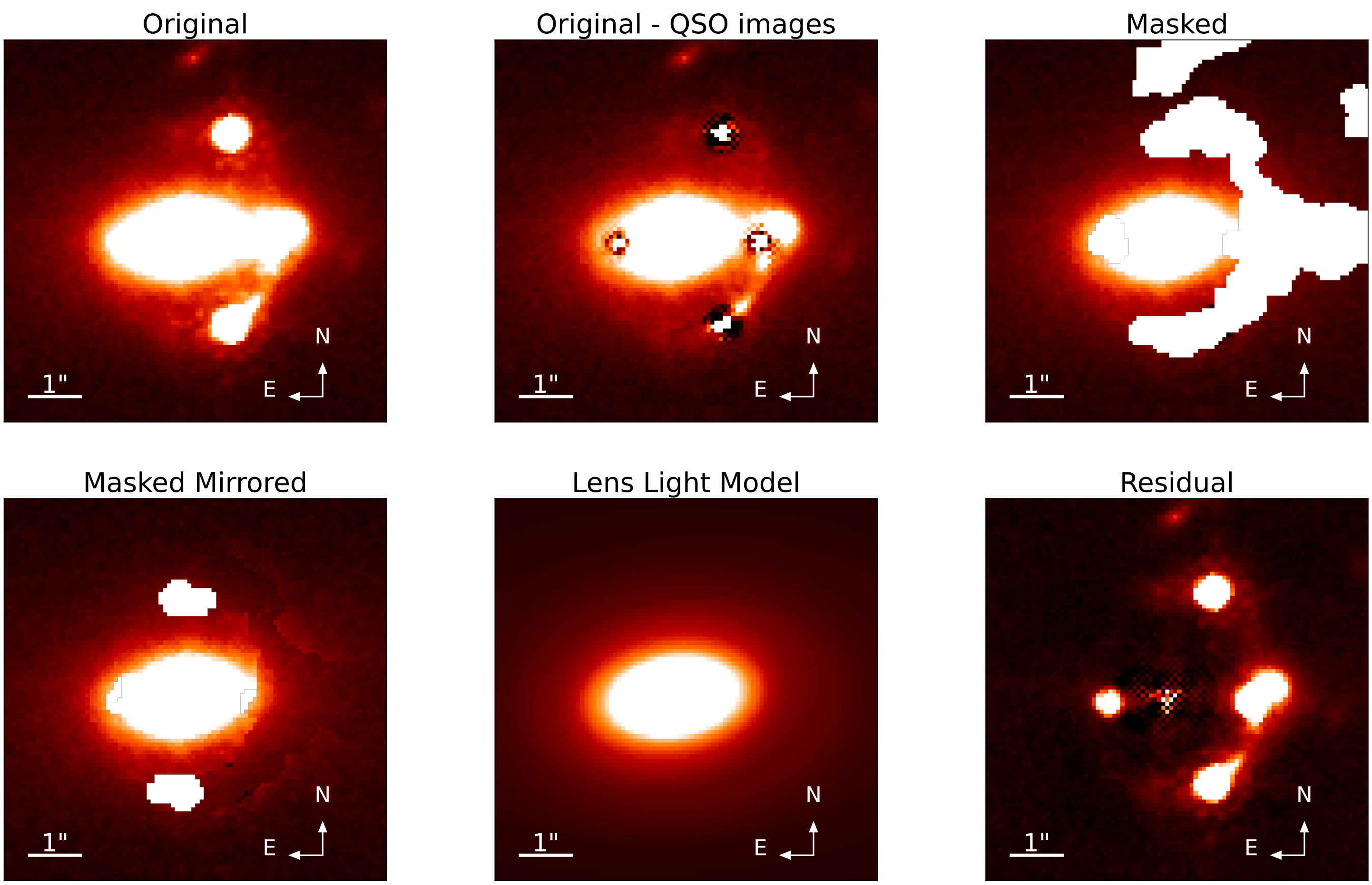

Figure 4.6: The required steps for the isophotal lens light modelling of F160W. Note that only the second and final iterations are shown here, as the first differs only in the masking and the final result.

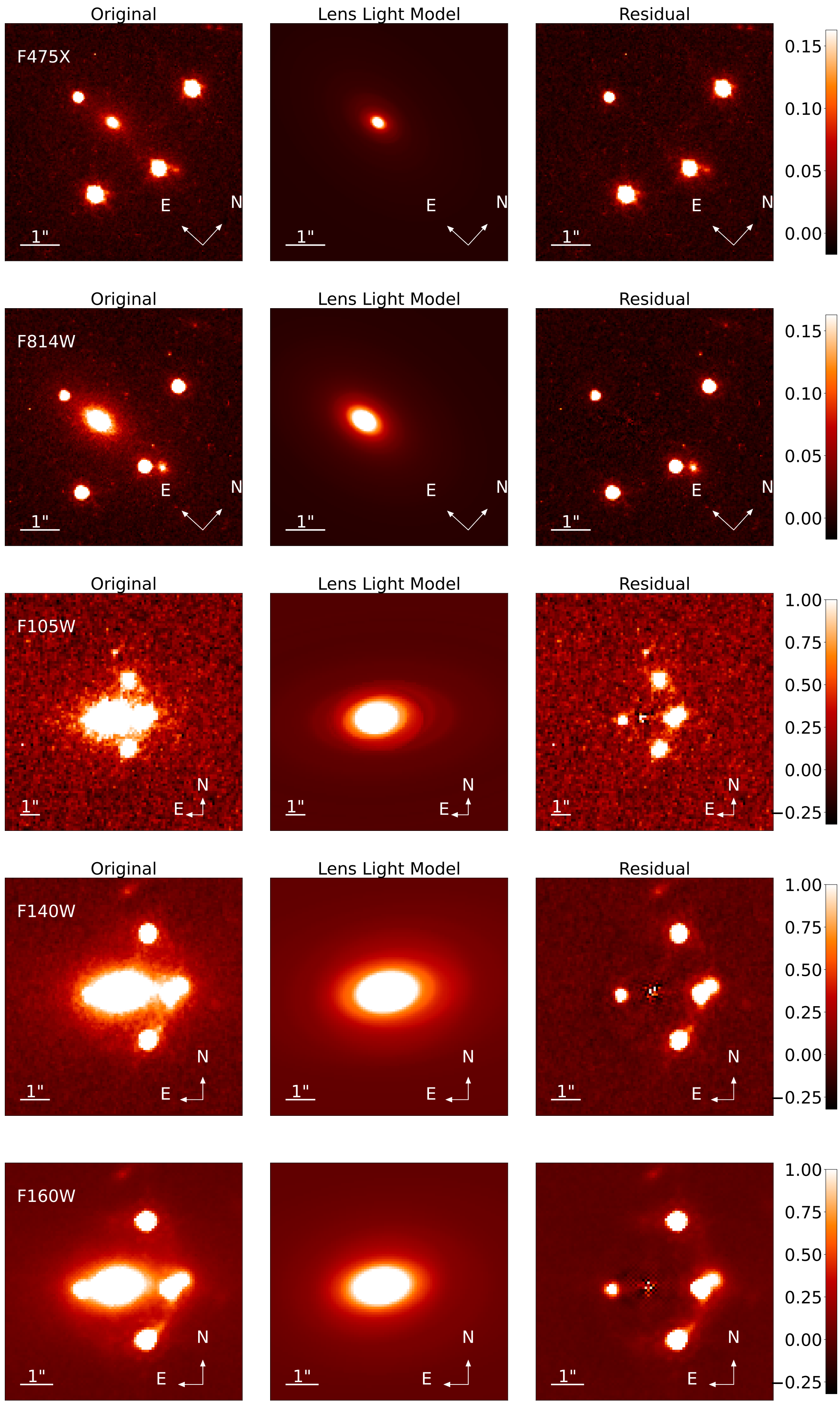

Figure 4.7: The resulting isophotal lens light modelling for the remaining filters.

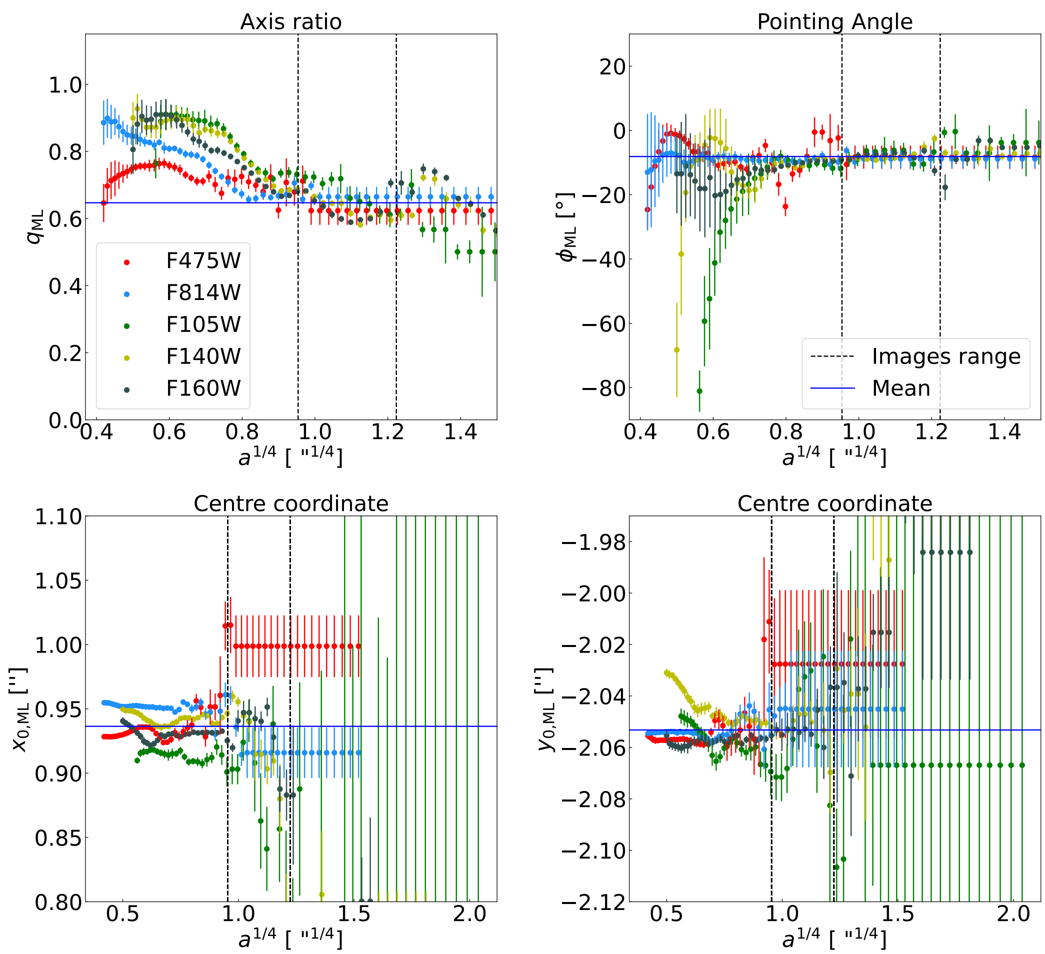

Figure 4.8: Parametric results of the isophotal fitting for the main lens light with respect to the axis ratio to the power of 1/4, $a^{1/4}$: axis ratio q_{ML}, pointing angle ϕ_{ML} and coordinates x/y_{ML}. For each plot, the vertical lines indicate the range within which the QSO images are located. The horizontal blue line indicates the average of the given parameter taken within the QSO images region. These values are then considered as additive information to the likelihood during the lens mass modelling, as later described in Section 4.3.5. Differences in starting value for $a^{1/4}$ are due to the varying pixel resolution between the different filters.

Parameter	Value
$\langle q_{\rm LL} \rangle$ [ ]	0.65 ± 0.04
$\langle \phi_{\rm LL} \rangle$ [°]	-8.1 ± 1.6
$\langle x_{\rm LL} \rangle$ ["]	0.94 ± 0.01
$\langle y_{\rm LL} \rangle$ ["]	-2.05 ± 0.01

Table 4.4: The resulting values obtained for the axis ratio $q_{\rm LL}$, the pointing angle $\phi_{\rm LL}$ and the central coordinates $\vec{c}_{\rm LL} = (x_{\rm LL}, y_{\rm LL})$ of the main lens luminous profile from the isophotal as shown in the Figure 4.8. Note that the coordinates are defined with respect to image A.

range 1.1 arcsec$^{1/4} \lessapprox a^{1/4} \lessapprox 1.25$ arcsec$^{1/4}$, thus after fixing the shapes parameters. This comes to the advantage of the fitting, as this region is heavily masked for most of the filters, and thus the fit might be underconstrained if all parameters were allowed to vary. Moreover, this is the most informative region, as it is where all lensed images (QSO and host galaxy) are located, and thus most of the constraining power of the data is contained within it. For this reason, when taking the average of the shape parameter (q, ϕ and coordinates x and y), I only consider the isophotes within this range. The reason is that the model used to describe the mass distribution is only a valid approximation in a limited region around the images' positions (usually indicated by the Einstein angle $\theta_{\rm E}$). Therefore, any correlation between the mass shape and position should be done with respect to the light distribution at a corresponding radius. The obtained average is reported in Table 4.4, where the uncertainties are obtained by error propagation of the parametric uncertainty, but are not considered further for the model.

For completeness, I report here also the a_4 parameter, although this will not be further considered in the analysis. As shown in Figure 4.9, the results are also in agreement in this parameter, and the average is compatible with zero.

4.2.5 Masking

In order to limit the modelling to the lens system and avoid being biased by the light of nearby sources, interlopers and residuals of the lens light subtraction, a mask file is given as input to `lenstronomy`. The mask is a boolean pixel map representing which pixels are considered for the modelling. This step becomes particularly significant due to the pre-subtraction of the lens light presented in Section 4.2.4. This leaves residuals which can be seen by the eye in the centre of the image, as seen in Figure 4.7 and Figure 4.6, especially due to the lower resolution of the infrared filters. Given that these filters will be considered for the modelling, it is therefore important to mask all traces of light which do not belong to the modelling. I then focused on F160W, being the most informative filter due to its high wavelength and thus extended source, as well as having the longest exposure time and therefore significant SNR. I created a large mask, which covers most of the central region of the lens and everything outside the location of the QSO images and the arclets. The comparison between the original image and the masked one is presented in Figure

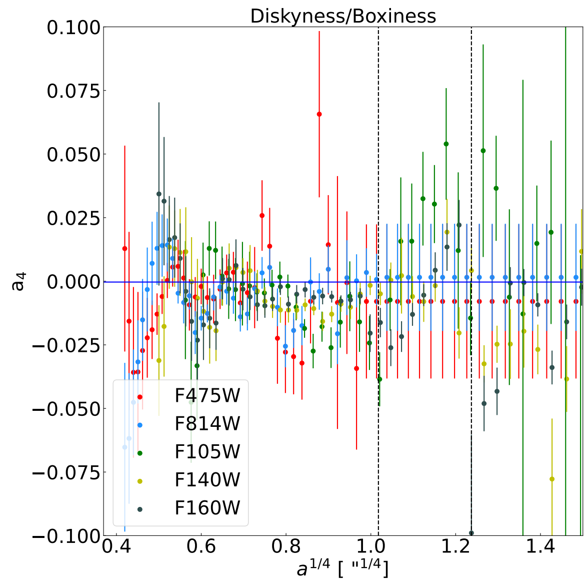

Figure 4.9: The diskyness/boxiness parameter a_4 obtained from the isophotes model in the different filters. We can see that the result agrees with 0 for all filters.

4.10.

To limit the variation introduced in the modelling of the different filters, an equivalent mask is applied to F140W. The mask transformation from one frame of reference to the other is obtained by aligning the file using its relative coordinate information. The change is minimal, however, thanks to the identical drizzling procedure which ensures that the two images share a similar pixel grid. The result is also shown in Figure 4.10.

A particular region of interest for the masking is later discussed in Section 4.4.1 in connection with an excess of light visible in F160W.

4.3 Mass and Light Profiles

In modelling gravitational lenses, our primary focus is on the overall mass distribution of the lens, encompassing both luminous (baryonic) matter and dark matter. While the observed light can provide insights into the mass distribution, it must be taken with caution. Given the presence of dark matter, the principle that 'light traces matter' can not be applied unconditionally. Instead, informed by the results of the lens modelling of the Sloan Lens ACS (SLACS) Survey lenses Bolton et al. (2006, 2008); Auger et al. (2010), I followed a similar approach as described in Schmidt et al. (2022), i.e. adding limited constraints to the mass model from the lens light profile.

In this section, I discuss the profiles used for the mass (see Section 4.3.1) and light (see Section 4.3.2) modelling of the system. I also explain the ties between said profiles implemented by joining some parameters (see Section 4.3.3). I then describe the prior uses for this parametric modelling (see Section 4.3.4) and how the likelihood is computed (see Section 4.3.5).

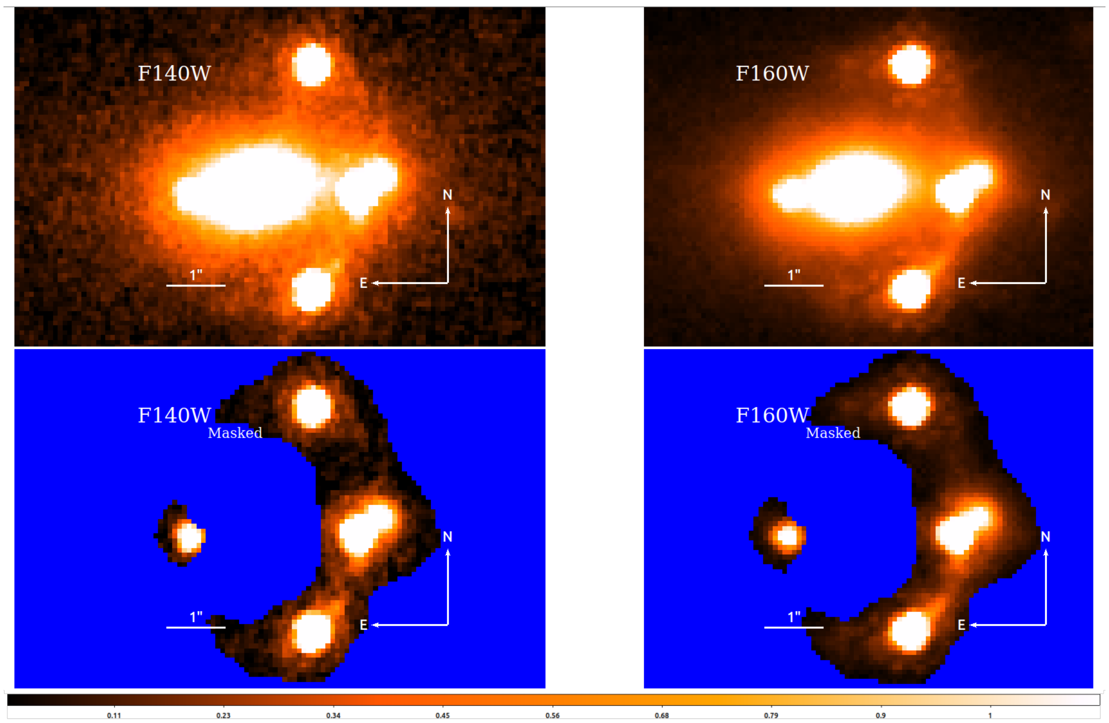

Figure 4.10: Mask for the two filters F140W and F160W. The mask is defined for the F160W image and applied to both filters, considering a coordinate transformation to adapt to the small differences in the pixel grid. Note that this is the initial mask, which will be updated as discussed in Section 4.4.1

4.3.1 Mass Profiles

The most common approach for the mass model of galaxy-scale lenses is to consider a parametric profile for the mass density. This is due to the simplicity of such profiles and corresponding low computational cost, as well as the empirical confirmation of optimal results in reconstructing the observed lens configuration with said models. For example, the Singular Isothermal Spherical (SIS) profile is a special type of profile:

$$\kappa(x, y) = \frac{\theta_E}{\sqrt{x^2 + y^2}} = \frac{\theta_E}{r}. \tag{4.5}$$

While being unphysical due to its singularity at $r = 0$ and the fact that the total integrated mass diverges at infinity, this profile has several advantages. Firstly, its simplicity makes it one of the first density profiles to be defined and used. Secondly, this profile is analytically obtained by considering a distribution of massive point particles in thermal equilibrium. The relation between its velocity dispersion and the Einstein angle can therefore be obtained: $\sigma_{\mathrm{v}} = c\sqrt{\theta_E \frac{D_{\mathrm{s}}}{4\pi D_{\mathrm{ds}}}}$, where D_{s} and D_{ds} represent the angular diameter distance between the observer and the source and the lens and the source, respectively. Lastly, although simplistic, this profile is in rough agreement with several astrophysical structures, such as the stellar dynamics and X-ray halo of elliptical galaxies, the model of individual lenses and joint lensing and dynamical analyses, see Keeton (2001). The natural evolution of the SIS profile is represented by additional degrees of freedom. Firstly, the ellipticity of the profile might be considered. Secondly, the power of the profile γ can be set free to vary. When considering both these degrees of freedom, we obtain the elliptical power law profile:

$$\kappa(x, y) = \frac{3-\gamma}{2}\left(\frac{\theta_E}{\sqrt{qx^2 + y^2/q}}\right)^{\gamma-1}. \tag{4.6}$$

Note that the axis ratio q is the only ellipticity parameter written explicitly. However, the pointing angle ϕ is also implicitly present, as it can be defined by the orientation of the coordinate system, which in equation 4.6 is aligned with the major and minor axes of the lens. A different approach can be obtained by considering composite models, consisting of an elliptical NFW profile (Navarro et al., 1997) for the dark matter halo and a baryonic component linked to the observed light distribution by a scaling factor. This approach, for example, has been implemented in the lens modelling of the TDCOSMO collaboration (e.g. Rusu et al., 2020), where the light distribution is modelled by a Chameleon (or pseudo-Jaffe) profile Kassiola, Kovner (1993); Dutton et al. (2011) and scaled by a mass-to-light ratio factor.

In this study, I will limit the analysis to an elliptical power law for the main lens profile. Additionally, the perturber is also considered to be a significant component of the mass distribution. This is modelled by a SIS, as it appears to be too small to be fitted by a more complex model. Finally, to account for the effect of external massive components which are not explicitly modelled, I consider a two-component external cartesian shear (see Section 2.2.7). This represent a shear field defined by strenght γ^{Shear} and direction ψ^{Shear} .

Note, however, that I will not take into account the mass-sheet degeneracy (MSD), i.e. the external convergence, described in Section 2.3.3. This is due to the lack of two fundamental datasets for such a study: a 2D spectroscopically derived stellar velocity dispersion of the lens system, which would inform the mass distribution within the system, and a map of the neighbourhood of the system (i.e. weighed number count of the nearby galaxies), which can constrain the external component of the convergence. For this reason, is important to consider that this analysis is still affected by this unconstrained bias, and a future improvement upon this work must consider it. Also note that the choice of a fixed mass profile formally breaks the mass sheet degeneracy (e.g. Schneider, Sluse, 2013), as the MSD transformation described in equation 2.80 does not return an exact power law.

To conclude, while the choice of lens models here described appears successful in fitting the data (see Section 4.6), further study can be done to test the validity of other models.

4.3.2 Light Profiles

I considered multiple profiles to fit the light observed in the images. These are the main lens, the perturber and the "sky brightness" from the lens plane, while the QSO images and the arclets of the host galaxy are from the source plane. As described in Section 4.2.4, the main lens light is modelled and subtracted a priori, and is not modelled further. The perturber is modelled as a circular Sérsic profile, whose centre is fixed (see following Section) to be the centre of the corresponding mass profile described previously in Section 4.3.1. The "sky brightness" is, in principle, subtracted a priori, see Section 4.2.2. However, I still consider it here as an ulterior free parameter to take into account any residual of the lens light subtraction. This is modelled as a uniform background over all pixels, and in all models, it has a negligible impact on the final result. It might be discarded in further analyses without significant loss of accuracy. For the source plane, the QSO is modelled as a point source, while the host galaxy is represented by a circular Sérsic profile Sérsic (1963):

$$I(r) = I_e \exp\{-b_n[(\frac{r}{R_e})^{\frac{1}{n}} - 1]\}, \tag{4.7}$$

where I_e is the intensity at the half-light radius R_e, n is the "Sérsic index" which controls the curvature of the profile and b_n is a function of n.

As previously mentioned, the host galaxy was only visible in the NIR filters, F140W and F160W, and was therefore modelled in these filters only. Additionally, the centres of the host galaxy and the QSO are fixed to coincide in the source plane, as we expect the QSO to be powered by the supermassive black hole residing at the centre of the host galaxy. A more complex profile for the host galaxy might be considered, as adding shapelets (Refregier, 2003; Birrer et al., 2015). These functions are localised basis functions with varying shapes, designed to capture small variations in luminosity, adding detail to the otherwise smooth profile of the source. For this system, preliminary tests in which the addition of shapelets was considered did result in overfitting. This was characterised by a lower χ^2 and unphysical properties of the source (such as negative flux). However, while discarded for the current analysis, it might be interesting to test this fitting procedure further. Similarly, an elliptical Sérsic profile was also considered, but it proved to be too little information in the lensed arclets for an accurate model, and it also resulted in an overfit of the data. Given that the simple addition of these degrees of freedom resulted in overfitting, I did not take into account more complex profiles.

4.3.3 Joint Parameters

The program `lenstronomy` allows users to join parameters between different profiles for both the luminous and massive components. This ensures that the values of these parameters are consistently equated at each step of the analysis. As mentioned previously, for this system, the following parameters were joined:

- the coordinates of the centre of the host galaxy and the QSO have to coincide in the source plane,
- the coordinates of the centre of the perturber light profile and its mass profile have to be the same.

The first point is physically motivated, as explained in Section 4.3.2. The second, tying the mass and the light centre of the perturber, is instead data-driven. The perturber is not significantly extended, ranging to $\approx 0.25''$, thus $\approx$ 1.4kpc on the lens plane, assuming ΛCDM cosmology (with $\Omega_{m,0} = 0.3$, and $H_0 = 70\frac{\mathrm{km}}{\mathrm{Mpc\ s}}$, as described in the Section 4.1). Therefore, while it is unlikely that its dark matter and luminous profiles are perfectly aligned, I assumed that the deviation should be negligible given the available resolution. While this approach was motivated by the perturber, I decided against fixing the centre of the main lens mass to the centre of the isophotal model previously obtained (see Section 4.2.4). This is to avoid overconstraining the model and because a shift between the centre of the light and the mass might be physically motivated due to the dark matter halo. In order to weakly inform the centre of the mass profile, I instead added an additional likelihood term, later described in Section 4.3.5.

4.3.4 Priors

I present here the resulting parameter of the profile considered in the model and its prior. The prior ranges were initially physically motivated or based on previous results, such as (Agnello et al., 2018), (Shajib et al., 2019) and (Schmidt et al., 2022). However, they were iteratively optimised over time to expedite model convergence by avoiding extreme parameter ranges and to enhance the model constraints by applying tight and informative priors for the QSO image positions.

Given the individual modelling of each filter, it is paramount to define a prior which is common to all models. For the two NIR filters, the priors are identical due to the similarities of the data. The optical filters, which are modelled as a verification, differed only in the lack of the host galaxy. Note, however, that this analysis is focused on constraining the Fermat potential differences $\boldsymbol{\Delta\phi}$, thus, the prior might differ when it comes to "noise terms" such as, in this case, the presence of the host galaxy. It will be shown in Figure 4.11 that this difference does not affect the final prior, as the prior $P(\boldsymbol{\Delta\phi})$ is compatible between all modelled filters. This derives from the fact that for the mass profiles, all filters are modelled consistently. For the main lens (ML) the elliptical power law is described by the Einstein radius θ_E^{ML}, the exponent of the power law γ^{ML}, the two polar components of the ellipticity e_1^{ML} and e_2^{ML}, the centre coordinates x^{ML} and y^{ML}. Similarly, for the perturber (P) we have θ_E^{P}, x^{P} and y^{P}. Finally, the two polar components of external shear are indicated by γ^{Shear} and ψ^{Shear}.

Additionally, there are eight positional free parameters for the QSO's image positions in the image plane, x_i^{QSO} and y_i^{QSO} for $i = \{A, B, C, D\}$, along with their independent intensities $I_{0,i}^{\mathrm{QSO}}$ (see Section 4.3.3). Furthermore, the model requires additional parameters for the luminous profiles: for the perturber and host galaxy (HG) Sérsic profiles there are the half-light radii R^{P}, R^{HG}, the Sérsic indexes n^{P}, n^{HG}, and the central intensities I_0^{P}, and I_0^{HG}, respectively. Such profiles normally also require two coordinates for the centre position; however, these parameters are not free. As previously mentioned in Section 4.3.3 they are defined to coincide with the centre of the mass profile (for the perturber) and with the backwards ray-traced position of the QSO images (for the host galaxy).

Finally, the uniform background has one component, a constant intensity I^{Bkg}. The parameters used in the modelling are summarised in Table 4.5.

By default, the priors in `lenstronomy` are uniform, and for most of the parameters considered, I maintained such an uninformative approach. However, I followed a different approach for what concerns the positions of the various components. The reasons stem from the aforementioned fact that the optical filters had the highest resolution while containing the least amount of information due to the lack of visible arclets of the host galaxy. Thus, they were the perfect candidates to constrain the positions of the QSO images, as these were unperturbed by the additional light of the lensed host galaxy. However, the frame of reference (FoR) had to be aligned with high precision between the different filters to compare the coordinates between them. Firstly, in each filter, the QSO positions are measured using `SExtractor` (Bertin, Arnouts, 1996) and their relative PSF model obtained in Section 4.2.3. The first order correction is to subtract the position of the QSO image A, which is one of the brightest and least perturbed images. This image will then be considered as the reference point. Due to the shared high resolution, it can be seen that the optical coordinates are in agreement with an average 0.0018" difference, or 0.04 pixels in the opticals. I take the average of these coordinates as the reference position $\vec{c}_{\langle\mathrm{VIS}\rangle} = (x^{\mathrm{QSO}}_{\langle\mathrm{VIS}\rangle,\mathrm{i}}, y^{\mathrm{QSO}}_{\langle\mathrm{VIS}\rangle,\mathrm{i}})$, where $x^{\mathrm{QSO}}_{\langle\mathrm{VIS}\rangle,\mathrm{i}} = \langle x_{\mathrm{i}}^{\mathrm{QSO}}\rangle_{\mathrm{VIS}}$ and $y^{\mathrm{QSO}}_{\langle\mathrm{VIS}\rangle,\mathrm{i}} = \langle y_{\mathrm{i}}^{\mathrm{QSO}}\rangle_{\mathrm{VIS}}$ are the average of the corresponding coordinates for i-th image measured from the optical filters. While $\vec{c}_{\langle\mathrm{VIS}\rangle,\mathrm{i}}$ is defined as the centre of the prior for the QSO positions in all models, I also use it to align the FoR. In order to accomplish this, I compare $\vec{c}_{\langle\mathrm{VIS}\rangle,\mathrm{i}}$ with the corresponding coordinates of the i-th QSO image for each given NIR j-th frame, $\vec{c}_{\mathrm{NIR_j},\mathrm{i}} = (x_{\mathrm{NIR_j},\mathrm{i}}, y_{\mathrm{NIR_j},\mathrm{i}})$ (excluding image C, which in the NIR is often too severely blended with the perturber to be accurately measured). I first measure and correct for the difference between the centroids of these coordinates, thus further correcting for a shift. This is a minor correction since taking image A as the FoR origin already corrects for it at the first order. Secondly, I run a Kabsch algorithm (Kabsch, 1976) between the resulting positions to measure the rotation matrix. Both the centroid shift and the rotation matrix are used to correct the FoR of the j-th frame in order to be compatible with the opticals.

The prior of the QSO images is then defined to be a Gaussian prior, centred on $\vec{c}_{\langle\mathrm{VIS}\rangle,\mathrm{i}}$

Model	Parameter	Variable			
Filters		F475X	F814W	F140W	F160W
PEMD$^{\mathrm{ML}}$	Centre	x^{ML}			
		y^{ML}			
	Einstein radius	$\theta_{\mathrm{E}}^{\mathrm{ML}}$			
	Power	γ^{ML}			
	Polar Ellipticity	e_1^{ML}			
		e_2^{ML}			
SIS$^{\mathrm{P}}$	Centre	x^{P}			
		y^{P}			
	Einstein radius	$\theta_{\mathrm{E}}^{\mathrm{P}}$			
External Shear	Shear Power	γ^{Shear}			
	Shear Angle	ψ^{Shear}			
QSO	X Coord.s images (i={ABCD})	$x_{\mathrm{i}}^{\mathrm{QSO}}$			
	Y Coord.s images (i={ABCD})	$y_{\mathrm{i}}^{\mathrm{QSO}}$			
Sérsic$^{\mathrm{P}}$	Half-light Radius	R^{P}			
	Sérsic index	n^{P}			
	Intensity	I_0^{P}			
Sérsic$^{\mathrm{HG}}$	Half-light Radius	//	//	R^{HG}	
	Sérsic index	//	//	n^{HG}	
	Intensity	//	//	I_0^{HG}	
QSO Intensities	Intensities (i={ABCD})	$I_{\mathrm{i}}^{\mathrm{QSO}}$			
Background	Intensity	I_0^{Bkg}			
Total N of Parameters		27		30	
N of non-linear Parameters		20		23	
N of used Pixels		5636	5601	1667	1520
Degree of Freedom (DoF)		5609	5574	1637	1490

Table 4.5: Parameter used for the lens mass modelling. The double lines separate the list in mass parameters, light parameters correlated to the mass profiles and light parameters independent from the mass profiles. ML refers to the Main Lens (mass), P to the Perturber (mass and light) and HG to the Host Galaxy of the QSO (light). All intensities (I_0^{P}, I_0^{HG}, I_0^{Bkg} and $I_{\mathrm{i}}^{\mathrm{QSO}}$) are linear parameters and are not sampled explicitly by the non-linear solver (see Section 4.4). The centre of the Sérsic profile of the host galaxy is not a free parameter, as it is defined to be identical to the position of the QSO in the source plane (Section 4.3.3). The number of pixels refers to the non-masked pixels, and the DoF results from the difference between pixels and the total number of parameters.

with a standard deviation of 0.004" (equivalent to 0.1 pixels in the optical filters). The prior is furthermore implemented within a certain range, thus effectively representing a truncated normal distribution. The range is defined to be $\vec{c}_{\langle\mathrm{VIS}\rangle,\mathrm{i}}\pm0.012$" (equivalent to 0.3 pixels in the optical filters). These hyperparameters are selected with a slightly less stringent constraint to account for any unknown errors in the FoR alignment, primarily influenced by the uncertainty of the QSO position measured from the NIR frames. The precision of these measurements is impacted by the lower resolution and the presence of external lights: arclets, perturbers and main lens light residuals. Thus, the bounds and standard deviation of the prior are larger than the previously mentioned of 0.0018". The other coordinates are the centre of the main lens mass and the centre of the perturber (which is the centre of both its mass and light profile). The first is obtained from the lens light model described in Section 4.2.4. The parametric results of the model, combined with the filters, yield the centre of the light distribution, which we assume to be a reasonable prior for the mass as well. I consider a large prior range to allow for a shift between the mass and the light of 0.198", corresponding to 5 pixels in the optical frames. As described in the following Section 4.3.5, I will also introduce an ulterior likelihood component to tie the centre between mass and light profiles. For what concerns the perturber, the approach is similar to prior of the QSO images. The central position is measured with `SExtractor` in the two optical frames, it is averaged between them obtaining $\vec{c}^{\mathrm{P}}_{\langle\mathrm{VIS}\rangle} = (\langle x^{\mathrm{P}}\rangle_{\mathrm{VIS}}, \langle y^{\mathrm{P}}\rangle_{\mathrm{VIS}})$. This is then used as the centre of the prior for the modelling, considering a uniform prior with a range of $\pm0.04''$, equivalent to 1 pixel in the optical frames. The resulting prior values are reported in Table 4.6.

Note that the QSO positions $\vec{c}^{\mathrm{QSO}}$, once adjusted for the different reference frames, are in optimal agreement with Schmidt et al. (2022), hereafter S22, Table 3. Also, be aware of the different naming conventions, where image D in S22 is named A in this study and vice versa.

Once the prior is defined for the lens parameters, it is straightforward to compute the corresponding prior for the Fermat potential difference at the position of the images $\Delta\phi$. This is done numerically by doing an MCMC sampling of the prior of the lens parameters, as described in Table 4.5 and Table 4.6. For each resulting particle, I compute the corresponding Fermat potential ϕ at each QSO image position. I then take the differences with respect to image A, as in $\Delta\phi_{\mathrm{A}j} = \phi_j - \phi_{\mathrm{A}}$, where j = B, C, D. This results in a three-dimensional distribution of points which samples the prior $P(\boldsymbol{\Delta\phi})$. This is reported in Figure 4.11, and it can be seen that, as expected, the prior is compatible between all models. In this Figure, the prior for the optical filters are reported as well, even though it will not be explicitly used for the modelling. The prior obtained here will be used in Section 4.5 to compute the combined posterior.

4.3.5 Likelihood

In `lenstronomy`, the likelihood is computed as log-likelihood, such that every term results in a simple addition. The default initial value is a simple χ^2 function: $\chi^2_{\mathrm{mod}} = \mathrm{Mask}\cdot(\mathrm{Reconstructed\,Image} - \mathrm{Data})^2/\mathrm{Error\,Frame}^2$. The Reconstructed Image results from the light modelling from the source plane, lensed by the lens mass model on the lens plane, added to the light profile present on the lens plane, and finally convolved with the PSF model. Thus χ^2 depends on all lens model parameters. In addition to this standard term, I considered the

Lens Mass Models			
Model	Parameter	Variable	Prior
PEMD^{ML}	Centre	x^{ML}	$\mathcal{U}$(x_0^{ML} -0.198",x_0^{ML} +0.198")
		y^{ML}	$\mathcal{U}$(y_0^{ML} -0.198",y_0^{ML} +0.198")
	Einstein radius	θ_E^{ML}	$\mathcal{U}$(1.5",1.8")
	Power	γ^{ML}	$\mathcal{U}$(1.5,2.5)
	Polar Ellipticity (1)	e_1^{ML}	$\mathcal{U}$(0.0,0.3)
	Polar Ellipticity (2)	e_2^{ML}	$\mathcal{U}$(-0.12,0.12)
SIS^{P}	Centre	x^{P}	$\mathcal{U}$(x_0^{P} -0.04", x_0^{P} +0.04")
		y^{P}	$\mathcal{U}$(y_0^{P} -0.04", y_0^{P}+0.04")
	Einstein radius	θ_E^{P}	$\mathcal{U}$(0.0",0.35")
External Shear	Shear Power	γ^{Shear}	$\mathcal{U}$(0.0,0.4)
	Shear Angle	ψ^{Shear}	$\mathcal{U}$(-π/2,-0.2π)

Lens Light Models			
Model	Parameter	Variable	Prior
Sérsic^{P}	Half-light Radius	R^{P}	$\mathcal{U}$(0.1",2.0")
	Sérsic index	n^{P}	$\mathcal{U}$(1,6.5)

Host Galaxy Models			
Model	Parameter	Variable	Prior
Sérsic^{HG}	Half-light Radius	R^{HG}	$\mathcal{U}$(0.1",0.5")
	Sérsic index	n^{HG}	$\mathcal{U}$(1,6.5)

QSO Model			
Model	Parameter	Variable	Prior
QSO	Image Coordinates	x_i	$\mathcal{U}$($x_{i,0}$ -0.012",$x_{i,0}$+0.012") $\cdot\mathcal{N}$($x_{i,0}$,0.004")
		y_i	$\mathcal{U}$($y_{i,0}$-0.012", $y_{i,0}$+0.012") $\cdot\mathcal{N}$($y_{i,0}$,0.004")

Table 4.6: Parametric bounds for the Prior implemented in the lens light model. Note that for all coordinates apart from the QSO images, the uniform prior is defined as a function of (x_0^{prof},y_0^{prof}), which are the initial estimates for the (x,y) positions of the given profile. For their numerical values, see Table 4.7. All parameters that are not reported here are either joint parameters (see Section 4.3.3 and Table 4.5) or intensities, which do not have a prior constraint since they were not sampled.

Parameter	Variable	Value[”]
Main Lens Centre	x_0^{ML}	0.9653
	y_0^{ML}	-2.029
Perturber Centre	x_0^{P}	-1.126
	y_0^{P}	-1.863
Image A	$x_{A,0}$	0.000
	$y_{A,0}$	0.000
Image B	$x_{B,0}$	-0.004
	$y_{B,0}$	-3.756
Image C	$x_{C,0}$	-0.762
	$y_{C,0}$	-2.130
Image D	$x_{D,0}$	2.043
	$y_{D,0}$	-2.175

Table 4.7: Starting value for the coordinates of the various luminous components. Note that these are reported in arcseconds with respect to the position of the QSO image A, which is therefore (0,0). The main lens coordinates are obtained from the isophotal modelling described in Section 4.2.4. The QSO positions $\vec{c}^{\mathrm{QSO}}$ and the perturber positions are instead measured in the optical filters.

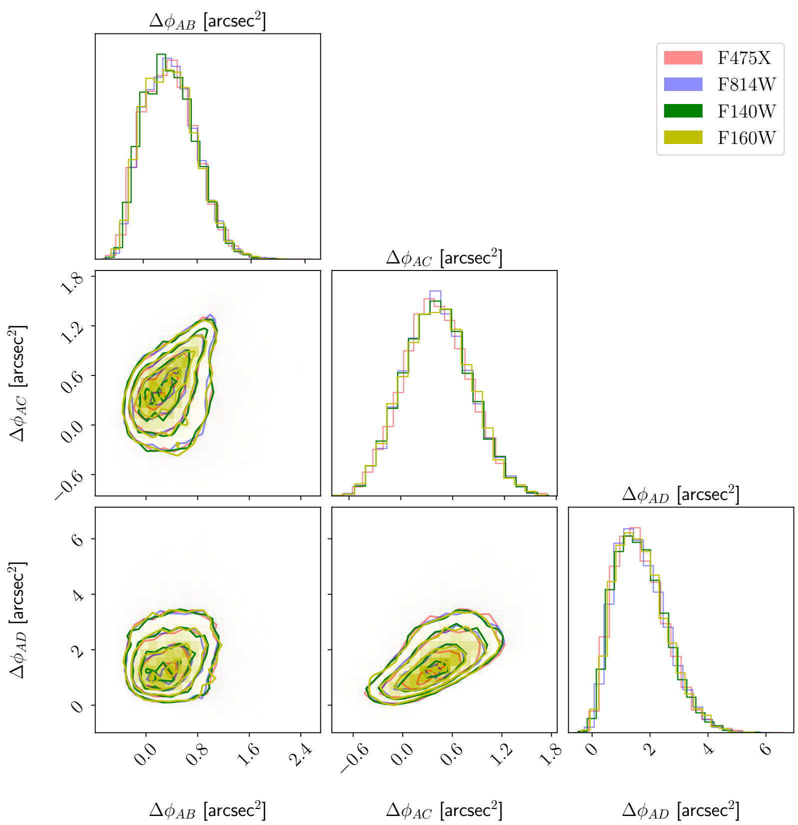

Figure 4.11: Superposition of the corner plot for the sampling of the prior $P(\boldsymbol{\Delta\phi})$ for all models. As expected, these are identical. The prior for the optical filters, since they will not be considered in the final result, are shown for comparison only and thus are transparent.

"`check_matched_source_position`", which adds a punishing factor χ^2_{source} for models in which the backwards ray-traced images do not match in the source plane. In theory, this forces the model to consider each image to have originated from the same source. In practice, this is implemented by taking the positions of the QSO images $\boldsymbol{\theta}$ resulting from the model and assuming a default uncertainty of σ_θ =0.001". The corresponding error matrix is $\Sigma_\theta = \begin{pmatrix} \sigma_\theta & 0 \\ 0 & \sigma_\theta \end{pmatrix}$. By ray-tracing the i-th image $\boldsymbol{\theta}_i$ into the source plane we obtain the source position $\boldsymbol{\beta}_i$ relative to this image. Similarly, by transposing Σ_θ through the Jacobian matrix into the source-plane following equation 13 of (Birrer, Treu, 2019) we obtain the corresponding error matrix Σ_β. Now on the source plane, the χ^2_{source} term is computed as

$$\chi^2_{\mathrm{source}} = \sum_i \Delta\vec{\beta}_i \cdot \Sigma^{-1}_{\beta,i} \cdot \Delta\vec{\beta}_i, \tag{4.8}$$

where $\Sigma^{-1}_{\beta,i}$ is the inverse of the uncertainty matrix $\Sigma_{\beta,i}$ and $\Delta\vec{\beta}_i$ is the difference between $\vec{\beta}_i$, and the mean of all ray-traced images positions, $\langle\vec{\beta}_i\rangle$. Furthermore, the likelihood is computed with a "hard bound" $\sigma_{\beta,\mathrm{hb}} = 0.001''$, meaning that any model resulting in $|\Delta\vec{\beta}_i| > \sigma_{\beta,\mathrm{hb}}$ for any i is discarded, thus forcing the matching of the images.
An ulterior term to the likelihood computation is an additional punishing term that I introduced through the "`custom_logL_addition`" argument. This is used to add information to the mass profile of the main lens from its relative isophotal light model. In particular, this concerns the central position and the ellipticity parameters (axis ratio and pointing angle). The central position of the lens mass (LM) $\vec{c}_{\mathrm{LM}} = (x_{\mathrm{LM}}, y_{\mathrm{LM}})$ is tied to the corresponding coordinates measured from the isophotal lens light $\vec{c}_{\mathrm{LL}} = (\langle x_{\mathrm{LL}}\rangle, \langle y_{\mathrm{LL}}\rangle)$ (reported in Table 4.4) by computing their distance $d_{\mathrm{L}} = \sqrt{(x_{\mathrm{LM}} - \langle x_{\mathrm{LL}}\rangle)^2 + (y_{\mathrm{LM}} - \langle y_{\mathrm{LL}}\rangle)^2}$ and adding a simple gaussian log-likelihood, centred in 0 and with standard deviation σ_d = 0.4".
For the ellipticity parameters, I follow a similar approach as discussed in Schmidt et al. (2022) (S22). In their paper, they underline a weak relation between mass and light profile axis ratio q and pointing angle ϕ. In S22, this relation is based on 63 of the gravitationally lensed system of the SLACS lenses (Bolton et al., 2006, 2008; Auger et al., 2010), and is represented in their Figure 4 and Figure 5, here reported in Figure 4.12 and Figure 4.13, respectively.

Considering that this analysis treats a single lens, from which luminous axis ratio and pointing angle have been measured, the reported relation can be translated into a top-hat likelihood and a Gaussian likelihood, respectively. However, to avoid overconstraining the model, I considered a weaker addition to the likelihood. For what concerns the axis ratio, I instead implemented a hybrid likelihood distribution, considering a uniform distribution for q$\leq \langle \mathrm{q_{LL}}\rangle - 0.1$ and a normal distribution otherwise. Such normal distribution would be centred on $\langle \mathrm{q_{LL}}\rangle - 0.1$ and have a standard deviation $\sigma_{q_{\mathrm{LL}}}$ obtained from the propagated uncertainty on $\langle q_{\mathrm{LL}}\rangle$ as reported in Table 4.4 scaled by a factor of 3.

Lastly, for the pointing angle ϕ the punishing term is a normal χ^2 function, assuming a

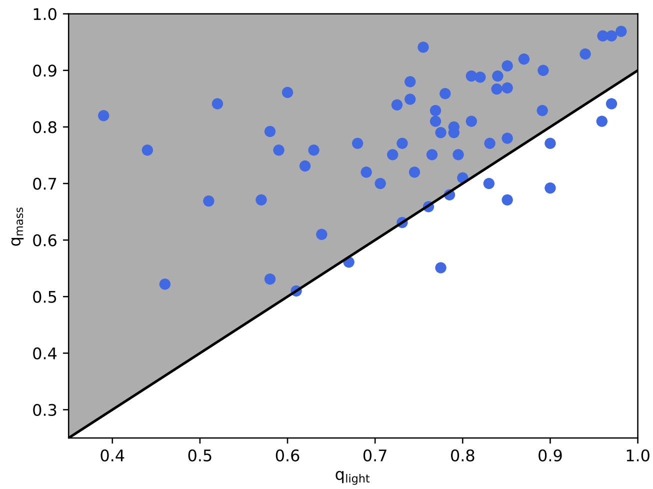

Figure 4.12: Figure 4 from S22: The relation between the axis ratio q of mass (q_{mass}) and light (q_{light}) profiles for 63 of the SLACS lenses. The shaded area corresponds to $q_{\mathrm{mass}} > q_{\mathrm{light}} - 0.1$ and is taken as additional information implemented in the likelihood of this study.

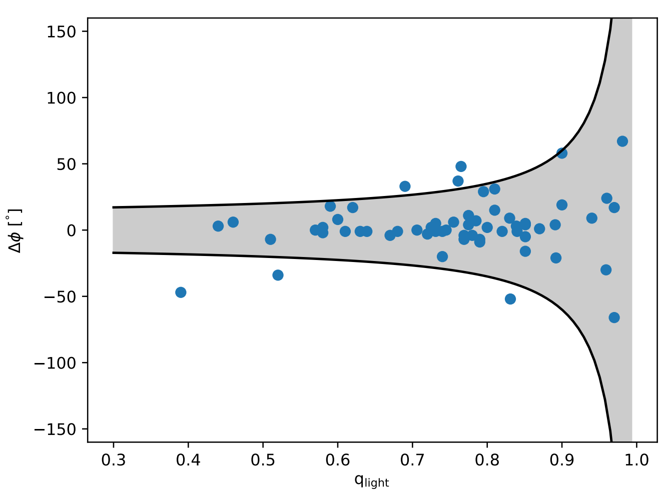

Figure 4.13: Figure 5 from S22: The relation between the difference of pointing angle between the mass and light profile, $\Delta\phi = \phi_{\mathrm{mass}} - \phi_{\mathrm{light}}$ and the axis ratio of the light profiles (q_{light}) for 63 of the SLACS lenses. The shaded area corresponds to $\Delta\phi < |10 - 5/(q_{\mathrm{light}})|$. While considered as such for the lens modelling prior in S22, in this study I will implement a simpler Gaussian likelihood relation between the pointing angle of the two profiles, as described in Section 4.3.5.

$\sigma_{\phi_{\rm LL}} = 4.5\,^\circ$. This can be summarised as follows:

$$\chi^2_d = -\frac{{d_{\rm L}}^2}{2\sigma^2_{d_{\rm L}}} \tag{4.9}$$

$$\chi^2_{\rm q} = \begin{cases} 0 & \text{if } q \leq (\langle q_{\rm LL}\rangle - 0.1) \\ -\frac{[q-(\langle q_{\rm LL}\rangle - 0.1)]^2}{2\sigma^2_{q_{\rm LL}}} & \text{else.} \end{cases} \tag{4.10}$$

$$\chi^2_\phi = -\frac{(\phi - \phi_{\rm LL})^2}{2\sigma^2_{\phi_{\rm LL}}}. \tag{4.11}$$

$$\tag{4.12}$$

I argue that the constraints are here imposed to be less stringent to account for deviation from the rule described in S22, especially regarding the hard limit they implement for the axis ratio.

To conclude the total $\chi^2_{\rm tot}$ results from the sum over all χ^2:

$$\chi^2_{\rm tot} = \chi^2_{\rm mod} + \chi^2_{\rm source} + \chi^2_{\rm d} + \chi^2_{\rm q} + \chi^2_\phi. \tag{4.13}$$

4.4 Modelling Run

Once I preprocessed the data (see Section 4.2) and defined the models (see Section 4.3), I then proceeded to run the models. `lenstronomy` is implemented with two non-linear solvers; first, a Particle Swarm Optimisation (PSO, Kennedy, Eberhart, 1995) is run to recover the approximate optima of the parameter space, then a Markov chain Monte Carlo implemented with `emcee` (Foreman-Mackey et al., 2013). The PSO, as the name suggests, is implemented by selecting a sample of points in the multi-dimensional parameter space, referred to as particles, which are given a direction and velocity. The optimisation then runs iteratively by moving the "swarm" of particles and measuring the likelihood (following equation 4.13) of the new position for each of them, called "fitness". Each particle then adapts its velocity and direction in relation to its own "personal best", i.e. the position it has reached with the highest likelihood, and the "global best" of the swarm, i.e. the highest likelihood position found by any of the particles. The particles are then expected to converge to the global optimum of the parameter space. The convergence of the PSO in `lenstronomy` is defined by two main criteria. First, the first $p = 70\%$ of particles is selected with respect to their "personal best", and the mean of their "personal best", $\langle pb \rangle$, is computed. The first criterion of convergence is that the difference between $\langle pb \rangle$ and the "global best" of the swarm must be less than $m = 10^{-3}$. If the first criterion is passed, the program selects the best $p = 70\%$ of particles - note that this considers the "fitness" of the particle at each iteration, not the "personal best" that a given particle has ever reached. It then computes the distances within the parametric space between the position of these particles and the recorded position of the "global best". If the maximum distance is smaller than $n = 10^{-2}$, the second criterion is deemed to be reached. Note that all hyperparameters p, m, n can be adjusted, but in this analysis, the default values appeared to be well-defined and therefore I did not modify them.

In this work, the PSO was run with 2000 particles and 5000 iterations. This setup, combined with the boundaries of the prior defined in Section 4.3.4 (see Table 4.6), ensured that all models reached convergence. Having well-defined convergence criteria meant that the PSO was interrupted once these were reached. Thus, implementing long iterations was not deemed an expensive task. Note that the PSO by construction does not sample effectively the parameter space, but it is a fast optimiser, i.e. reaches convergence relatively quickly.

Once completed, the PSO outputs the "global optimum" position. This is then used as a starting value to sample the posterior with the MCMC. The MCMC implemented within `emcee` with an affine invariance property (Goodman, Weare, 2010). Such a property allows the sampler's efficiency to be independent of the anisotropic shape of the sampled distribution. Given the unknown shape of our posterior, this is a valuable property. This sampler is initialised with a burn-in of 2000 steps, followed by an initial sample of 4000 steps. The "walker ratio" is set to 10, i.e. the ratio of walkers to the number of non-linear parameters, thus corresponding to 230 walkers for each run of the NIR filters and 200 for the VIS filters. The sample after the burn-in is occasionally extended with a few thousand steps to obtain a smoother distribution, but this has no significant effect on the final result.

Contrary to the PSO, the MCMC does not have a uniquely defined convergence criterion, although several have been introduced, such as the integrated autocorrelation time or the Gelman–Rubin statistic Gelman, Rubin (1992). In this study, I considered a few different approaches. Firstly, the convergence can be visually verified by plotting the "MCMC behaviour", i.e. the average position of the walkers during the run for each parameter. Any strong trend in the last part of the run would indicate a lack of convergence. An example of this can be seen in Figure 4.14, for the main lens θ_{E} obtained from the model of F160W. The discontinuity appears due to the additional MCMC run, which doubled the length of the chain. However, such discontinuity is a minor numerical effect and has no impact on the results. Nevertheless, it can be seen that the trend is approximately flat, which is the expected behaviour for a converged MCMC.

Similarly, the corner plot of the MCMC chain can be observed. These show the correlation between the modelling parameters. Distributions that showed multimodality or centred around the prior limits are deemed not converged. Note that there is an exception to this general rule, which in this analysis is represented by the position of the QSO images in the NIR filters. As discussed, these parameters are not easily recovered from the modelling of such filters alone; thus, I introduced a very informative prior obtained from the optical filters. This implied that the posterior was necessarily affected by the prior bounds, and in some cases, hit the prior limits. Thus, this phenomenon is physically motivated and is not considered an indication of a lack of convergence. Moreover, not all parameters were given the same weight, as this analysis is focused on lens mass parameters; therefore non-convergence of light parameters (most notably for the source R^{HG} and n^{HG}) was not highlighted as an issue.

These tests had the advantage of being computationally cheap and of giving a larger understanding of the results obtained. However, they were not quantitative in nature. I then considered the criterion described in Ertl et al. (2023), equation 7:

$$\Delta \log P \leq 5, \tag{4.14}$$

where $\Delta \log P$ represents the difference in log-likelihood between the median of the first and

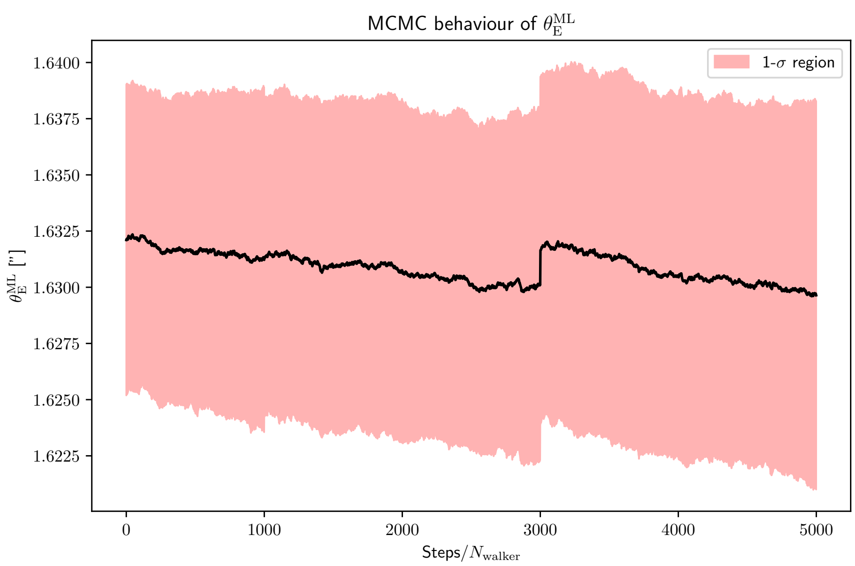

Figure 4.14: Result of the MCMC of the model F160W for the θ_{E} parameter of the main lens model. For readability, the steps shown are averaged over the walkers: the black line indicates the average and the red shaded area indicates the scatter. Note the discontinuity in the middle: this is due to the additional MCMC chain after the first 3000 steps. Such discontinuity is, however, negligible when considering the intrinsic scattering.

last 2000 points of the MCMC.

To correct for non-convergence, the model could be run with a longer MCMC or the parametric freedom could be changed. Note that in some cases, the lack of convergence might indicate some problem in the data, usually tied to an imperfect PSF model, or in the modelling choices.

If the MCMC was deemed converged, the resulting distribution of points is taken as a reasonable sample of the posterior. Given the focus of this analysis on the Fermat potential, I then computed the corresponding posterior distribution in the $\Delta\phi$ parametric space, following the same computation as described for the prior in Section 4.3.4. This resulted in the distribution $P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)$, where $\mathcal{D}_i$ indicate the i-th dataset, i.e. the set of data of a given filter i: science frame, error frame, PSF model and mask.

Once obtained $P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)$ of multiple filters, I could compare them. This step is crucial since the results must not show tension between each other. This is physically motivated, as each filter is an independent representation of the same phenomenon. Thus, the constraining power of each filter may vary, but the underlying truth should be constant. Therefore, a lack of convergence would be a serious indication of a bias or a lack of convergence in one or more of the models and would require an in-depth analysis to be solved. Most commonly, this could be explained by wrong parametric constraints, sub-optimal PSF model, non-converged PSO or MCMC and finally an under-constrained modelling approach. While the first three problems are easily tackled, the last is the most delicate and requires careful correction. In particular, this was the indication that my initial model required the additional likelihood terms discussed in Section 4.3.5 and led me to the finding of the contaminant light in F160W (see Section 4.4.1). Moreover, the posterior has to agree in order to be combined. I discuss this combination of the posteriors in Section 4.5.

4.4.1 Contaminant Light in F160W

When comparing the posterior resulting from F160W with the ones from other filters (including the optical), it became clear that there was a significant tension. This is shown in Figure 4.15, where I report the comparison on the $\boldsymbol{\Delta\phi}$ posterior for F160W and F140W. The two distribution are in tension of $\tau(\text{F140W}, \text{F160W}) \approx 2\sigma$, where the tension is computed as $\tau(\text{F140W}, \text{F160W}) = \sum_i \frac{|\Delta\phi_i^{\text{F140W}} - \Delta\phi_i^{\text{F160W}}|}{\sqrt{\sigma^2_{i,\text{F140W}} + \sigma^2_{i,\text{F160W}}}}/3$, and i indicate the image three pairs $i = [AB, AC, AD]$. $\Delta\phi_i^j$ and $\sigma_{i,j}$ are the median and standard deviation of the posterior distribution of $\Delta\phi_i$ considering the model for the j filter, respectively.

Even after testing different PSF models and increasing the MCMC run, the tension remained, thus indicating that it was independent in both aspects. It was therefore not caused by one of the most obvious points of failure and needed a more detailed study. Given the similarity between F160W and F140W, as discussed in Section 4.3.4, the same modelling approach was applied to both filters. Consequently, divergences due to modelling factors, such as the presence of the lensed host galaxy and the prior subtraction of the Main Lens light, can be ruled out. Similarly, data quality can not be a plausible explanation for divergence due to the shared pixel resolution. While F160W had a higher exposure time (see Table 4.2), this is unlikely to explain the significant tension observed. It is most likely that the tension arose from the different wavelength coverage of the filters and the corresponding observed features in the luminous components.

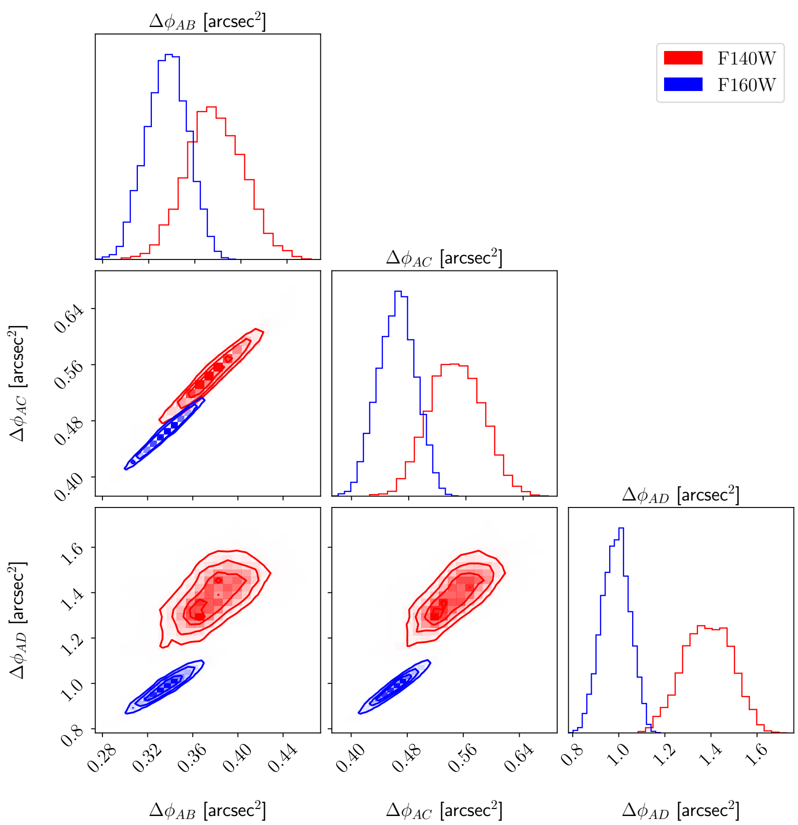

Figure 4.15: Corner plot of the posterior distributions $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F140W}})$ and $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F160W}})$. Note the significant tension between the two results obtained from the two NIR filters.

I therefore focused on the data, and I wanted in particular to understand what difference there was between the data in F140W and F160W. To do so, I considered the parametric model obtained from the lens modelling of F140W. I then reconstructed the image by convolving it with the PSF model obtained from F160W, and let the amplitude of each luminous component of the model free to vary (i.e. the perturber, the lensed host galaxy and the four individual QSO images). I then fit the science frame of F160W with this given model. The normalised residual is shown in Figure 4.16.

In this Figure, several residuals were present. Firstly, the less surprising ones are the residuals at the image positions, which are almost always present in these filters due to the imperfect PSF model, which is not tailored to fit the QSO images, as discussed in Section 4.2.3. More interestingly, two "contaminant" structures appeared, East of image A and of image B, as indicated in the Figure by the red circles. These residuals, while not as significant in amplitude as the residuals at the positions of the QSO images, are interesting due to their location. They are roughly in the same region of the arclets; however, they are not expected from the model of F140W, as their orientation is not compatible with the expected shape of said arclets. Moreover, there is no other obvious source (e.g., main lens, perturber, or other nearby sources) in this region that is expected to produce light. These excesses of light, which I will refer to as "contaminants" hereafter, are therefore not expected from the model, and given their position, residing close to the arclets and the positions of the QSO images, they appear to heavily affect the model of F160W. Other residuals are also present, such as a similar excess of light South of image C. I will show, however, that these residuals are not as significant, and while lowering the final χ^2 (see Section 4.6.1), they do not create tension in the final model. Interestingly, such contaminants were first found in the earlier model iteration, where the data of F160W and F140W were not drizzled, and the main lens light was not subtracted a priori. This indicates that this signal is not due to an artefact of the data analysis but is produced by a physical source. This is also supported by the fact that the light is even visible from the original data: for example, looking East of image A and East of image B in Figure 4.10, although heavily blended with the lens light.

I then considered multiple hypotheses to explain the source of such contaminants. Generally, these can be separated into two categories: either it is a "noise component" or an "informative component". In the first case, the source of the contaminant has no significant relation with the lens or the source and therefore is only a source of noise for the modelling. This, for example, could be a foreground object with negligible mass or be undersubtracted lens light. Initially, when modelling the data before drizzling, the contaminant near image B was not as clearly present and the one near image A appeared circular. Thus, a plausible hypothesis was that the object was a foreground star, absent in other observations due to its relative motion with respect to the system and the *HST* telescope, combined with the different times of the exposures for the other two NIR filters, which were taken two years later (see Table 4.2). However, this hypothesis, and, in general, all hypotheses which do not account for the presence of the second contaminant near image B, are now invalidated. The second category assumes the contaminant to be an informative component, which would mean that it is part of the lens system, either by being a part of the lens or the source (i.e. the host galaxy or a second source). The most appealing scenario would be the second, as it would explain the presence of both contaminant sections (i.e. both the one East of image A and the one East of image B) and would be an addition of information to the

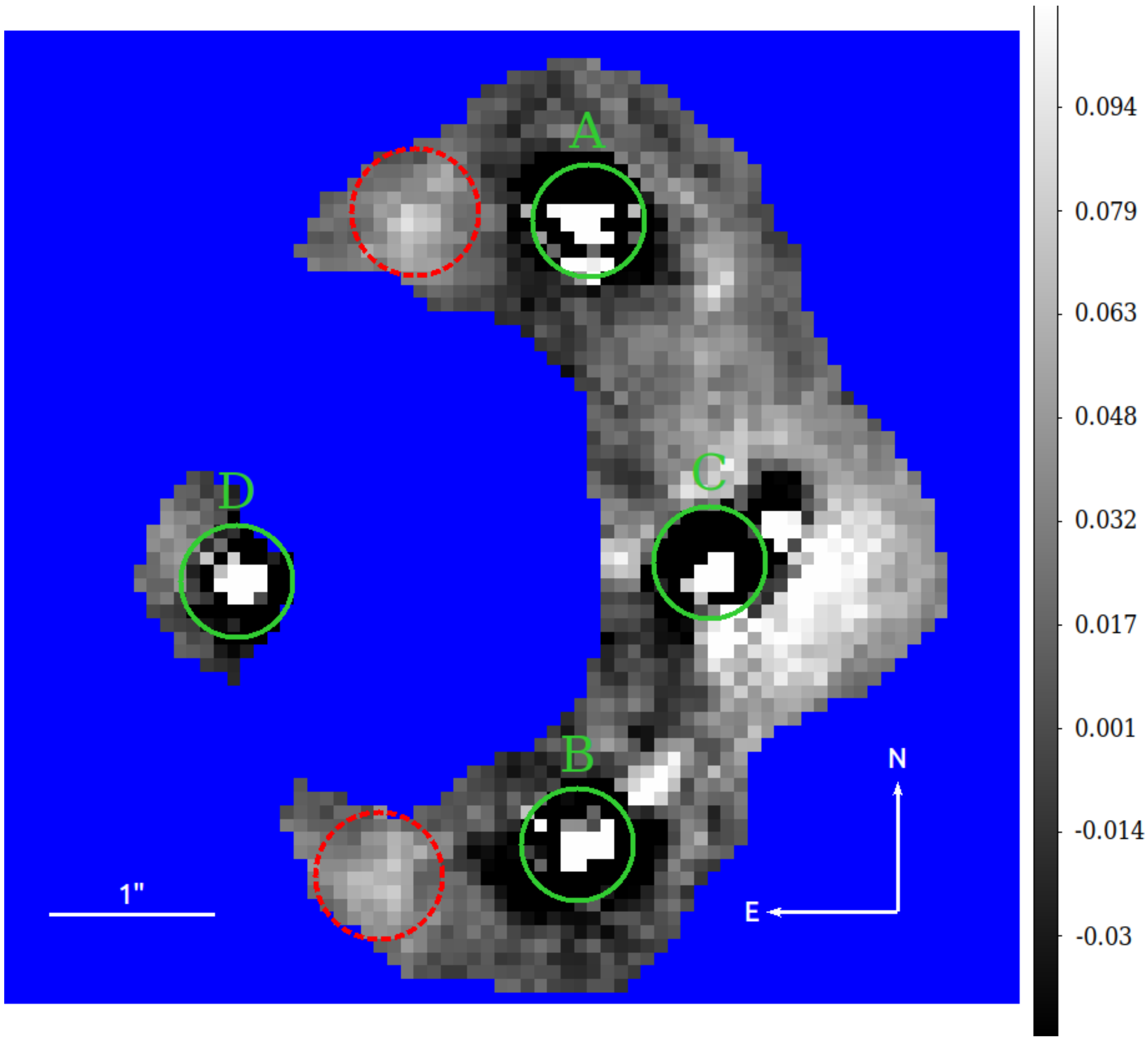

Figure 4.16: Residuals from subtracting the F160W exposure from the reconstructed image, based on the lens modelling results of filter F140W. The red circles indicate the "contaminants", while green circles and corresponding letters mark the positions of the images. Additionally, a significant residual appears west of image C; however, this does not affect the F160W model or its residuals, unlike the contaminants as later shown in Figure 4.17, and is, therefore, less significant for the model.

model. While in principle this would be useful to further constrain the model, I was unsuccessful in fitting this light. I verified that adding complexity to the modelling of the host galaxy, such as considering an elliptical Sérsic profile, a Sérsic profile + shapelets and even a second Sérsic profile, failed to improve the modelling and to fit the contaminants. While further testing could be possible by increasing the complexity of the model, it must be done with caution in order not to overfit the data. This is even more important when considering the low resolution of this filter and the unknown origin of this light. I conclude that, within the limitations of this work and of the available data, the contaminants can not be fitted and must be masked.

I then implemented an extended mask, which approximately covers the red circles shown in Figure 4.16. As described in Section 4.2.5, this had to be adapted to the pixel grid of each filter, so that the models could be compatible. The comparison between the original mask and the updated one, along with the corresponding normalised residual map for the modelling of F160W, is shown in Figure 4.17.

This completely solves the tension, as the resulting posteriors $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F140W}})$ and $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F160W}})$ are now in agreement within 0.6σ, as seen in Figure 4.18. Note that the same mask has now been applied to both models, but F140W remains mostly unchanged by the difference.

I conclude that the presence of unknown complexity in the data, likely corresponding to the extended source, biases the model if not correctly taken into account. The masking of the contaminants solves the problem, although it limits the constraining power of the data.

These effects are predominantly observed in F160W due to a combination of factors. These are the higher brightness of the redshifted host galaxy at the observed wavelength, the large PSF FWHM (see Table 4.3) and the lower pixel resolution of NIR filters.

While in this analysis I was therefore limited to masking, I believe that future higher-resolution observations in wavelengths similar to F160W or further in the infrared, e.g. observed from James Webb Space Telescope Near-infrared Camera (Beichman et al., 2012), could be employed to further study the contaminants and understand their origin. This would, in turn, be useful to constrain the lens mass model further.

Nevertheless, these results are in agreement, converged and with reasonable residuals. I thus consider them to be ready to be combined in a common posterior, as described in the following Section.

4.5 Combination of Posterior

Given the independent modelling of each filter followed in this analysis, the result of the modelling are two independent posteriors, $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F140W}})$ and $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F160W}})$, shown in Figure 4.18. I then applied a Bayesian approach to combine the posteriors in order to obtain the joint posterior $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F140W}}, \mathcal{D}_{\mathrm{F160W}})$. The problem can be generalised to any number of posterior $P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)$ given the dataset $\mathcal{D}_i$, as long as the datasets are independent. In this case, the NIR datasets are independent of each other. However, this can not be said of the opticals, given that I used the positions of the QSO images measured in the latter to constrain the prior of the modelling. Therefore, this approach is only valid to combine the posterior of the NIR filters, while I will

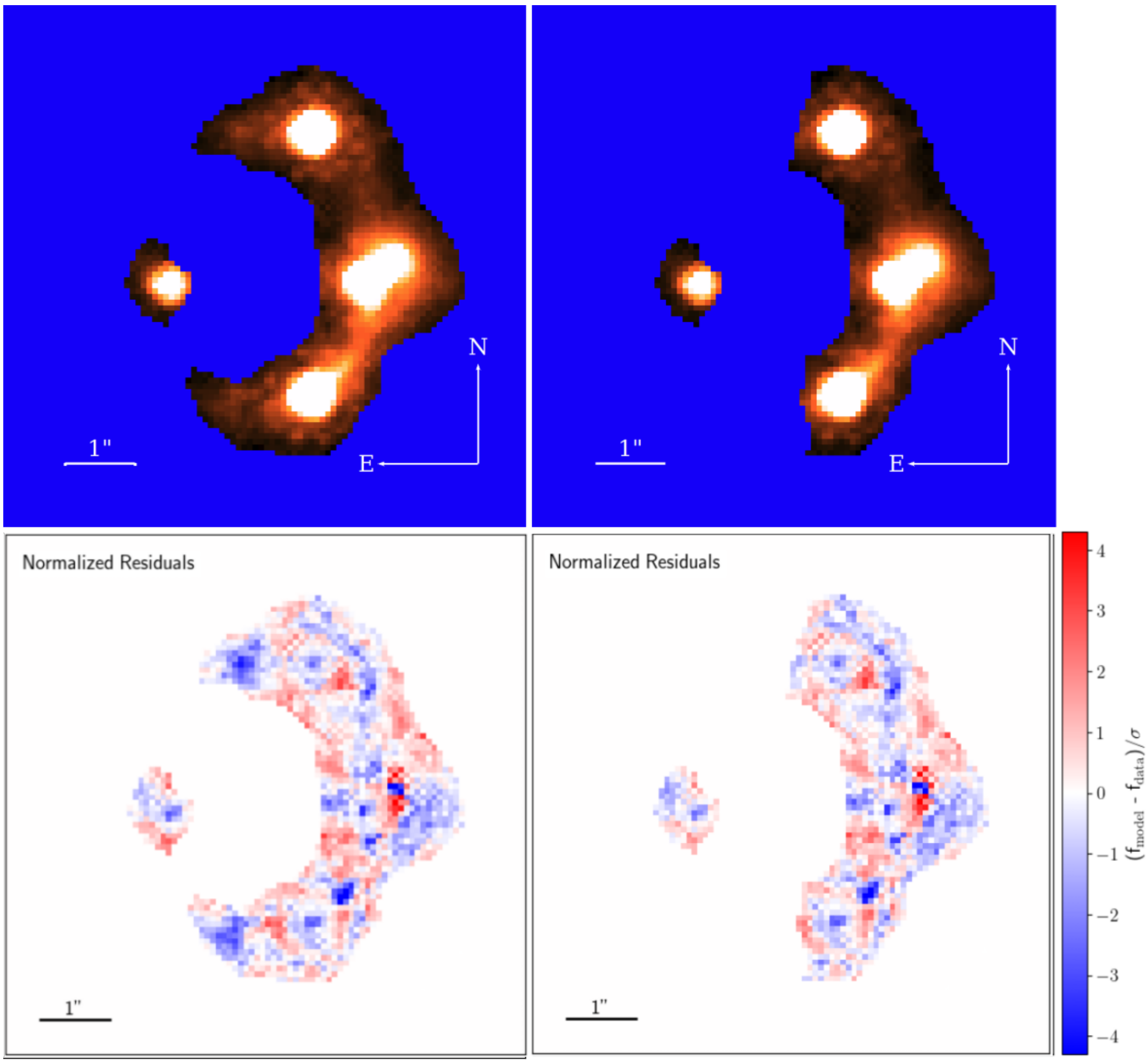

Figure 4.17: Comparison of the original mask (top left) and the updated mask (top right) applied to F160W and the corresponding residual after modelling (bottom row). Notice the clear residual East of image A and image B that indicate the presence of the contaminants.

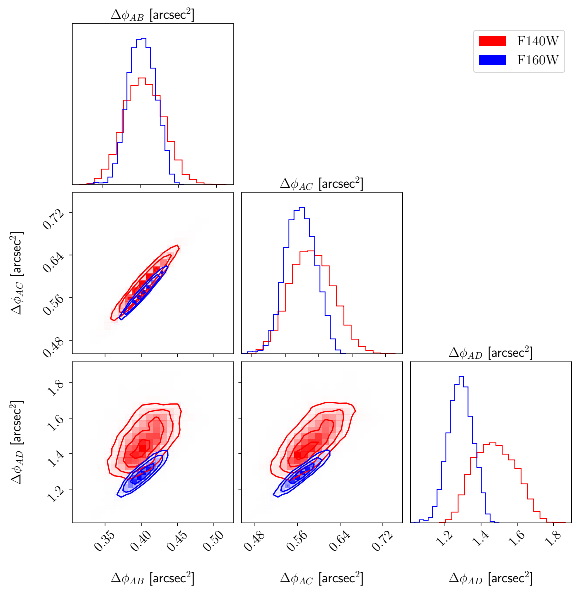

Figure 4.18: A corner plot representing the posterior of the Fermat potential obtained from the model of F140W and F160W after the masking of the contaminant, shown in Figure 4.17. Note that the tension has been solved.

further discuss the combination of the optical posterior in Section 4.5.1. Given the Bayesian theorem, we can write:

$$P(\boldsymbol{\Delta\phi}|\mathcal{D}_i) \simeq P(\mathcal{D}_i|\boldsymbol{\Delta\phi}) \cdot P(\boldsymbol{\Delta\phi}) \tag{4.15}$$

$$P(\boldsymbol{\Delta\phi}|\boldsymbol{D}) \simeq \prod_i^N P(\mathcal{D}_i|\boldsymbol{\Delta\phi}) \cdot P(\boldsymbol{\Delta\phi}) \tag{4.16}$$

$$P(\boldsymbol{\Delta\phi}|\boldsymbol{D}) = a\frac{\prod_i^N P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)}{P(\boldsymbol{\Delta\phi})^{N-1}}. \tag{4.17}$$

Where $\boldsymbol{D} = [\mathcal{D}_1, \mathcal{D}_2, ...]$, $P(\mathcal{D}_i|\boldsymbol{\Delta\phi})$ is the likelihood and a is a normalisation constant which takes into consideration the evidence. While equation 4.17 can be applied more generally for any number of independent datasets, in this analysis $N = 2$ and $P(\boldsymbol{\Delta\phi}|\boldsymbol{D}) = a\frac{P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F140W}})P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F160W}})}{P(\boldsymbol{\Delta\phi})}$.

In practice, this computation can be divided into two parts, computing the density of the posterior distributions $P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)$ and of the prior $P(\boldsymbol{\Delta\phi})$. The posterior is computed by measuring the distribution density by binning the data into a histogram. In order to ensure that the binning is identical for each posterior, I first compute the range of each posterior. I then define a uniform binning within the maxima and minima of this range. Following this, I measure the density of points of the MCMC distribution for each model within each bin, obtaining the density for each posterior $P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)$. Subsequently, I can multiply for each bin j the densities, obtaining $P(\boldsymbol{\Delta\phi}_j|\boldsymbol{D}) = \prod_i P(\boldsymbol{\Delta\phi}_j|\mathcal{D}_i)$. The prior $P(\boldsymbol{\Delta\phi})$ is computed the same way as in Section 4.3.4, by sampling the prior of the lens parameters and computing the corresponding distribution in the Fermat parameter space. I then use the kernel density estimator (KDE) available within `sklearn` (Pedregosa et al., 2011) to fit the sampled points of the prior at the position of the bin grid previously described. The KDE implementation to compute the density distribution of the prior $P(\boldsymbol{\Delta\phi})$ is needed to solve a sampling problem. The prior range is significantly larger than the range of the posterior by definition and in practice as seen by comparing Figure 4.11 and Figure 4.20. Thus, sampling the whole range of the prior might result in only a few points within the range of the posterior. This would result in a badly constrained density if it were to be measured by considering a histogram with the same binning of the posterior. Instead, a KDE can be trained over the whole prior range and then measured at the centre of the bins. This allows for a fast and safe computation of the prior. Note that the KDE could be a method to compute the density for the posterior $P(\boldsymbol{\Delta\phi}|\mathcal{D}_i)$. However, this is rendered impractical due to the large number of samples and the dimensionality of the problem, which significantly increases the computational cost for the KDE fitting. Instead, the histogram results to be significantly faster and easier to implement. Future improvement might consider another approach to the problem, but I deem this solution to be satisfactory within the required precision. Once obtained the posterior multiplication $P(\boldsymbol{\Delta\phi}_j|\boldsymbol{D})$ and the KDE trained on the prior $P(\boldsymbol{\Delta\phi})$, I can compute the combined posterior following the equation 4.17. The factor a is taken into account by normalising the posterior distribution. The result is shown in Figure 4.19

I report the resulting numerical values in Table 4.10, where I compare it to the results reported by Schmidt et al. (2022). The values reported are the median of the distribution. The upper and lower uncertainties are determined by the differences between the median and the 84th/16th percentiles, respectively. This distribution reported in Figure 4.19 is then considered as the final posterior $P(\boldsymbol{\Delta\phi}|\boldsymbol{D_{HST}})$ which will be use in Chapter 6 to constrain H_0.

4.5.1 Joining the Posterior with the Optical Models

While combining the posterior $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F140W}}, \mathcal{D}_{\mathrm{F160W}})$, I first wanted to verify that the modelling of the optical filter would be in agreement with the result obtained from the NIR. To do so, I run the modelling for F475X and F814W. These models were identical but differed from the NIR model for two reasons. Firstly, the host galaxy of the QSO is not visible as lensed arclets and therefore is not modelled. Moreover, the combined effect of their higher resolution, lower brightness of the main lens and lack of arclets meant that after subtracting the isophotal lens light (see Section 4.2.4), I could use the QSO images as a reference for the PSF model, as described in Section 4.2.3. All the other specifics of the model were in common with the NIR modelling.
The resulting posterior $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F475X}})$ and $P(\boldsymbol{\Delta\phi}|\mathcal{D}_{\mathrm{F814W}})$ are shown in Figure 4.20, where they are shown in comparison with the NIR posterior previously obtained. As expected, the results are in optimal agreement, although the optical posteriors show a larger uncertainty. This is expected due to the lower information contained in the images, due to the absence of visible arcs, which greatly increases the constraining power of the modelling. Also, note that the difference in SNR of the host galaxy arclets between F140W and F160W is reflected in the scatter of the posterior. F140W has a lower exposure and covers a lower bandwidth, resulting in a lower SNR of the red arclets, and thus in lower constraints for the mass model.

It is also interesting, while not properly mathematically correct, to consider the combination of

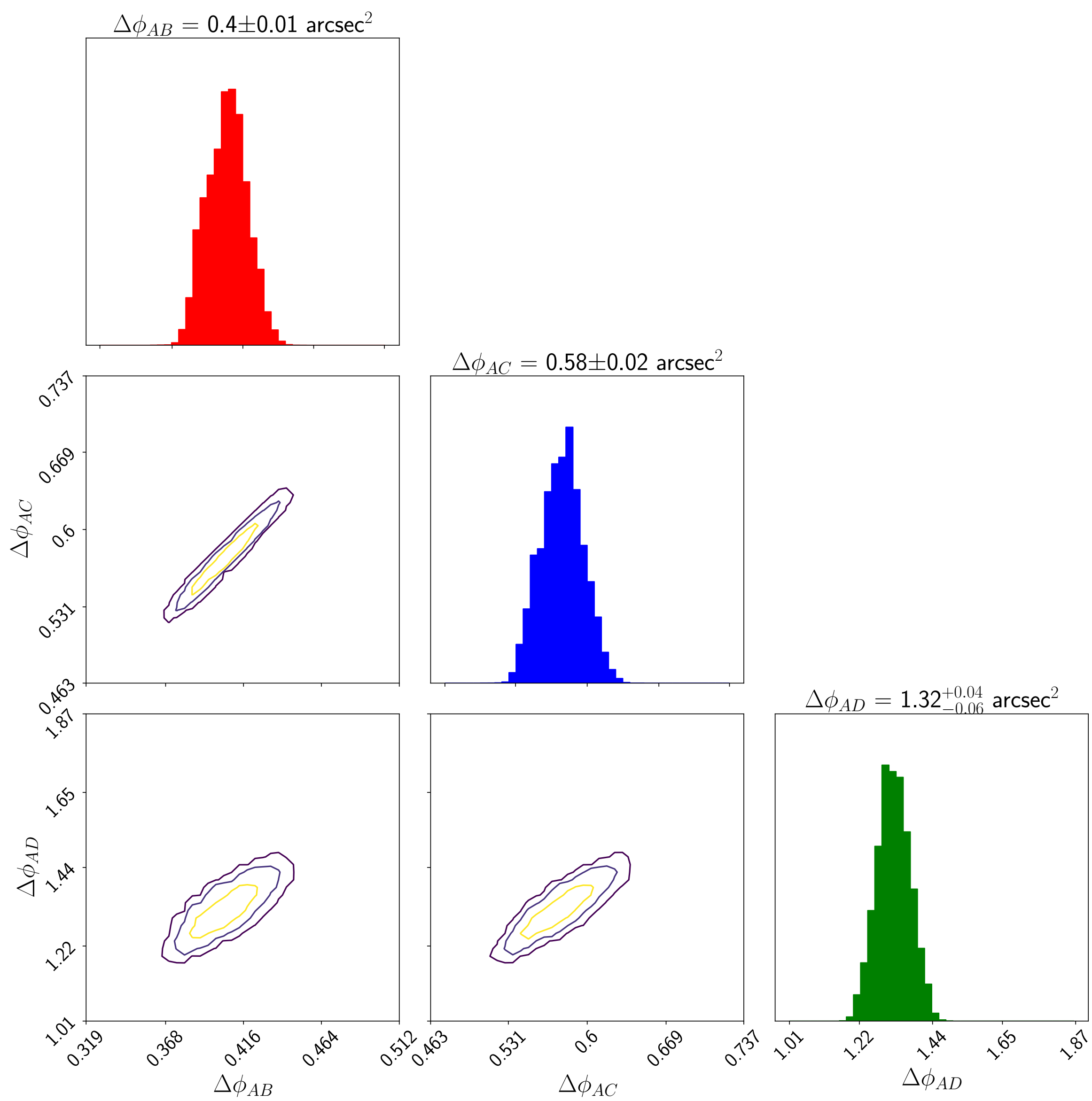

Figure 4.19: Resulting combined posterior $P(\boldsymbol{\Delta\phi}_j|\mathcal{D}_{\mathrm{F140W}}, \mathcal{D}_{\mathrm{F160W}})$ from the modelling of Section 4.4 following equation 4.17. The contour levels indicate the 68, 95 and 99.7 % confidence levels, whereas the reported values indicate the median and the 1-σ confidence level.

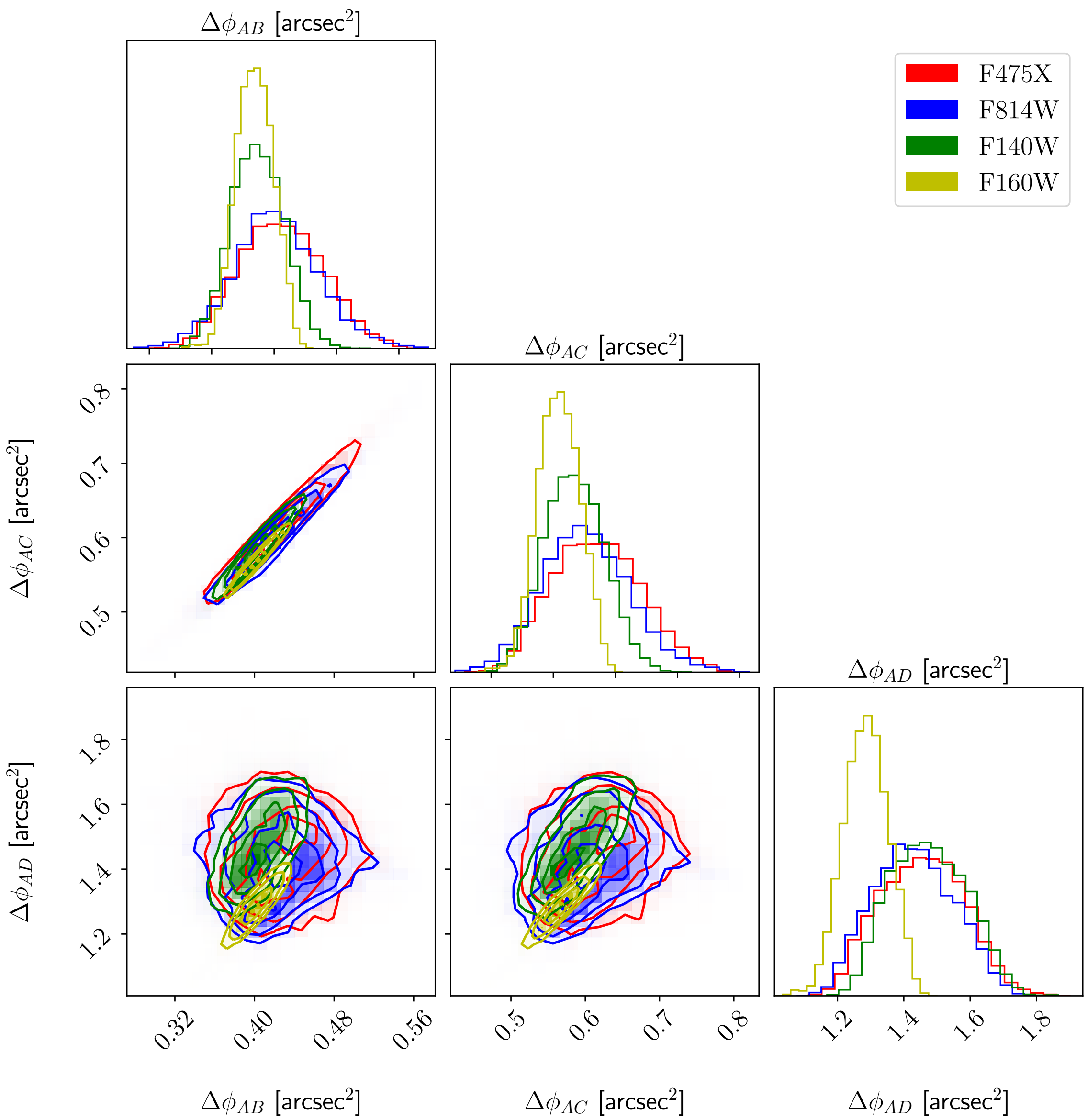

Figure 4.20: Corner plot comparing the $P(\boldsymbol{\Delta\phi}|D_i)$, for D_i = {F475X, F814W, F140W, F160W}. Note the agreement between the posteriors.

these posteriors as obtained from equation 4.17. In practice, this would be the correct combination if we ignore the dependency between the NIR modelling and the optical filter, represented by the information on the positions of the QSO image and of the perturber. The result is shown in Figure 4.21.

It can be seen that the obtained posterior is fully compatible with the one shown in Figure 4.19, thus indicating that the modelling of the optical filters does not further constrain $\boldsymbol{\Delta\phi}$. I consider this further proof that the optical filters do not contain other useful information other than the QSO and perturber position.

4.6 Results

Once I obtained the posteriors, I estimated the individual performance in Section 4.6.1. I present here some ulterior results of the mass and light model, such as the total mass and mass-to-light ratio (in Section 4.6.2) and the main lens colour profile (in Section 4.6.3). Finally, I compare the obtained model with the available literature; while few other models are available, such as Agnello et al. (2018) and Shajib et al. (2019), I will focus on Schmidt et al. (2022), which present the most updated model on J1433, and share the largest similarities in the modelling procedures.

4.6.1 Goodness of Fit

An intuitive way to define the goodness of a given fit is given by the normalised residual map. This is obtained from the formula $\mathrm{p_i} = \frac{(\mathrm{Model_i}-\mathrm{Data_i})\mathrm{Mask_i}}{\sigma_i}$, where p is the normalised residual map, Model is the final model, Data is the observed image in a given filter and σ is its corresponding error frame. All such values are computed pixel-wise for each non-masked pixel i. The obtained maps are shown in Figure 4.22, where I included the optical results for completeness. Notice that the latter has a significantly lower residual, which is explained by the lack of extended sources of light and the higher pixel resolution. As it can be seen, F160W present higher residuals overall, due to the lower resolution and higher brightness of the multiple luminous components.

While indicative of the presence of structure in the residual, the residual maps are not a quantitative indication of the goodness of fit. For this, we relied on the total reduced χ^2 of the non-masked pixels, i.e. the normalised residual divided by the respective number of degrees of freedom (DoF, reported in Table 4.5 for each model), also indicated as χ^2_{red}. These vary significantly between filters as they mostly depend on the number of pixels considered. The number of free parameters also changes between filters (as different light profiles are considered for different filters, see Section 4.3.2), but such a change is negligible compared to the total number of pixels. The χ^2_{red} is reported for each filter in Table 4.8. For the optical filter, $\chi^2_{\mathrm{red}} \sim 0.7$ indicates a slight overfit of the data, most likely due to the significant prior constraint on the QSO position. Do consider that these models are reported here for completeness, but do not enter the computation of the final posterior, as described in Section 4.5.

F140W also results in a low chi-square, with $\chi^2_{\mathrm{red}} \sim 0.5$. This would indicate a significant overfit of the data; however, it is likely to be driven by an overestimation of the error frame and possibly by the lower number of pixels taken into account, while only minor residuals are present.

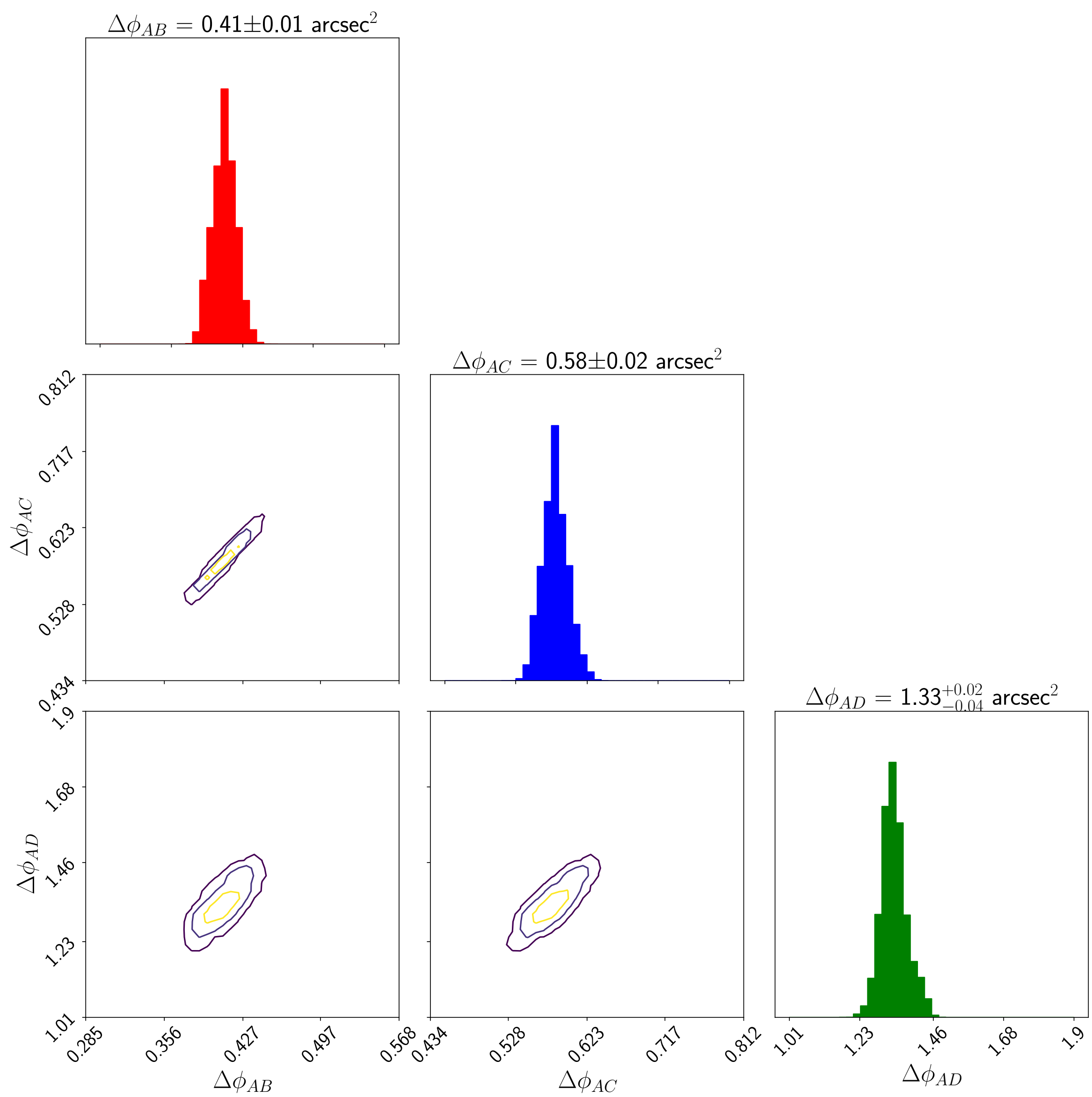

Figure 4.21: The resulting combined posterior following equation 4.17 considering the posterior of both NIR and optical filters as shown in Figure 4.20. Notice that the comparison between this image and Figure 4.19: the two posteriors are fully compatible.

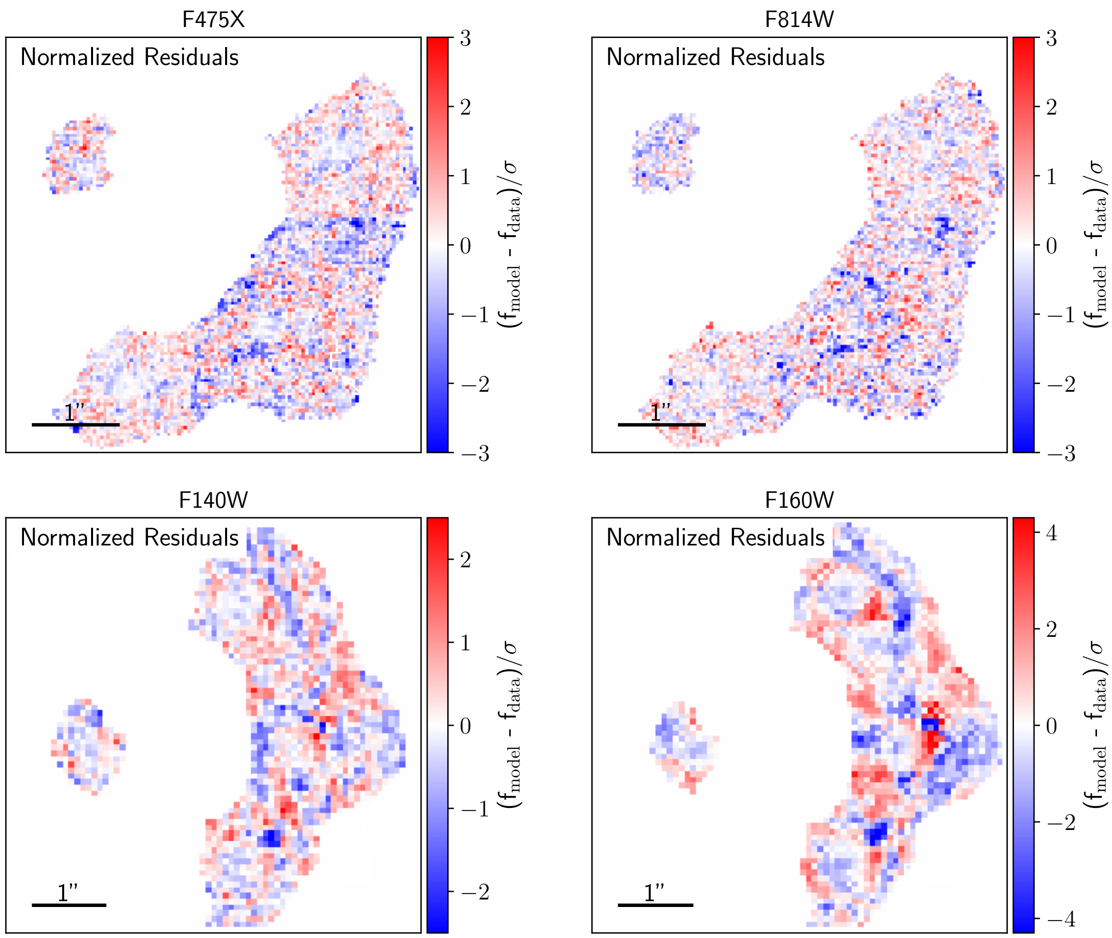

Figure 4.22: Residual maps for all observed filters, normalised by the uncertainty. The masks are also applied. Note the different orientations of the image between optical and NIR, resulting from the different drizzling.

Filter	QSO images RMSE ["]	Reduced χ^2
F475X	0.0018	0.65
F814W	0.0025	0.67
F140W	0.0075	0.51
F160W	0.0091	1.37

Table 4.8: Root Means Square Error of the QSO's images positions and total reduced χ^2_{red}.

On the other hand, the moderately high resulting $\chi^2_{\text{red}} \sim 1.4$ for F160W can be explained by the aforementioned presence of significant residual structure in the image along the arclets.

In Table 4.8, I also indicate the "Root Mean Square Error" (RMSE) of the QSO images position. This indicates the average difference between the modelled and expected positions of the QSO images. It is defined as $\text{RMSE}_{\text{QSO}} = \sqrt{\Sigma_i^N (x_{\text{obs,i}} - x_{\text{mod,i}})^2 + (y_{\text{obs,i}} - y_{\text{mod,i}})^2 / N}$ where $x_{\text{obs,i}}, y_{\text{obs,i}}$ are the reference coordinates of the QSO images in the optical filters (see Section 4.3.4), while $x_{\text{mod,i}}, y_{\text{mod,i}}$ are the ones resulting from the lens model and $N = 4$ is the number of images. Note that this is a measure in arcseconds. Unsurprisingly, the recovered positions are less accurate in the NIR models, even in the presence of the constraining prior informed by the optical filters.

4.6.2 Total Mass and Mass-to-Light Ratio

To obtain the Mass-to-Light ratio, I will combine the measurement of the mass profile from the most accurate mass model and the light profile from the most accurate lens light modelling. The first is obtained from the modelling of F160W, due to the presence of the lensed host galaxy as described in the previous sections. The most accurate lens light model, which has been obtained in Section 4.2.4, is deemed to be the one relative to F814W. This takes advantage of the higher resolution and high exposure time of this filter. Moreover, this filter covers higher wavelengths compared to F475X, where the lens galaxy appears brighter, thus increasing the SNR. Finally, the lens is not blending with the QSO images, thus allowing for an easier subtraction/masking of their light.

For each mass model, the cumulative mass profile of the lens system, defined as the total projected mass within a given aperture) can be estimated, following the standard lens equations:

$$\text{M}(\vec{r} \leq) = \Sigma_{\text{crit}} \int \kappa(\vec{\text{x}}')dx'. \tag{4.18}$$

Here, Σ_{crit} is the critical mass density, a cosmological function dependent on the redshifts of the lens and the source: $\Sigma_{\text{crit}} = \frac{\text{D}_\text{s} c^2}{4\pi \text{G} \text{D}_\text{d} \text{D}_\text{ds}}$. $\text{D}_{\text{s/d/ds}}$ indicate the angular diameter distances between us and the source, us and the deflector and the source and the deflector, respectively. G and c are

the gravitational constant and the speed of light, respectively. $\kappa(\vec{x})$ is the convergence at a given position $\vec{x}$, which is defined as the projected mass density $\Sigma(\vec{x})$ normalised by $\Sigma_{\rm crit}$. The integral is then a two-dimensional integral over the circular region of radius $\vec{r}$.
Note, however, that the density profile used to model the mass distribution of the lens system (described in Section 4.3.1, equation 4.6) diverges at $\vec{x} = 0$ and does not converge to zero fast enough for $\vec{x} \to \infty$, meaning that the total integrated mass at infinity diverges as well. Both cases are nonphysical but are not critical as the lens model is considered accurate only within the regions where the images appear, indicatively at the Einstein radius $\theta_{\rm E}$. Very far from such a region (both at very small radii or very large), the model is merely an extrapolation and should not be trusted. Such divergences are avoided in practice by defining cut-off radii. To define these cut-off radii, I also take into consideration the light profile obtained from the isophote models (see Section 4.2.4), in order to allow for a meaningful comparison between the mass and light profile and hence the Mass-to-Light ratio profile. Therefore, I restrict the focus on the mass profile results within the range of [$a_{\rm min} \lesssim a \lesssim a_{\rm max}$ ”], where a is the major axis of the main lens light profile. I defined the minimum radius to be the major axis of the first isophote larger than the $\rm R_{EE}$, the EE radius defined in Section 4.2.3, in order to avoid PSF effects. For F814W, $\rm R_{EE}$(F814W) = 0.075″ (as seen in Table 4.3), corresponding to $a_{\rm min}$ = 0.088”, equivalent to 2.2 pixels for F814W, and 0.48 kpc on the lens plane. For the outer cut-off, I limited the plot to $a \leq 5\theta_{\rm E} \approx 7.8$ ”, i.e. $a_{\rm max} \approx 42.4$ kpc on the lens plane.
I compute the flux of the main lens by considering elliptical apertures using the Python package `photutils` (Bradley et al., 2022) matching the isophotes of the lens model. I then compute the enclosed luminosity from the model, as well as the corresponding error from the error frames. The flux is then converted into magnitude considering the equation (Institute, 01/31/2024):

$$m_{\rm AB}(r) = -2.5\log_{10}\left(\frac{{\rm Flux}(r)}{{\rm EE}(r)}\right) + {\rm ZP_{AB}}, \tag{4.19}$$

where the subscript AB indicate that the magnitude is in the AB magnitude system. while r refers to the radius and Flux(r) is the flux enclosed within it. Note that I correct for the Encircled Energy (EE), which is reported from the literature, as discussed in Section 4.2.3. The zeropoint $\rm ZP_{AB}$ is obtained following the equation described in STScI (05/30/2024):

$$\rm ZP_{AB} = -2.5\log_{10}(\texttt{PHOTFLAM}) - 5\log_{10}(\texttt{PHOTPLAM}) - 2.408. \tag{4.20}$$

`PHOTFLAM` and `PHOTPLAM` are header keywords used to derive the instrumental zeropoint; specifically, `PHOTFLAM` is, quoting STScI (05/30/2024), “the inverse sensitivity (units: erg cm^{-2} Å^{-1} electron^{-1}). This represents the scaling factor necessary to transform an instrumental flux in units of electrons per second to a physical flux density”, while `PHOTPLAM` is the pivot wavelength, in units of $\mathring{A}$. To convert the apparent magnitude into the absolute magnitude, I correct for the distance modulus of $5 \cdot \log_{10}(D_{\rm L}(z_{\rm lens})/10{\rm pc}) = 41.73$ mag (where $D_{\rm L}$ is the luminosity distance), for cosmic dimming of $2.5 \cdot \log_{10}((1 + z_{\rm lens})^4) = 1.48$ mag, galactic extinction and K correction. The last two corrections are wavelength-dependent and are obtained from the literature. I obtain K(F814W) = 0.55 mag using the tool available at the site `http://kcor.sai.msu.ru/` based on Chilingarian et al. (2010) and Chilingarian, Zolotukhin (2012). The galactic extinction is retrieved from NASA/IPAC (2023), which results to be $A_{\text{F814W}} = 0.014$ mag. Finally, we convert to solar magnitude by considering the sun magnitude M= 4.52 mag in AB magnitude for F814W, as reported by Willmer (2018).

The resulting Mass-to-Light ratio $\Upsilon(a)$ is shown in plotted in Figure 4.23. Note how the trend of the enclosed mass is growing linearly, whereas the luminosity flattens, thus yielding a positive trend for $\Upsilon(a)$. It can be seen that the $\Upsilon(\theta_{\text{E}}) \approx 1$, which is in line with the expected Mass-to-Light ratio for such an early-type galaxy, at least considering the inner region sampled. Note that we are here taking into account only the light of F814W. Moreover, the corresponding mass $M(\theta_{\text{E}})$ results to be $4.28^{+0.04}_{-0.05} \cdot 10^{11}$. Doing some significant approximation, considering the model as a singular isothermal sphere, it is possible to extrapolate an analytical estimate of the velocity dispersion σ_{v} using the equation $\sigma_{\text{v}} = c\sqrt{\theta_{\text{E}} \frac{D_{\text{s}}}{4\pi D_{\text{ds}}}} \approx 275\,\frac{\text{km}}{\text{s}}$ (refer e.g. Meylan et al., 2006, , eq. 52). Considering that this is a rough approximation, this result is well in agreement with the result obtained from spectroscopy of Mozumdar et al. (2023), who reports in Table 3 $\sigma_{\text{v}} = 261 \pm 6(\text{stat.}) \pm 7(\text{sys.})\,\text{km/s}$.

4.6.3 Main Lens Colour

Given the isophotal model obtained in Section 4.2.4, it is then possible to obtain the magnitude of the main lens and its colours. I will focus here in particular on the integrated colour profile obtained from F475X and F814W. I measure the luminosity of the main lens as described in the previous Section 4.6.2, by measuring the flux within elliptical apertures and converting it first into apparent magnitude (following equation 4.19), then into absolute magnitude. The last conversion is divided between geometric corrections (distance modulus and cosmic dimming), which are independent of the wavelength, and wavelength-dependent corrections, such as galactic absorption and K-correction. The latter are computed in the same way for F475X as for F814W, obtaining $K(\text{F475X}) = 0.25$ and $A_{\text{F475X}} = 0.029$ mag.
I then compute the enclosed luminosity difference in the two filters for a given major axis value a. Note that this profile represents the integrated colour, i.e. the colour given by the enclosed luminosity within the isophote, rather than the colour at a specific isophote. The result is shown in Figure 4.24. The plot is cut after $a > 3$ " from the centre of the lens, as it reaches the limiting surface brightness.

The grey area, i.e. $a < 0.05$ " is ignored as it represents where the differences in the PSFs most strongly affect the model. The negative trend with radius, i.e. the colour becoming "bluer", is expected from elliptical galaxies (Saglia et al., 2000).

The resulting integrated colour within the isophote with major axis 0.05 " $< a <$3 " is $\Delta\text{mag}_{\text{F475X}-\text{F814W}} = 2.17 \pm 0.07$ mag.

Given the scope of this analysis, I do not consider a more precise estimation of the error. Instead, it is approximately obtained from photon noise and thus ignores any uncertainty due to other isophotes parameters (centre coordinates, ellipticity and boxiness). However, these uncertainties should be confined to the area we are considering here. Thus, while the reported error may underestimate the full uncertainty, it can still be regarded as a reasonable approximation.

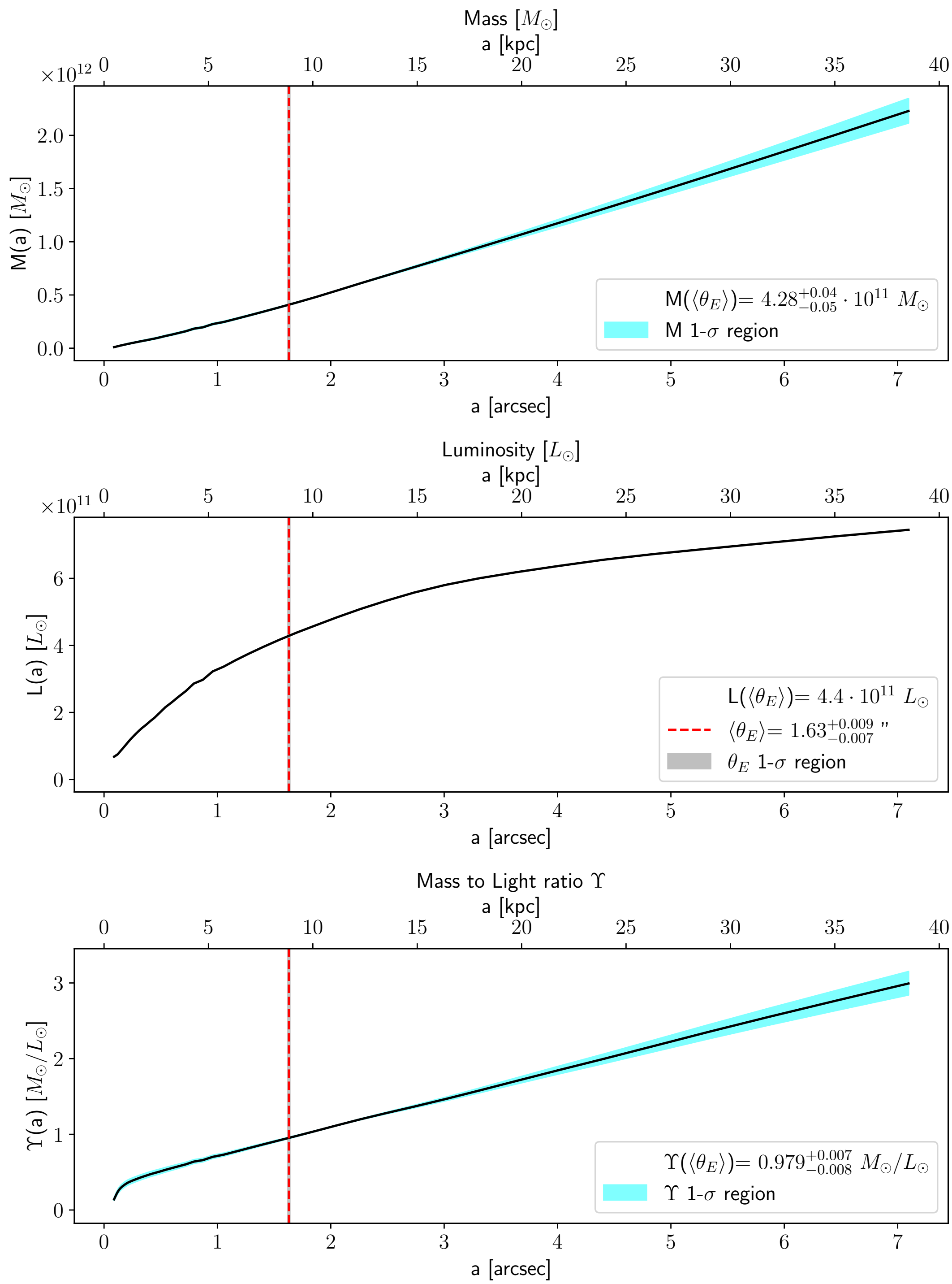

Figure 4.23: Plot of the cumulative mass from lens mass modelling (for filter F160W, see Section 4.4.1), the enclosed luminosity from isophotal fitting (for filter F814W, see Section 4.2.4), and the corresponding Mass-to-Light ratio Υ. The plots are with respect to the semi-major axis of the isophotes, a. The 1-σ region computed from the posterior of the lens modelling is indicated by the shaded cyan band. The luminosity uncertainty, while computed, has a negligible effect on the error budget and is not plotted. $\theta_{\rm E}$, obtained from the mass model, is indicated by the red dashed line, along with its corresponding grey 1σ region.

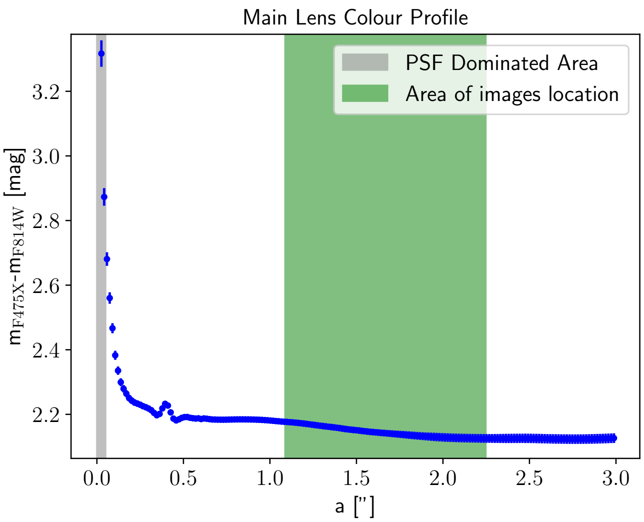

Figure 4.24: The integrated colour profile obtained from isophotal light models of the main lens in filters F475X and F814W, plotted with respect to the isophotes' major axis a. The green vertical region indicates the location of the QSO images, while the grey region approximately corresponds to the FWHM of the PSF of F475X (~ 0.05"), highlighting where the modelling is most affected by the PSF.

4.6.4 Comparative Analysis with Existing Literature

I compare the results of the lens modelling with the available literature, specifically Schmidt et al. (2022) (hereafter S22). I disregard Agnello et al. (2018) due to the lower resolution of the data used in their preliminary analysis, while I ignore Shajib et al. (2019) since S22 present the results of an improved modelling procedure based on this paper.
S22 shared some similarities and differences with the current work. They used the same modelling program (`lenstronomy`) and the main lens mass profile (an elliptical power law). While not explicitly described in the article, following their pipeline shown in Figure 3 and the resulting convergence map in Figure C5, here reported in Figure 4.25, it appears that the perturber was also modelled considering a singular isothermal sphere. Furthermore, an external shear was considered. However, the analysis developed in S22 diverged from the current work in other aspects; firstly, S22 aimed to introduce an automated pipeline for lens modelling, while I focus here on the analysis of a single system. In S22, they considered multifilter modelling, based on the `HST` exposures from filters F475X, F814W and F160W. I instead implemented a single filter modelling, and considered F140W and F160W for the modelling, while employing the optical filter only to constrain the position of the QSO and the perturber.

In Table 4.9, the results between this model and the results of S22 are compared, along with their tension. Note that I will report only the results relative to the modelling of F160W, as it is the most precise result obtained. Only a limited number of parameters are in agreement, while most results present varying degrees of tension. Partially this is due to the very high precision claimed by S22, such as for θ_E and possibly for y^{ML}, while others stress a significant difference

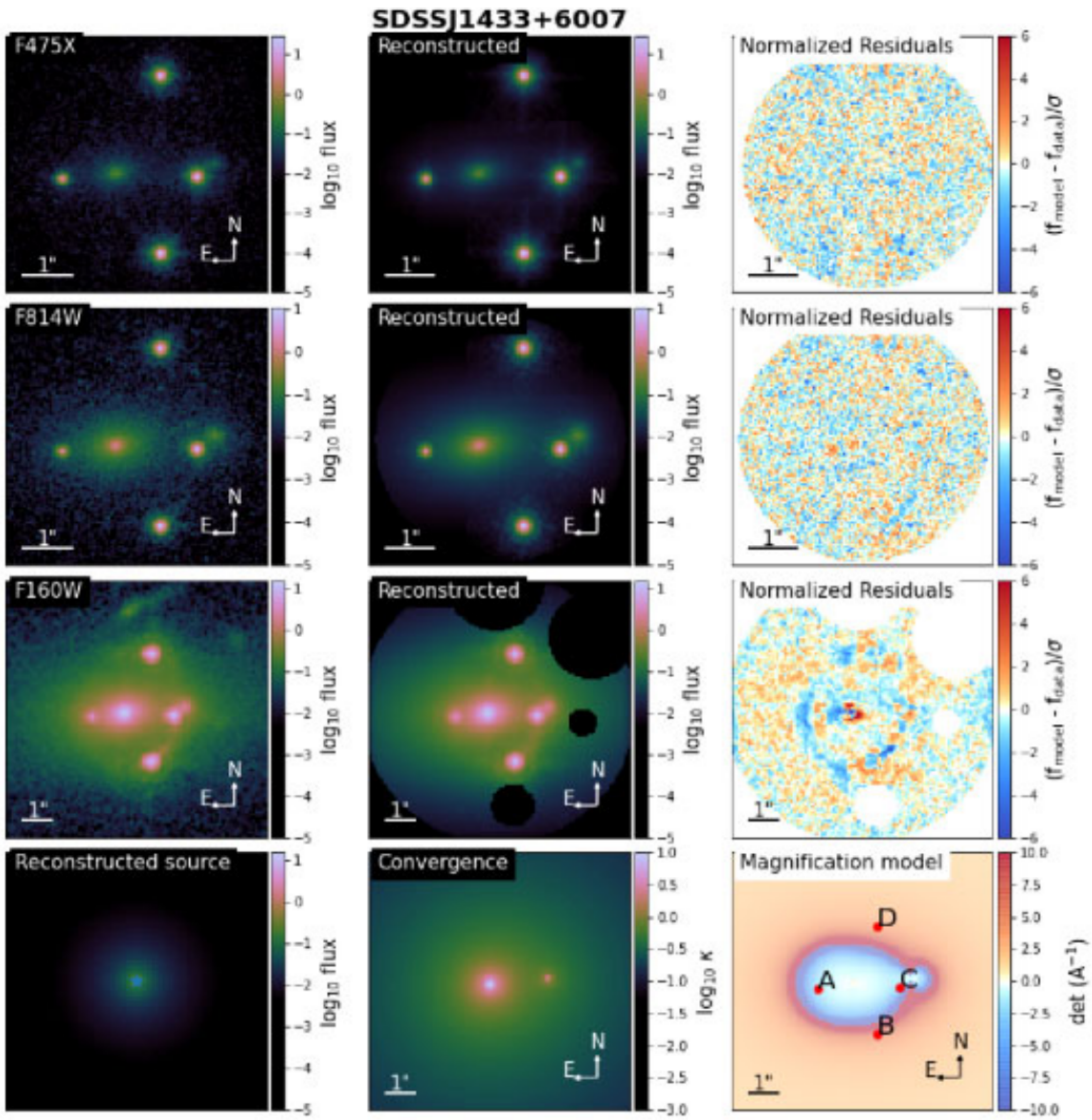

Figure 4.25: Resulting model reported in Figure C5 of S22. Note the perturber modelled as a SIS, the host galaxy modelled as a circular Sérsic profile and the significant residual in F160W at the position of the contaminant, described in Section 4.4.1.

		This paper (F160W)	S22	Tension
θ_E^{ML}	["]	$1.63^{+0.007}_{-0.008}$	$1.581^{+0.003}_{-0.002}$	6.2
γ^{ML}	["]	$1.88^{+0.04}_{-0.05}$	1.92 ± 0.03	0.8
q^{ML}	[]	0.69 ± 0.02	0.96 ± 0.01	10.6
ϕ^{ML}	[°]	-4.8 ± 0.5	$-28^{+4.5}_{-2.6}$	6.5
x^{ML}	["]	0.93 ± 0.01	0.931 ± 0.006	0.3
y^{ML}	["]	-2.061 ± 0.002	-2.038 ± 0.006	3.6
θ_E^{Pert}	["]	0.186 ± 0.009		
γ^{Shear}	[]	$0.061^{+0.007}_{-0.008}$	0.127 ± 0.004	7.8
ψ^{Shear}	[°]	$-70.3^{+2.7}_{-3.8}$	-82 ± 0.4	3.7

Table 4.9: Comparison between the mass-profile results for F160W and S22, adapted to the same frame of reference. It should be noted that S22 does not explicitly report the value for θ_E^P.

in the analysis. For example, the high level of tension in the ellipticity parameter q^{ML} and ϕ^{ML} is not explained only by the small uncertainty of the result of S22. In this regard, the pointing angle ϕ^{ML} reported by S22 presents a difference larger than 18 ° from the lens light distribution. In this analysis, it would be unlikely to happen as such large differences are suppressed (see Section 4.3.5). This tension can be explained by the fact that the axis ratio q^{ML} resulting from S22 is very large, indicating an almost spherical distribution. Thus ϕ^{ML} is almost degenerate with respect to the model; therefore, this tension is not significant. On the other hand, the tension in q^{ML} strongly impacts the mass model and therefore the cosmological inference based on it.
This tension might be due to the modelling of F160W in S22, as it has no mask accounting for the "contaminant" described in Section 4.4.1. As seen in Figure C5, reported here in Figure 4.25, the modelling of this filter leaves significant residuals at similar positions, as well as in other regions, which are masked in this model.

Given the focus of this analysis on the Fermat potential differences, I compared $\Delta\phi$ (Table 8 of see S22). In this case, I consider the combined Fermat difference posterior shown in Section 4.5 and Figure 4.19. This was done to compare our final result with a similarly informed model from the literature. Note, however, that the result of this analysis is formally only on the two NIR filters F160W and F140W. The information about the optical is reduced to the prior of the positions of the QSO images and the perturber. Note also that the naming of the QSO images differed from this work and S22, as the names of images A and D are inverted. I remain consistent with the naming of this work.

The results are shown in Table 4.10. As expected from the tension in the individual lens

	This paper	S22	Tension
$\Delta\phi_{AB}$ [arcsec^2]	0.4±0.01	0.36 ± 0.03	1.11
$\Delta\phi_{AC}$ [arcsec^2]	0.58±0.02	0.49 ± 0.03	2.19
$\Delta\phi_{AD}$ [arcsec^2]	$1.32^{+0.04}_{-0.06}$	1.03 ± 0.03	4.83

Table 4.10: Comparison of results for the combined Fermat potential difference $\Delta\phi$ between S22 and this work.

QSO Images	F475X			F814W			F160W		
	This Paper	S22	Tension	This Paper	S22	Tension	This Paper	S22	Tension
A	20.27 ± 0.07	20.265 ± 0.001	0.06	20.2 ± 0.1	20.175 ± 0.003	0.04	20.7 ± 0.1	20.518 ± 0.0045	2.0
B	20.09 ± 0.07	20.095 ± 0.001	0.02	20.0 ± 0.1	20.048 ± 0.003	0.07	20.7 ± 0.1	20.369 ± 0.004	3.5
C	20.447 ± 0.078	20.468 ± 0.002	0.3	20.4 ± 0.1	20.365 ± 0.004	0.2	20.6 ± 0.1	20.455 ± 0.007	2.0
D	21.9 ± 0.2	$21.972^{+0.004}_{-0.005}$	0.5	21.8 ± 0.2	$21.782^{+0.008}_{-0.011}$	0.1	22.5 ± 0.2	$21.793^{+0.012}_{-0.011}$	3.1

Table 4.11: QSO images luminosities obtained from lens modelling and compared to S22 with their relative tension.

parameters shown in Table 4.9, all resulting $\Delta\phi$ are in tension, in all cases higher than 1.1 σ.

For completeness, I compare the luminosities of the QSO images in the shared filters, F475X, F814W and F160W. To measure them, I computed the total flux for each modelled QSO image (in practice, by taking an aperture of 100 arcseconds). For the uncertainty, I considered a Poissonian noise, and took the flux of the corresponding aperture on the root-squared image. Note that this error computation ignores all other sources of error, such as the modelling and the presence of other sources.

The results are reported in Table 4.11, where the S22 data are taken from Table A2. The large tension present for the F160W is dominated by the high precision claimed by S22, along with some differences in the results obtained, which are likely to be due to differences in the modelling. The photometry of image D, which is the dimmest image and the most blended with the lens light, is particularly sensitive to the modelling of the lens light. I conclude this section by comparing the source reconstruction. This is, however, not easily done due to different parametrisation of the source. While both works consider a circular Sérsic profile, S22 consider this model in common between all filters, plus a set of two-dimensional Cartesian shapelets. Thus, S22 differ from the current work as it considers the source modelling in all filters. Nevertheless, the central position of the source with respect to image A can be recovered. I then compare them in Table 4.12. As expected, given the overall results, I find a significant tension in these results as well. The difference between the two positions is ≈0.13".

Overall, the results of S22 differ significantly in lens parameters (see Table 4.9) and thus in the Fermat potential constraints (see Table 4.10). This is likely due to the differences in modelling approach and most importantly to the difference in the masking employed.

Host Galaxy Coord.	This Paper (F160W)	S22	Tension
δRA	0.46±0.02	0.493 ±0.010	1.5
δDEC	-1.83±0.02	-1.92±0.003	4.5

Table 4.12: Comparison of the centre of the host galaxy between S22 (Table 4) and this work for F160W, along with their relative tension. The positions are relative to image A.

In particular, it appears that modelling in parallel without a robust approach for the detection and treatment of unknown systematics, such as the "contaminants" of F160W, would lead to severely biased results. This approach is therefore strongly discouraged within the context of cosmological inference. It is also to be considered that S22 refers to the results from the model of J1433 as "far from cosmography grade" based on a metric of the stability of the model. Therefore, while the tension with their result is significant, it should not be a cause to doubt the findings of this study.

Chapter 5

TDC@W: Time Delay Estimate from Lightcurves Analysis

The second cardinal part of the Time Delay Cosmographic method consists of measuring the time delay between the various strongly lensed images. This can then be used to correlate the measured Fermat potential obtained from the lens mass model, as described in Chapter 4, to the cosmological parameters, as seen in Chapter 2 (in particular, see Equation 2.79).

As previously described, the time delay with respect to the unperturbed image is not observable in this case, but the relative time delay (i.e. the time delay between two images) can be measured.

In practice, a time delay measurement of a continuously variable source (such as a QSO, in contrast to "una tantum" sources such as SNae) would require observing the lightcurve of every image with high photometric and temporal accuracy. This can be done by means of an observational campaign from a ground-based telescope. Given the substantial cost of any observational campaign, several aspects have to be considered to plan and successfully carry out the observations. Firstly, the length of the observational campaign, $t_{\rm obs}$ depends on the maximum expected time delay $\Delta t_{\rm max}$ and the variability of the source. $\Delta t_{\rm max}$ would require t_{obs} to be at least twice as large as $\Delta t_{\rm max}$, preferably a few times. $\Delta t_{\rm max}$ is not known precisely without measures but can be approximately gauged a priori from the image separation, as $\Delta t(\vec{\theta}_i, \vec{\theta}_j) \propto \Delta\phi(\vec{\theta}_i, \vec{\theta}_j) \propto (\vec{\theta}_i - \vec{\theta}_j)^2$ (as seen in Equations 2.79 and 2.73), thus indicating that larger image-separation system would require longer campaigns. It can also be estimated more precisely following the TDC equation and assuming a given value of H_0; as long as this is considered to be a rough measurement to gauge $t_{\rm obs}$, this assumption is valid and would not bias the result.

The second aspect affecting $t_{\rm obs}$, the variability of the source, is instead harder to define a priori. Given the random nature of the variability, the length of the campaign can not be defined a priori by such a parameter but rather should be prolonged in case the variability is low in order to increase the overall signal and the chance of observing a larger-scale variability.

A second aspect to consider is the frequency of observations, $\nu_{\rm obs}$, or inversely its cadence, i.e. how many days between two observations. Firstly, there is a lower bound indicated by the shortest time delay to be measured, Δt_{min}, as the cadence has to be smaller than such value.

More generally, higher frequency, i.e. higher sampling, also correlates with lower time delay uncertainty. The observation strategy for the exposures, such as the number of exposures, their length in time and the filters used, as well as the dependence of such choices on the characteristic of the single night (mostly weather and brightness of the sky) should be tailored to the specifics of the telescope and of the system.
For the system under consideration, the initial t_{obs} was indicatively defined on the order of 1 year, with the option to extend it depending on the requirement of the analysis (i.e. in case of absence of significant variability) and freedom of the telescope. Furthermore, the observation cadence was defined to be 1 observation per night, to sample even smaller scale variation. Note that such frequency is just an indication for the best-case scenario, as weather and down-times of the telescope, as well as quality limitation of the data, limit the number of observed nights to a significantly higher cadency, as later discussed in Section 5.1.
Initially, the observations were carried out in two Sloan filters, g' and i'. The multifilter observation should have been useful to further correct for microlensing effects (Rojas et al., 2020). The principle was based on the idea of taking advantage of the so-called chromatic microlensing, i.e. the change of colour in the lightcurves due to microlensing. This effect is caused by the small caustic size of microlenses, which can therefore lens a specific region of the QSO. Given our current model of QSO, these regions have different colours and thus the magnification appears different in different wavelengths. In practice, the infrared wavelength resulted in a lower SNR compared to the g' filter and had to be discarded from the analysis as it carried little information. This allowed us to interrupt the observation in such a filter, and therefore prolong and increase the number of exposures in the g' filter, thus increasing the SNR for the leading filter.
Moreover, as it will be discussed in the time-delay analysis in Section 5.3, the effect of microlensing appears not to be severely impacting the data compared to the intrinsic variability of the source, and therefore, a more sophisticated analysis was not required.
The final requirement to define a priori of the observations or during its first stages (once a few observations have been carried out) is the minimum quality requirement to satisfy for an observation to be carried out. Namely, this refers to the weather conditions of the night: the maximum seeing and minimum transparency necessary for an observation to be carried out and be useful for the analysis. Such requirements limit the number of data points in favour of their higher quality while optimising the cost of the campaign in terms of observation time. In this category, it can also be defined the exposure length of the single exposure as well as their total number relative to the sky brightness, which in turn is mostly dependent on the location and period of the moon or other human-made light sources. For the analysis at hand, it was defined to have a maximum seeing of 1.9" and a minimum transparency of 70%. Furthermore, the length of the exposures was seen not to be significantly affected by the "bright time", but the number of exposures per night varied from 6 ("dark time") to 12 ("bright time").
Once such parameters were set, the observational campaign was carried out at the LMU Wendelstein Observatory using the 2.1-meter Wendelstein Telescope equipped with a Wide-Field Imager (WWFI) (Hopp et al., 2014; Kosyra et al., 2014). The details of the observations and their data reduction are described in Section 5.1. I introduce the method employed for the time delay analysis in Section 5.2. I then present the time delay analysis in Section 5.3. The error estimation on the time delay results is described in Section 5.4. Given the specific parametric

approach taken, such analysis results in a set of results; how these results are combined is then described in Section 5.5.

5.1 WST Data for Lightcurve Compilation

The campaign lasted from 7/02/2020 till 15/06/2023; this resulted in significantly longer than the originally planned campaign of one year due to the requirement of the observation (as the QSO presented low to moderate variability for the first ~2 years of observations, as later discussed) and availability of the telescope itself. Each night of observation was then reduced with the standard data-reduction pipeline developed for the `WST` described in detail in Kluge (2020) and Kluge et al. (2020). While the pipeline described encompasses the most general cases, we here require a specific data reduction, for which few options are superfluous. An overview of the process can be seen in the flow chart 5.1, which is Figure 10.1 of Kluge (2020) tailored to the observation at hand. The pipeline first subtracts the bias voltage, obtained by a monthly calibration (referred to as the "monthly masterbias"), and subtracts the overscan values. It then aligns the CCD regions and computes the relative gains of each read-out region. Subsequently, the image is divided by the flatfield exposure, which is taken daily (referred to as the "daily masterflat"). After this, the programme then identifies and masks bad pixels (considering hot pixels, cosmic rays and saturated pixels). It then converts the counts into photon counts by multiplying the gain. The propagation of the statistical uncertainty provides the corresponding error frames for each exposure. A first approximated astrometry is obtained from the observation coordinates. The charge persistence is masked, and the bias offset residuals are matched between the 4 read-out regions of the CCD. Finally, the pipeline creates a star catalogue for each night. At this point, the programme is paused and each exposure is visually inspected to verify the quality of the observation, identify possible failures in the computations and manually mask the satellite traces. To speed up this tedious process, the masking is only applied to the CCD quadrant where the QSO is located. After this manual intervention, the pipeline is restarted. The reduced single exposures are resampled and stacked together, producing the final combined image for the night. The same process is applied to the error frames, which return the corresponding error image.

By comparing it with the original, it can be seen that for the observation of J1433 there is no need for modelling the bright stars' effects, such as stellar ghosts and halos, as the field around the object has few of them and most of them fall out of the interested quadrant (see Figure 5.3).

Once the single reduced night is produced, a tailored analysis of the reduced images is required to accurately extract the photometry. This falls within the general description of PSF-photometry, although with specifically high requirements to reach the necessary precision for time delay measurements.

During the data reduction, the PSF full-width half maximum (FWHM) is automatically estimated with a non-parametric approach from the surface brightness profile of field stars. This FWHM is the seeing estimate for the given exposure and, once averaged over the multiple exposures, represents a first estimate for the seeing of the given night. This is used to discard "bad" nights with a too large seeing, i.e. seeing $> 1.9''$, as previously mentioned. A first photometric

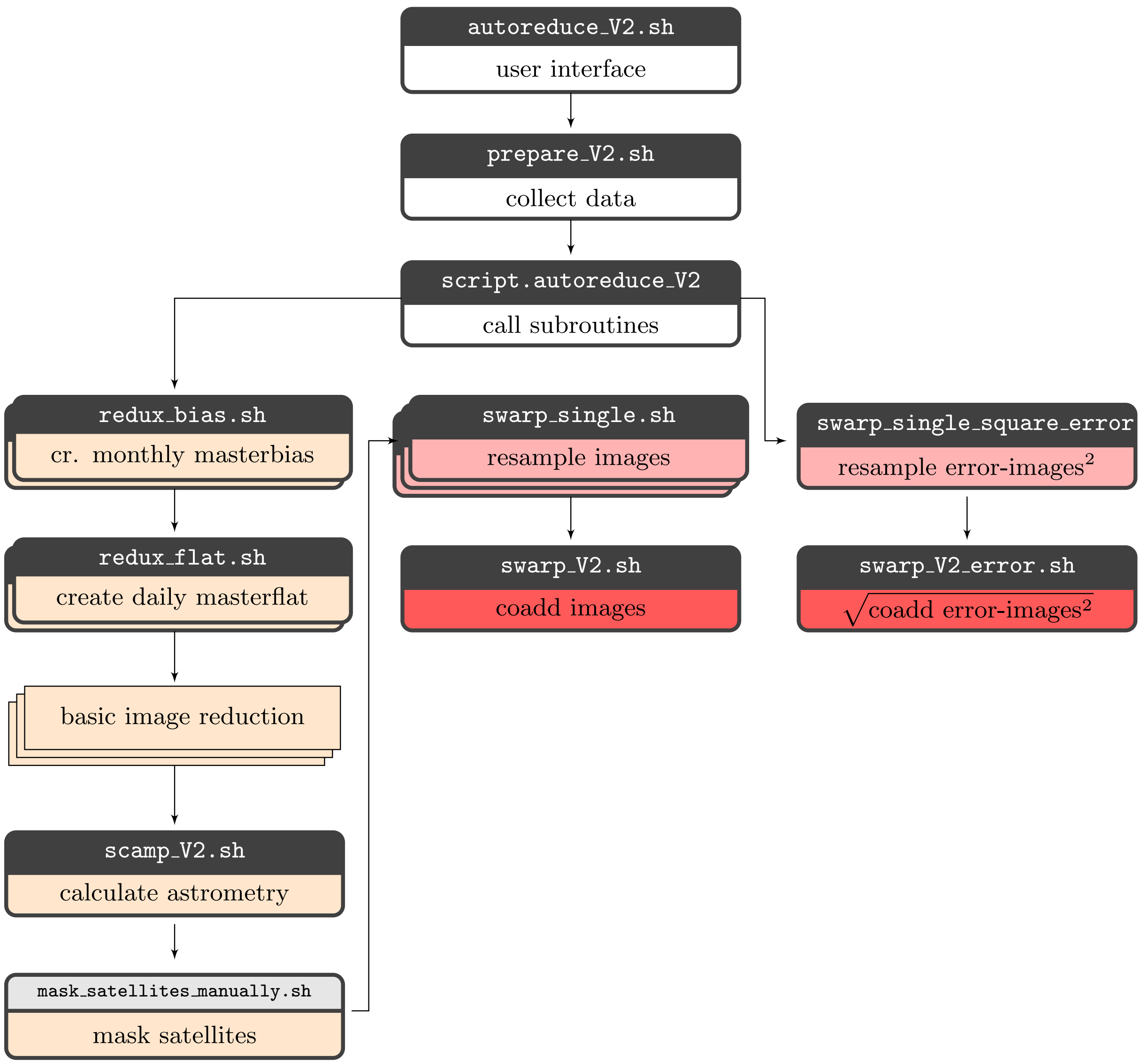

Figure 5.1: Flowchart for the data reduction of the single night observation for `WST`, tailored for the lightcurve analysis of J1433. This is a simplified version of the more general chart presented in Figure 10.1 of Kluge (2020).

zeropoint (ZP) estimate is also obtained by comparing approximately 100 stars in the field of each observation to the Pan-STARSS1 DR2 catalogue (Flewelling et al., 2016). This is done by measuring their aperture photometry using a circular aperture with 25 pixels diameter, to be compatible with the catalogues' methods. This can produce some seeing-dependent errors, as the smearing of the light can be severe enough that the flux may leak out of the aperture, affecting the precision of the ZP calibration. This method is estimated to be precise on the order of 0.2 mag, not enough for the purpose of this analysis. To obtain a higher photometric precision, another specific pipeline was developed based on PSF photometry. While this method would still rely on the first ZP calibration based on the PD2 catalogues, any unknown error would correspond to a systematic shift in magnitude. This would not affect the time delay analysis as it only depends on relative luminosity variations.

Firstly, the seeing of the night is measured again by taking the FWHM of the PSF model obtained from a "seeing reference star". This star is selected to be close to J1433, ensuring that the seeing remains unchanged between the two objects, and bright enough to be constantly visible with high SNR, while not too bright to cause a non-linear response in the detection. The selection of the same star for all observations ensures an unbiased measurement of the seeing. Next, a supersampled PSF model is created following a similar approach as described in Section 4.2.3, i.e. using the programme `psf` (Riffeser, 2006), based on eight stars in the field around J1433, referred to as the "PSF reference stars". In this case, the exact position of the centre for each of these stars is obtained by fitting them with a Moffat profile with an FWHM fixed to the one previously measured from the "seeing reference star". The centre coordinates are then used as a fixed parameter in the PSF modelling procedure, which relies on running `psf` on these "PSF reference stars". Note that the added step of measuring the centre improves the PSF model precision, as such ground-based observation presents a significantly larger PSF spread compared to `HST`. Having defined "a priori" the center allows for a more stable PSF model. The obtained PSF model is supersampled by a factor of 5 and is then used for the ZP calibration and further for the measurement of the QSO images' photometry.

The ZP calibration (or rather, re-calibration as it corrects for second-order deviation from the first estimate based on the Pan-STARSS1 catalogue) is done by measuring the luminosity of stable stars in the field, referred to as "ZP reference stars". To select such stars, a large sample of field stars in the vicinity of J1433 is considered, and their PSF photometry is measured for a few months at the beginning of the campaign. The stars that appeared significantly variable with respect to the average were discarded. This resulted in two stars being stable with respect to each other. This can be seen in Figures 5.2 where their magnitude in *g'* is plotted after the subtraction of their reference magnitude and the ZP correction. This corresponds to the equation:

$$\Delta\mathrm{mag}_{i,S_j} = \mathrm{mag}_{i,S_j} - \mathrm{mag}_{\mathrm{ref},S_j} - \delta\mathrm{ZP}_i, \tag{5.1}$$

where $\Delta\mathrm{mag}_{i,S_j}$ is the resulting data point for ZP star S_j for the i-th night, mag_{i,S_j} is the measured magnitude of S_j for the i-th night, $\mathrm{mag}_{\mathrm{ref},S_j}$ is the reference magnitude of the S_j star and $\delta\mathrm{ZP}_i$ is the ZP correction for the corresponding i-th night. Such correction is measured by averaging the difference of the measured magnitude of the ZP reference stars at the i-th night with their reference magnitude, i.e. $\delta\mathrm{ZP}_i\langle\mathrm{mag}_{S_j,i} - \mathrm{mag}_{\mathrm{ref},S_j}\rangle_j$.

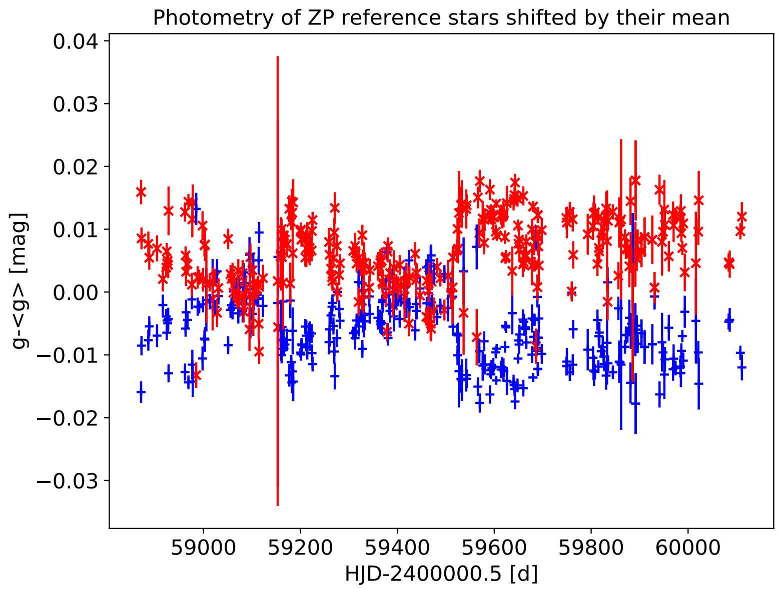

Figure 5.2: PSF photometry in *g'* filter of the “ZP reference stars” (see Table 5.1 and Figure 5.3) corrected for their reference magnitude and ZP correction (see Equation 5.1).

R.A. (J2000)	Dec. (J2000)	
14:33:26.46	+60:06:25.27	Seeing Reference Star
14:33:15.12	+60:07:46.45	ZP Reference Star 1 ($mag_{g'}$ = 18.064)
14:32:53.65	+60:08:35.26	ZP Reference Star 2 ($mag_{g'}$ = 18.306)
14:33:01.28	+60:08:38.29	
14:32:58.96	+60:09:01.82	
14:33:04.72	+60:06:10.16	
14:32:58.31	+60:05:20.49	
14:32:53.35	+60:05:06.44	

Table 5.1: Sexagesimal sky coordinates of the “PSF reference stars” used for the calibration of the observations. The first one is the star considered when estimating the seeing. The second and third are the reference stars used for the ZP calibration with the corresponding magnitude measured for the reference night (31/07/2020, see Figure 5.3) and taken as the reference magnitude for all observations.

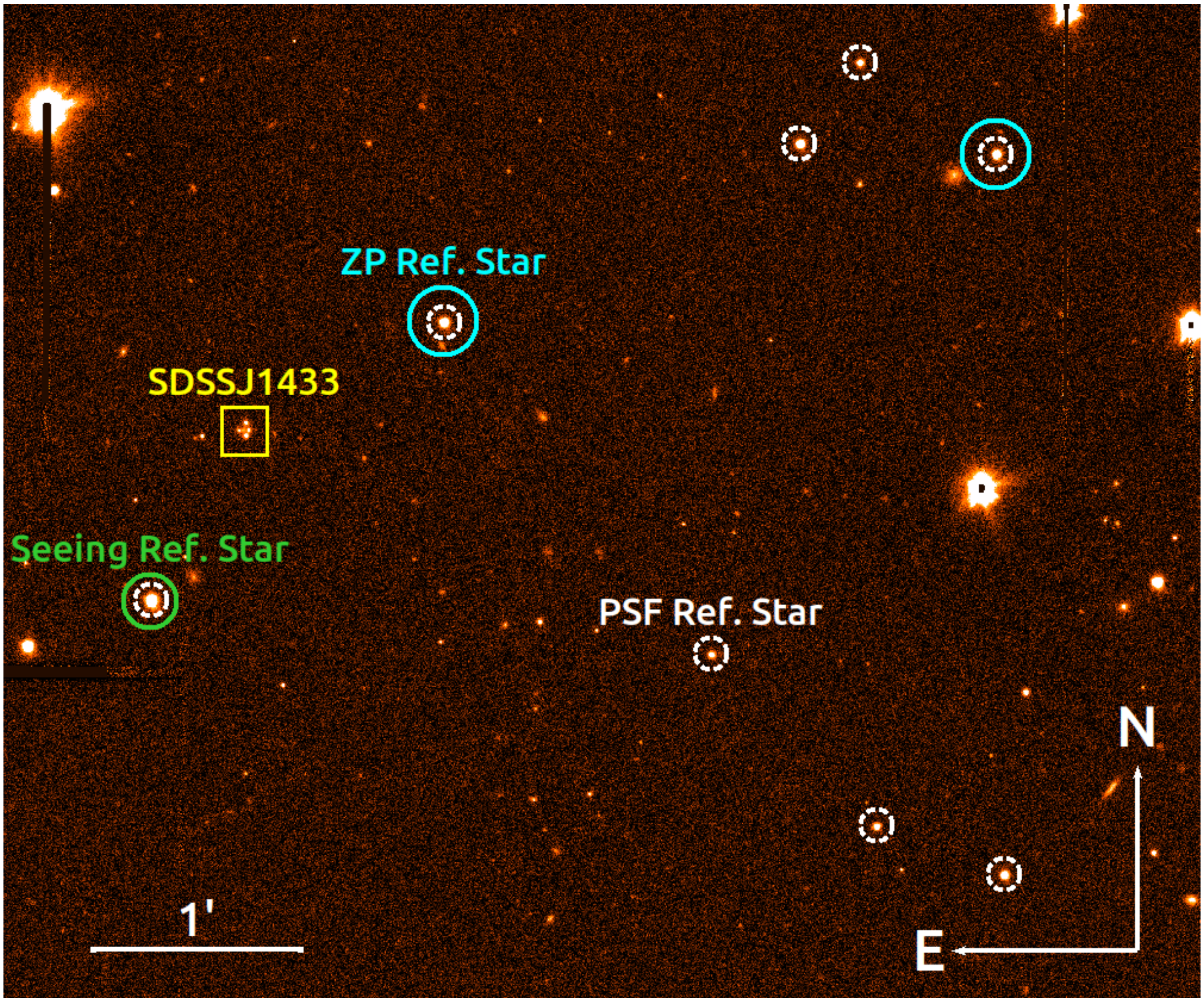

Figure 5.3: Reference image: this is the observation taken the 31/07/2020. Here are indicated the PSF reference stars (shown in white), ZP reference stars (shown in cyan) and seeing reference star (shown in green) for J1433 (in the yellow box). Their coordinates are shown in Table 5.1

The reference magnitude $\mathrm{mag}_{\mathrm{ref},S_j}$ is the magnitude of the given S_j star as observed on the reference night, 31/07/2020 (shown in Figure 5.3, values reported in Table 5.1). This night was selected at the beginning of the campaign for its good - but not optimal, in order to avoid it being an outlier - observational conditions (seeing of 0.88" and transparency of 99%).

As a validation measure, this process was replicated for the images obtained in filter *i'*, and the same set of stars resulted to be stable in that filter. The resulting standard deviation of $\Delta\mathrm{mag}_{i,S}$ is $\sim$ 0.006 mag for both ZP reference stars S. This has a negligible effect on the lightcurves error budget. Furthermore, note that any systematic error introduced here in the ZP calibration, i.e. a constant bias or shift in the magnitude of the lightcurves, would not have an effect on the time delay analysis, which only relies on the information carried in the relative shape of the lightcurves. Moreover, if such bias is a shared constant between the lightcurves, which is a

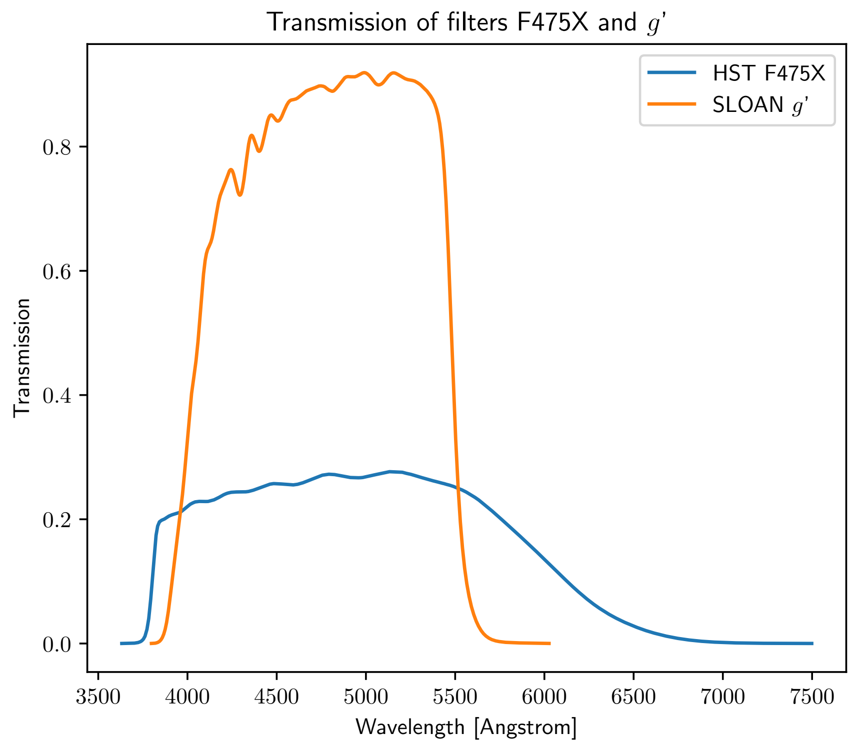

Figure 5.4: Comparison of transmission of F475X from camera WFC3, UVIS2 of *HST* and SDSS *g'* as a function of wavelengths. Note that, apart from the disparity in scale, which can be accounted for by adjusting the ZP, the two filters transmit very similar light. The data is obtained from Observatory (08/05/2024).

reasonable assumption, it would not affect the relative magnification shift, and thus would not propagate into the analysis of the flux ratio described in the Appendix A.
Before measuring the photometry of the images given our PSF model and the ZP calibration, it is necessary to subtract all interfering light, most notably the main lens light, the perturber light and all nearby sources that might blend with the light of the images. To do so, I used the *HST* exposure F475X, previously employed for the lens modelling (Chapter 4), as its bandwidth and centring are comparable to those of the *g'* band, as seen in Figure 5.4. Given the higher resolution of the *HST* images, deblending the light between the different sources is trivial, as was previously described in Chapter 4 with respect to the lens light subtraction. In this case, the objective is instead to remove the light of the QSO images. To do so, I produce a new PSF model for the F475X, now considering several bright but unsaturated stars in the field around J1433. This PSF model is regridded at the position of the images, rescaled to match the luminosity of the individual QSO images and subtracted. Note that the objective and therefore the methodology of such PSF modelling is here different from the one previously described in Section 4.2.3, as the PSF model obtained for the lens modelling was used to fit the position of the images, and

therefore required a very precise core and central inner part of the wings. Here, instead, the primary objective is to subtract the outer wings of the QSO, which are the ones blending with the other sources. The residuals at the centre are masked and interpolated from neighbouring pixels. Once successfully subtracted all light of the QSO images, the pixel resolution of the resulting *HST* image is degraded to that of *WST* of $0.2\frac{''}{\mathrm{pix}}$. The image is then shifted, rotated, cropped and finally resampled to the pixel grid of the *WST* observation. This image is indicated as the *HSTtoWST* image.
For every night of observation, this image has to be convolved to the observation seeing. This is done with a three-step process, first by convolving a first time the *HSTtoWST* image to the seeing of the night by means of the PSF model obtained from the "PSF reference stars", then the astrometry is measured again and corrected for small deviations between the observed night and *HSTtoWST*. This corrected astrometry is then applied to the initial *HSTtoWST* image, and the convolution kernel is computed again. Once this second, astrometry-corrected kernel is obtained, the *HSTtoWST* image can be accurately convolved to the seeing of the observed night. The reason for this iterative approach is that the convolution kernel is highly sensitive to precise astrometry, and in turn, small deviations of astrometry can be corrected only once the seeing is compatible. Thus, a first convolution is implemented for the astrometry to be corrected, and with that, a second, more accurate convolution kernel can be obtained.
Once the *HSTtoWST* image is brought to the correct seeing, its ZP is adjusted in order to the ZP of the observation night previously obtained. This is furthermore scaled by an additional factor, which is required due to the difference of *HST*'s ZP. The numerical value of this scaling factor is obtained empirically by minimising the residual and resulting to be 1.11.
The resulting *HSTtoWST* image, now convolved to the same seeing of the night and with the correct ZP, is subtracted from the observation, as shown in Figure 5.5 for the reference night (31/07/2020). Now that all interfering light has been subtracted, the QSO's images photometry can be computed. This is done by taking the previously obtained PSF model for the night, resampling it at the position of the images and rescaling it to match the flux of each QSO image while optimising for a common shift of their position. This last adjustment takes into account the remaining astrometric imprecision, on the order of 0.3" or 1.5 pixels. Such shift and scale factors are obtained by minimising the residual. The scale factor can then be converted into magnitudes with the ZP previously refined.
As previously mentioned, to limit the noise, the data was cut, discarding the exposures with sky transparency lower than 70% and/or seeing larger than 1.9". Note that this was applied to single exposures, and given the variability of such parameters during the night, thus in a few cases, this led to discarding a few exposures over the whole night of observations. However, many nights had to be completely discarded, leading to a cut of roughly 31% of the total number of observed nights. Note that this figure also includes cases where the observations were interrupted due to bad weather or other circumstances, which also resulted in bad observing conditions in the exposures taken. This is to say that the quality constraint might not be the primary cause for the loss of data. The resulting distribution for the quality of datapoints can be seen in Figure 5.6.

The final dataset comprised 297 data points obtained over three years with a median magnitude precision of $0.015 \leq \sigma_{\mathrm{mag}} \leq 0.055$ depending on the brightness of the QSO images. The light-

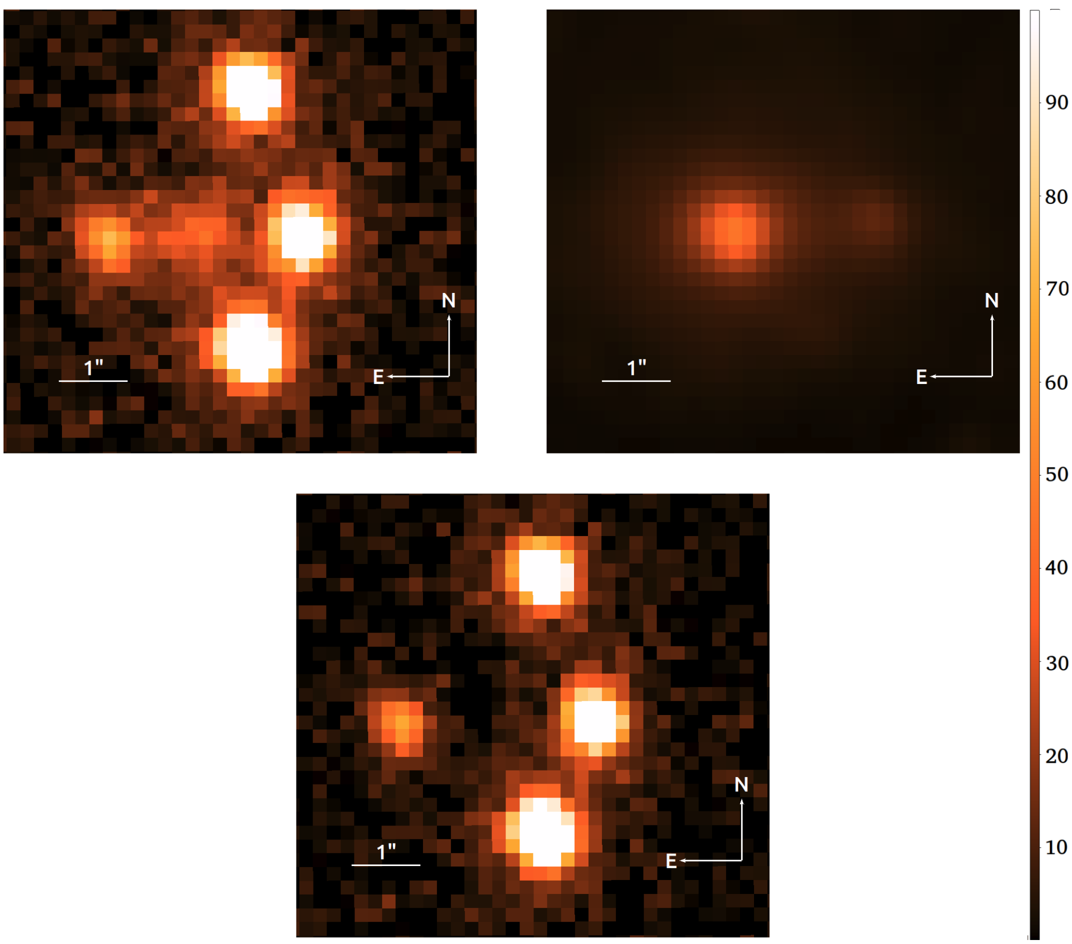

Figure 5.5: Example of the procedure followed to subtract all constant sources of light from the single observation night (in this case, the "reference night" 31/07/2020). Top left: reduced *WST* observation. Top right: the *HSTtoWST* image further convolved with the seeing of the observed night and rescaled to the correct ZP. Bottom: Resulting subtracted image, where only the QSO light remains.

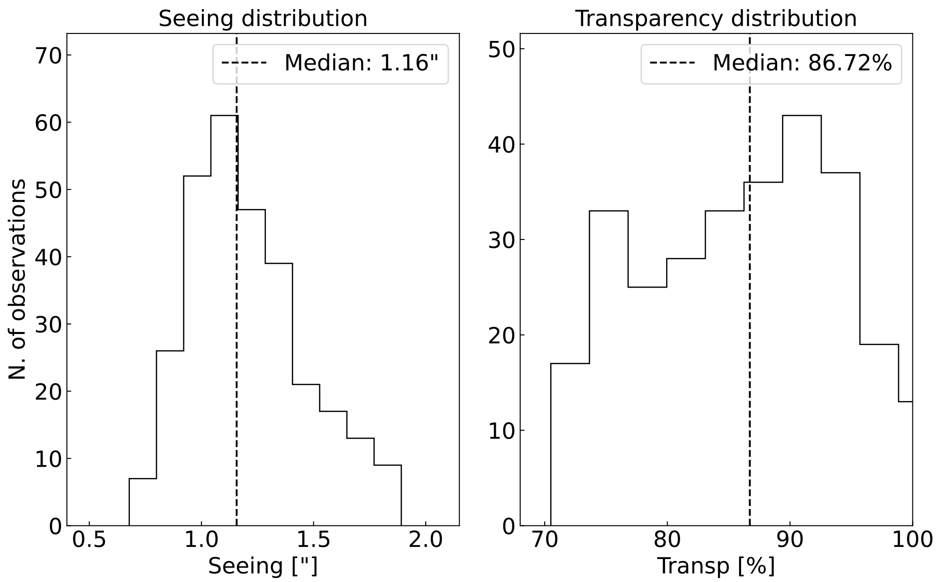

Figure 5.6: Distribution of seeing and sky transparency for the observations used in the light-curve analysis in Section 5.3 after the quality cutoff. The medians are 1.16 ” for the seeing and 86.72% for the transparency.

curves can be seen in Figure 5.7. The specifics of the observations are summarised in Table 5.2.

It can be seen that such lightcurves show clear signs of variability, so much so that even by eye one can see the time delay between them, although the most notable features (i.e. the minimum around HJD 2459600/February of 2022 and the maximum around HJD 2459900/December of 2022, depending on the lightcurves) only appeared on the last year and a half of observations. Also consider the different SNR between the images due to their intrinsic magnification, as image D is significantly dimmer and thus shows a clear scatter and higher photometric noise. Note several interruptions of observations, mostly due to technical problems and connected downtimes of the telescope, as well as more general problems such as the global COVID pandemic.
In conclusion, the data obtained is characterised by a high SNR, and high sampling and shows clear signs of intrinsic variations, with the overall observation period extended enough to cover all expected time delays, and is therefore an optimal dataset for time delay measurements, which will be presented in the remaining Sections of this chapter.

5.2 Introduction to PyCS3

For the time delay analysis, I relied on the Python library `PyCS3` (Python Curve Shifting) described in Tewes et al. (2013), the state-of-the-art time delay analytical procedure developed

Period of Observation	7 February 2020- 15 June 2023
Filter	Sloan g'
Total observed Nights	432
N. Datapoints	297
Sampling	4.1 days
Pixel scale	0.2 "/pixel
Median Seeing	1.16 "
Median Transparency	86.72 %
Median σ_{mag} [mag]	0.015 - 0.055
Exposure Time (bright time)	12×240 s
Exposure Time (dark time)	6×240 s
Relative Photometric Error	1.6 % - 5.7 %

Table 5.2: Specifics of *WST* observations. The final number of data points is 69 % of the total number of observed nights due to quality constraints. The sampling is the mean number of days between the observations and is computed with respect to the number of data points used. Note that the median σ_{mag} indicates the range of median uncertainty on the magnitudes of the light-curves. Thus, it depends on the luminosity of the images. Here, the two extremes are shown for images B and D, respectively, the brightest and dimmest. Due to the dark and bright time variation, the number of exposures varies by a factor of two.

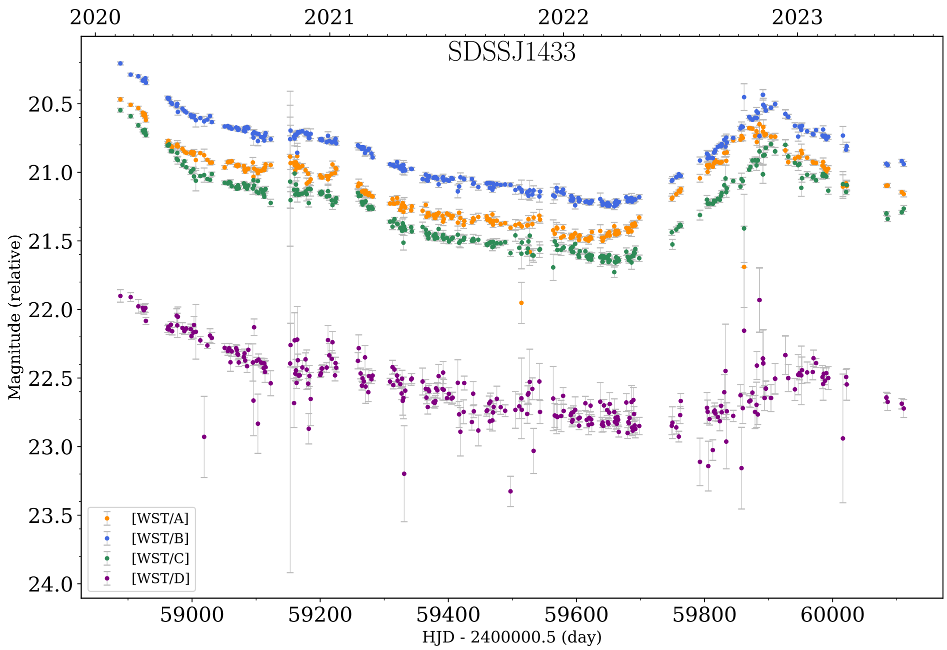

Figure 5.7: Resulting light-curves for the QSO's images observed from the WWFI at the 2.1 Meter Telescope of the Wendelstein Observatory in the g' band.

by the *COSMOGRAIL* collaboration [1]. This library allowed for two alternative time delay measurement methods: the spline fitting method and the Gaussian kernel regression method. These two methods are explained in detail in Tewes et al. (2013) and later refined in Millon et al. (2020a), but I will here briefly recapitulate their principle. The spline fitting method is based on the assumption that the intrinsic variability of the QSO (i.e. the lightcurve as it would be observed in the absence of the lens) can be modelled by a free-knot basis spline (or B-spline) of degree 3. These are specific splines, i.e. piecewise polynomials, whose knots, i.e. coordinates where the polynomial pieces connect, are free to vary and therefore have to be optimised alongside the polynomial parameters. The degree 3 means that the second derivative of the spline is continuous along the curve, thus ensuring a certain "smoothness" to the curve. Note that the freedom of the knots ensures a larger flexibility for the model, while also not forcing any discrete grid. However, this comes with the trade-off of a higher computational cost; in fact, optimising the spline parameters for a set of fixed knots is a linear problem, and therefore easily resolved. This is not true for the minimisation of χ^2 when the knot is a free parameter, as the computation becomes non-linear. The presence of several local optima and stationary areas in this parameter space further complicates the computation, and this method thus requires a specific approach to be successful. To solve this, the programme makes use of "bounded optimal knots" (BOK) described in Molinari et al. (2004), which optimise the knot position for a least-square spline approximation. In practice, the computation is divided into two steps; first, the knots are iteratively optimised within certain bounds; once fixed these are, the parameters are optimised

[1] `https://www.epfl.ch/labs/lastro/scientific-activities/cosmograil/`

linearly. This approach is iteratively repeated until convergence. The bounded approach to the knot position further ensures that the knots respect a minimum distance, avoiding the overlap of the knots, which would create discontinuities in the spline derivative.
However, the problem of lensed lightcurves is the presence of microlensing variation, referred to as the extrinsic variability of the lightcurves (see Chapter 2). These are not accounted for by a single intrinsic spline and have to be corrected differently. For the spline fitting method, this is obtained by considering, for each lightcurve necessary, an individual, "extrinsic" additive lightcurve to correct for such effects. The added microlensing correction results effectively in an added degree of freedom to adapt the individual lightcurve to the common "intrinsic" lightcurve. To avoid overfitting, the variability of such extrinsic correction has to be lower than the intrinsic lightcurve. Physically, this corresponds to the requirement that the intrinsic variability must dominate the signal in order to be able to measure it. In principle, this is not a given and depends on the system. In the case of lightcurves observed to be microlensing dominated, recovering a time delay with any significance would be impossible. When considering more moderate cases, where the microlensing is present but not dominant, it still raises the question of how much freedom to allow such extrinsic correction. In practice, these are modelled by low-degree polynomials or a degree 3 spline with a single, fixed knot, whose parameters are optimised with the intrinsic spline parameters. It has been seen that a free knot generates degeneracies with the fitting of the intrinsic lightcurve and is therefore avoided. Further note that, given the definition of such extrinsic correction, this is effectively a correction for all correlated noise in the single lightcurve. The most obvious case of noise is microlensing, and that is why this is considered a microlensing correction, but in principle, it is a correction for any correlated and smoothly varying noise in the data, known or unknown.
In conclusion, this is a parametric model that requires the definition of the number of knots for the intrinsic lightcurve, referred to as the knot step η, the minimum distance between such knots, ϵ, and the type and degree of freedom of the microlensing correction.
The Gaussian kernel regression model is instead a non-parametric model, in which each lightcurve is fitted by a Gaussian process regression. The modelled lightcurves are then subtracted pairwise, and their relative time delay is a free parameter optimised to minimise the variability in the residual.
These two methods are based on different approaches to the analysis of the lightcurves and result in similarly well-constrained time delays, in agreement between themselves (Millon et al., 2020a, see e.g. Figure 7 in). While it would be cautious to repeat the analysis with both to verify the stability of the results, preliminary analysis with the Gaussian kernel estimation resulted in weakly constrained results. Moreover, the uncertainty estimation for both methods, later described in detail in Section 5.4, requires a set of simulated lightcurves that are produced by a generative model based on the spline fitting approach. Thus, the Gaussian regression analysis can be considered not completely independent from the other method. It was therefore discarded and was left for a future expansion upon this work to test that such a method agrees with the results reported here. I will instead focus on the spline fitting method and explain in this and the following section my implementation for the analysis of J1433.
The overall structure of the analysis follows the method described in Millon et al. (2020a). This is divided roughly into three steps, corresponding to the three following Sections: the analysis

(Section 5.3), the uncertainty estimation (Section 5.4) and the combination of results (Section 5.5). This results from the parametric approach of this method and the impossibility of defining a priori the correct set of parameters for the analysis. Thus, a broader search is conducted, which is possible due to the low computational cost of the analysis and the reliable method for the combination of results, later described. In this approach, multiple sets of parameters, indicated by $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2 ..., \mathcal{S}_j, ...\}$ are considered, and for each of those, the analysis is repeated 1000 times. The resulting distribution is interesting for two aspects: its scatter $\sigma_{\Delta t,an.}$ and its median $\langle \Delta t \rangle$.

$\sigma_{\Delta t,an.}$ is considered a metric of the "goodness of fit": a scatter larger than a given threshold, in this analysis defined to be $\sigma_{\text{thresh.}}$ =2 days, indicating that the given $\mathcal{S}$ is unable to stably fit the dataset and does not converge to a single result. This is usually the case for over-constraining models, i.e. with too large freedom, e.g. see Figure 2 in Tewes et al. (2013). These $\mathcal{S}$ are flagged as "bad" and are therefore discarded from the following steps of the analysis. $\sigma_{\Delta t,an.}$ is saved as the intrinsic scatter for the given $\mathcal{S}$ and is later combined with the total uncertainty, although it is usually a negligible contribution (as expected, see e.g. Tewes et al., 2013).

The median $\langle \Delta t \rangle$ of the distribution obtained from the analysis is instead considered the resulting time delay.

For each given $\mathcal{S}$, the uncertainty is then estimated via a "Monte-Carlo" approach: a generative model is used to produce a set of 800 simulated lightcurves with similar constraining power as the real data, and the analysis is repeated with randomised initial conditions and the same parameters of $\mathcal{S}$ as for the real data. The resulting time delay is compared with the true time delay injected in the simulated lightcurves, and their difference produces an error distribution. Such distribution carries information on the uncertainty of the time delay estimation for that $\mathcal{S}$, where the systematic is given by the median of the distribution and the random error by its scatter. Subsequently, the uncertainty for this $\mathcal{S}$ is obtained by adding them and $\sigma_{\Delta t,an.}$ in quadrature.

Finally, the results of each remaining $\mathcal{S}$ have to be combined. This could be done by selecting the best absolute result or by marginalising over all the results. However, both approaches are suboptimal: the best result could be biased by unconstrained systematics and its error could be underestimated, while the marginalisation might inflate the uncertainty more than necessary, overestimating it. Instead, following the approach of Tewes et al. (2013), I opted for a hybrid approach that minimises the tension. I define the tension between two sets $\mathcal{S}_\alpha$ and $\mathcal{S}_\beta$ as follows. Given a lightcurve pair, the corresponding time delay estimates for each set are ${\Delta t_\alpha}^{+\sigma_\alpha^+}_{-\sigma_\alpha^-}$ and ${\Delta t_\beta}^{+\sigma_\beta^+}_{-\sigma_\beta^-}$. Their tension is then

$$\tau(\Delta t_\alpha, \Delta t_\beta) = \frac{\Delta t_\alpha - \Delta t_\beta}{\sqrt{{\sigma_\alpha^-}^2 + {\sigma_\beta^+}^2}}, \tag{5.2}$$

given $\Delta t_\alpha > \Delta t_\beta$. Else, the signs of the equation are inverted. This is computed for each lightcurve pair, and the total tension between the two sets $\mathcal{S}_i$ and $\mathcal{S}_j$ is taken to be the maximum tension between the different lightcurve pairs.

The result combination is therefore done by combining the results that minimise such tension. To do so, first, the best result is selected, $\mathcal{S}_{\text{best}}$. This corresponds to the lowest average uncertainty

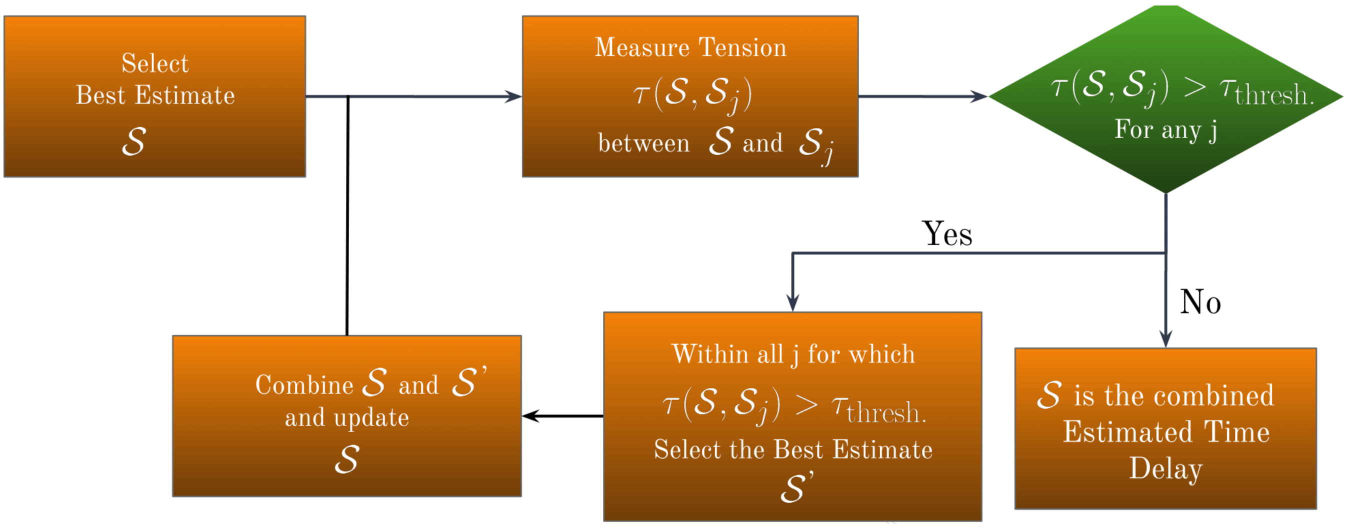

Figure 5.8: Flowchart for the combination of the time delay results given different sets of parameters $\mathcal{S}$ for which the analysis was run.

over all lightcurves couples. Then the tension between $\mathcal{S}_{\text{best}}$ and all other $\mathcal{S}$ is computed, and I defined a subgroup with all $\mathcal{S}_{\text{j}}$ that present a tension higher than a certain threshold, here defined to be $\tau_{\text{thresh.}} = 0.5$. Within this subgroup, I select the best $\mathcal{S}_{\text{best}'}$, following the same reasoning as before, i.e. the $\mathcal{S}$ which presents the lowest average time delay uncertainty. Following this, the $\mathcal{S}_{\text{best}}$ and $\mathcal{S}_{\text{best}'}$ results are combined, producing a new $\mathcal{S}_{\text{best}}$, and the process is repeated until there is no more tension between the final $\mathcal{S}_{\text{best}}$ and the other $\mathcal{S}$. I illustrate the flowchart in Figure 5.8.

5.3 Time Delay Analysis

The method chosen, the spline fitting method, is a parametric method that requires the user to define a few cardinal parameters. For the intrinsic spline, which is initialised with equally spaced knots, these are the knot step η, which indicates the initial time interval between the knots, and ϵ, the minimum distance between the knots. Both these parameters are defined in units of days. While ϵ is seen to have little effect on the results, η is the parameter that most crucially defines it. This is in fact inversely proportional to the number of knots present in the final spline and therefore dictates its flexibility. Its definition might be misleading, as it might indicate to the user that the knots are fixed. This is clearly not the case, but rather it can be seen as the first guess over the time-scale of the variability of the given lightcurve: a large η would indicate a slowly varying system, while a small η would correspond to a continuously varying intrinsic lightcurve. Given the freedom of the knots, however, it should be seen as the inverse of the number of knots, and rather than the absolute time scale of the variability, it would constrain the amount of variation over the observation.
The choice of η is therefore crucial in order not to overfit (η too small) or underfit the data (η

too large). A strongly/weakly variable curve requires a small/large knot step. In the case of the lightcurves obtained for J1433 and shown in Figure 5.7, intrinsic variations (i.e. common to all lightcurves) dominate its structure on medium to long time scales (~ 300 days). Smaller variations in shorter time scales (~10 days) can also be seen in the three brightest lightcurves. This led me to consider knot steps in a range between 30 and 45 days. I therefore selected $\eta = \{30, 35, 40, 45\}$. This corresponds to a number of knots of {41,35,31,27}, respectively.
Very short peaks, on the order of days or very long trends, over the whole observational period, might instead be caused by microlensing.
The very short peaks should not significantly affect the results as they only concern a few data points. Thus, this case is not accounted for in the analysis. Note that a different system, where such short-term, microlensing-induced variation was more common, might be analysed differently; e.g. by masking the affected data points.
The second microlensing effect, the very long-term one, is instead corrected explicitly by adding to every lightcurve a microlensing correction. As described in the introduction in Section 5.2, this is in fact an added lightcurve with some pre-determined degrees of freedom that is used to adjust every lightcurves to the common intrinsic spline. The flexibility of such an extrinsic lightcurve is therefore also important and depends on the time scale and intensity of the microlensing perturbance in the system. Due to the aforementioned assumption for the obtained lightcurve, I only considered a low degree of freedom for such corrections. Thus, I considered, for each lightcurve, either a 3-degree spline with one fixed knot or a polynomial with degrees ranging from 0 to 2. Regarding such choices, a few things are worth pointing out. Firstly, for each analysis, the same type of microlensing correction is applied to all lightcurves. This is done partly for simplicity, as it reduces the multiplicity of analyses carried out, and given the resulting low dependency on microlensing correction, it appears to be unnecessary. Furthermore, considering microlensing only for a subset of lightcurves, while possible, resulted in underconstrained results and was therefore discarded after a few tests. Secondly, the spline has a fixed knot, as leaving the knot free to vary is seen to overfit the data. In particular, it presents degeneracies with the intrinsic spline knot freedom, and the fit fails to converge to a stable solution in this configuration. Finally, the polynomial correction of degree zero corresponds to a constant value, i.e. a magnitude shift of the lightcurve. This is completely degenerate with the intrinsic magnification shift of the image and therefore can be viewed as fitting with no microlensing correction. It will be discussed later that this method appears to be too little constraining, i.e. underfitting the data; it is considered here as a benchmark, and its failure is a hint that extrinsic correction is necessary, even if with a low degree of freedom. Related to this case, it is important to note that this degeneracy, i.e. between the intrinsic magnitude shift of the lens and the average magnitude shift due to microlensing, is present for all microlensing corrections. This is a physical degeneracy: from the lightcurves alone, it is impossible to completely disentangle the luminosity variation due to microlensing or due to the lens. Two aspects come to help us here: microlensing magnification is limited to an order of magnitude lower than the intrinsic magnification, and it is time-variable. Thus, averaging the lightcurve over time would result in intrinsic magnification; this would be true if it were possible to average over timescales longer than the microlensing time scales, which in principle are undefined. It is believed, however, that microlensing should mostly appear on limited time scales. It is therefore reasonable to expect that the average magnification over

the observed period of over three years would be a fair estimate of the intrinsic magnification of the image. Thus I consider the total intrinsic magnification given by the magnification of the lightcurves, i.e. the shift $\Delta\mathrm{mag}_\mathrm{i}^\mathrm{intr.}$ applied to the individual i-th lightcurve to be matched to the intrinsic spline, plus the mean magnification of the extrinsic lightcurve for the given lightcurve $\langle\mathrm{mag}_\mathrm{i}^\mathrm{extr.}\rangle$. As with all measurements in this analysis, also this is a relative one, thus the measurement is the magnitude shift between two lightcurves, i and j, is given by $\Delta\mathrm{mag}_\mathrm{ij}^\mathrm{tot} = (\Delta\mathrm{mag}_\mathrm{j}^\mathrm{intr.} - \Delta\mathrm{mag}_\mathrm{i}^\mathrm{intr.}) + (\langle\mathrm{mag}_\mathrm{j}^\mathrm{extr.}\rangle - \langle\mathrm{mag}_\mathrm{i}^\mathrm{extr.}\rangle)$. This will be important, especially for the flux ratio anomaly discussed in the Appendix A, when comparing the obtained magnification to the predicted one from the lensing analysis, previously discussed in Chapter 4.
Given the aforementioned choices of parameters, the analysis resulted in a set of 16 different analyses, given 4 knot steps for the intrinsic spline and 4 possible microlensing corrections. To speed up the analysis and its convergence, the initial lightcurves are shifted in time and magnitude by a first rough estimate, in order to leave the optimisation a fine adjustment on the order of 10 days. The prior magnitude shift is obtained by taking the average of the difference in magnitude with respect to image A, weighted by the uncertainty of the data points. The prior time delay shift is instead obtained from the Fermat potential measured in the lens model, combined following equation 2.79 by assuming the default cosmology. This also means that such initial prior is dependent on an initial assumption of $H_0 = 70\,\frac{\mathrm{km}}{\mathrm{sMpc}}$. To avoid being biased by such a choice, and, more generally, to marginalise over all prior shifts and to estimate the stability of the analysis, each analysis is repeated 1000 times. For each iteration, the initial magnitude is randomised within a range of 0.5 mag from the initial prior, and the time shift is within a range of 10 days from the expected time delay. This time scatter approximately corresponds to varying H_0 between 50 to $100\,\frac{\mathrm{km}}{\mathrm{sMpc}}$ depending on the Fermat potential. This is equivalent to the later assumed prior on H_0 (see Chapter 6), and therefore the bias of this initial shift is rendered negligible.
The result of one of the analyses is plotted in Figure 5.9. In particular, this example corresponds to one of the 1000 analyses where the intrinsic knotstep is 45 days and the microlensing is modelled with a spline. It can be seen how the intrinsic variation of the lightcurves is well recovered by the black spline. In order not to clutter the text, the analyses for the other 15 sets of parameters are not reported here. The repetition of the analyses produces a distribution of time delays which can be analysed. An example of such distribution is shown in Figure 5.10. The set of parameters used for that analysis is the same as for the analysis shown in Figure 5.9. Note how the scatter of image pair AD is significantly larger than the other, due to the larger photometric uncertainty of image D coupled with its larger time delay. The latter implies that the overlap of the lightcurves is marginally limited compared to the others, and thus the constraining power of the lightcurves' data might be reduced as well.

As mentioned in the introduction in Section 5.2, this is used both to obtain an estimate of the time delay and to discard the "bad" $\mathcal{S}$ by verifying that the intrinsic scatter $\sigma_{\Delta t,\mathrm{an.}}$ is always lower than the chosen threshold of 2 days for all $\mathcal{S}$. In Figure 5.11, the $\sigma_{\Delta t,\mathrm{an.}}$ is reported for the different $\mathcal{S}$, and the different lightcurves pair.

Firstly, note that in all cases, image D causes a significant increase in time delay uncertainty. We can also see the importance of the cut-off. While such a threshold is mostly overestimated for the image pair AB and AC, it clearly distinguishes which set $\mathcal{S}$ are converging to a time delay

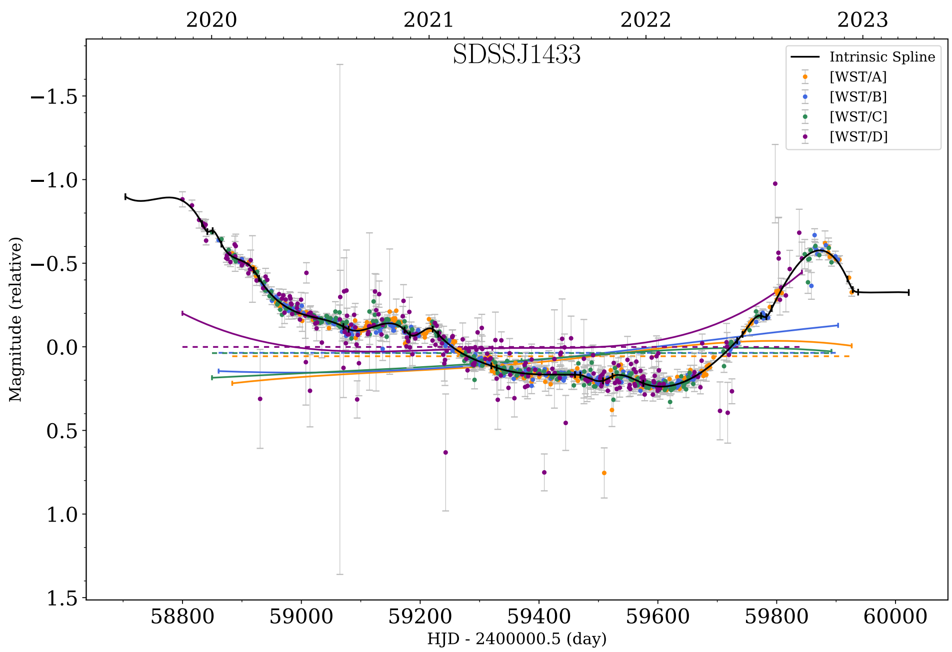

Figure 5.9: Time-delay analysis with `PyCS3` on the J1433 light-curves shown in Figure 5.7. The black curve is the resulting intrinsic spline, representing the common variability. The coloured points correspond to the data point of each lightcurve shifted by time delay, magnitude shift, and microlensing correction. The latter is plotted in the colour of the corresponding lightcurve. This is shown as an example of the analysis and is one of the 1000 iterations obtained considering an initial knot step of 45 days for the intrinsic spline and a spline microlensing correction

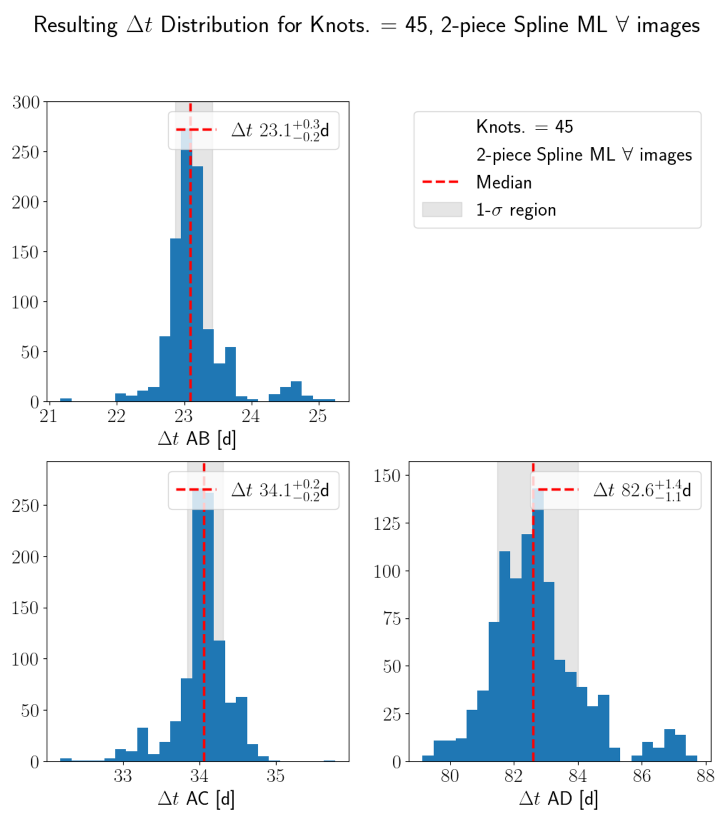

Figure 5.10: Time delay result distribution from the 1000 analyses of the lightcurves for a given $\mathcal{S}$. Knots. indicates the initial knot step of the intrinsic lightcurve. “Spline ML ∀ images” indicates that the spline microlensing correction is applied to all lightcurves.

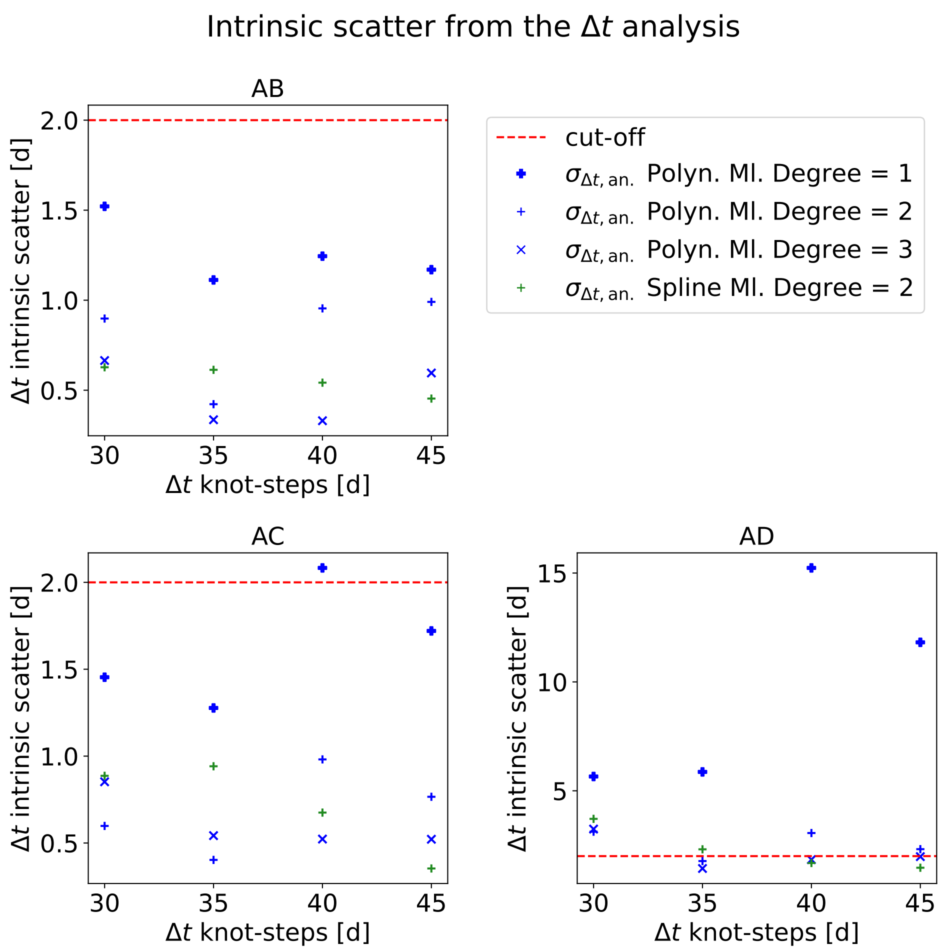

Figure 5.11: Resulting $\sigma_{\Delta \mathrm{t, an.}}$ for the various knot steps of the intrinsic spline and microlensing configuration of the $\mathcal{S}$. The blue crosses indicate polynomial microlensing, the different shapes depending on their degree, while the green ones indicate spline microlensing. The red dotted line indicated the 2 day cut-off.

for the image pair AD. In this case, we can see the polynomial correction of degree 1, i.e. the case where no effective microlensing is considered, appears higher (and for most of the chosen knot steps significantly higher) than the threshold. Also consider that, ignoring AD, a similar structure appears for the other image pair, i.e. a higher scatter when no microlensing is assumed. Thus, this model performs consistently worse than the others. This indicates that, while weak, the microlensing correction is required for this system. A similar case is the knot step 30, which is consistently unable to fit the image pair AD. Some other cases also present signs of instability in fitting the lightcurve D. As mentioned, all $\mathcal{S}$ that present a scatter higher than the cut-off are discarded from the analysis, and therefore the no-microlensing $\mathcal{S}$ is henceforth discarded. It might seem that such a cut is too strict, but it is necessary due to the required precision of the analysis.
From the time delay result distribution, I also obtain the time delay estimates by taking the median. The results of the analysis are reported in Table 5.3. From this table, it can be seen that the scatter $\sigma_{\Delta t,\mathrm{an.}}$ is a poor estimator for the uncertainty, as it severely underestimates the scatter between the different results. For this reason, an ulterior analysis is required to constrain the random and systematic uncertainty, which is described in the following Section 5.4.

5.4 Time Delay Uncertainty Estimation

The methodology to constrain the uncertainty of the time delay follows a "Monte Carlo" approach as described in Tewes et al. (2013), coupled with a stable generative model to simulate lightcurves. These simulated lightcurves will have, by construction, a similar constraining power as the real data. Thus repeating the analysis with the same set of parameters $\mathcal{S}$ as in the real data and comparing its result with the "real" simulated time delays, $\sigma_{\Delta t} = \Delta t^{\mathrm{real}}_{\mathrm{sim}} - \Delta t^{\mathrm{meas.}}_{\mathrm{sim}}$, would give us an indication of both random and systematic error for the given $\mathcal{S}$ applied to this dataset. Note that such analysis on simulated lightcurves has to be repeated multiple times, to marginalise over the random noise configuration of the simulated data and the randomisation of the initial time delay and magnification. As a byproduct of this method, the magnification ratio uncertainty can be estimated as well, which I will discuss in the appendix A when analysing the flux ratio anomaly.
A flowchart of the steps is shown in Figure 5.12. First, for each $\mathcal{S}$, the analysis is run once more on the real data. The output will now be the base for the simulations. The time delays and magnification ratio are uniformly randomised around the estimated values in a range of ± 10 days and ± 0.5 mag, respectively. The intrinsic spline is instead taken as the model for the intrinsic variability of the QSO. Once time-shifted and magnitude-shifted by the simulated values for each image, the obtained lightcurves are added to the extrinsic lightcurves obtained by the fitting, i.e. the modelled microlensing effects are added to the synthetic lightcurves. These lightcurves are then sampled on the same dates as the real observations. The photometric uncertainty of each data point is set to be identical to that of the corresponding observations. It can be pointed out that the extrinsic microlensing fit does not consider all "spectra" of the noise, as very fast extrinsic variability, correlated noise and shot noise are not modelled by it and therefore have to be injected in the simulated lightcurves in a different way. For such fast extrinsic variability, the code randomly generated a "power-law noise", which is an additive noise component to the lightcurve

Image Pair	Intrinsic Spline	Microlensing Correction		Δt [d]	$\sigma_{\Delta t,\mathrm{an.}}$ [d]
	Knot Steps η	ML Type	ML Degree		
AB	35	Polynom.	2	21.09	0.42
AC				34.67	0.40
AD				76.72	1.77
AB			3	21.51	0.33
AC				32.53	0.54
AD				78.57	1.43
AB	40	Polynom.	3	21.58	0.33
AC				32.69	0.52
AD				78.92	1.83
AB		Spline		22.73	0.54
AC				34.03	0.68
AD				83.20	1.68
AB	45	Polynom.	2	22.09	0.60
AC				32.88	0.52
AD				78.85	1.98
AB		Spline		23.09	0.45
AC				34.06	0.35
AD				82.59	1.47

Table 5.3: Time delay results and their scatter for different analysis configurations after the cut-off.

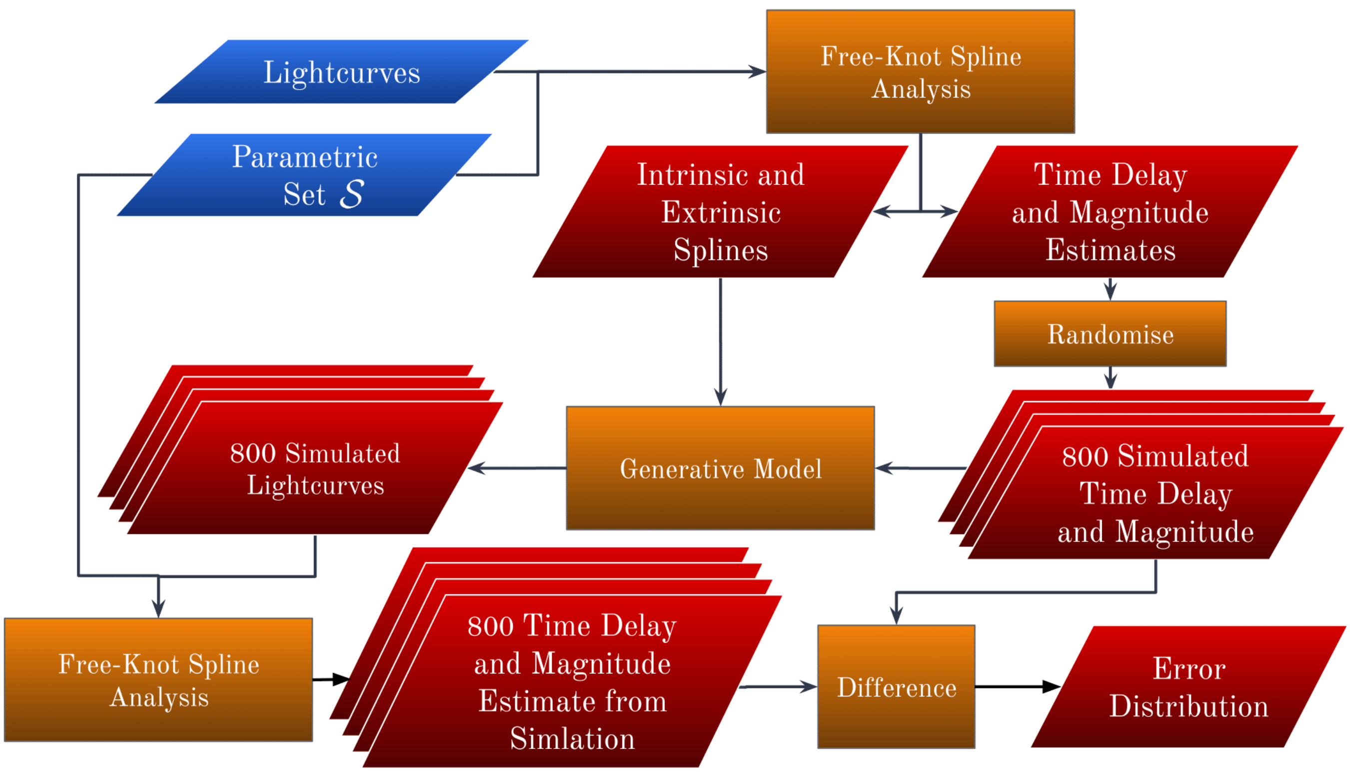

Figure 5.12: Flowchart for the error estimation of the time delay measurement for a given set of parameters $\mathcal{S}$.

whose Fourier spectrum follows a power law. The parameters of this additive component are iteratively optimised for every set of synthetic lightcurve and for every individual QSO lightcurve within such a set. The procedure, described in detail in Tewes et al. (2013) is the following:

- The power-law noise is drawn for each lightcurve with a fine regular sampling.
- The corresponding signals are linearly interpolated at the observation time.
- The noise is rescaled locally such that its amplitude follows the observed scatter.
- The analysis is repeated on this synthetic lightcurve, and the residuals are compared to those obtained from the analysis of the observed lightcurves.
- The power-law parameters are iteratively optimised until the residuals are statistically compatible.

The power-law is characterised by the minimum and maximum frequency the exponent, β, of the power-law and the scaling factor for the noise, A. The first two are defined to be f_{min} =1/500 days^{-1}, as correlated signals with lower frequency are well fitted by the explicitly defined extrinsic microlensing, and f_{max}=0.2, defined by the maximum Nyquist frequency from the sampling. Interestingly, such frequency parameters, which affect the sampling, have little effect. On the other hand, the two "shape" parameters, β and A, influence directly the efficiency of the time delay estimation on the simulated lightcurves, and are therefore optimised in order to achieve a similar constraining power as for the real data.
Once the power-law is optimised, the noise is converted into a real signal via inverse Fourier transformation, and such "lightcurve noise" is interpolated at the observation date, yielding ϵ_i. This is further locally rescaled by a factor s_i, in order to match the scatter of the observed nights for the given date. This factor is obtained by analysing the residuals of each of the observed lightcurves, $r_{i,\mathrm{obs}}$. Taking their absolute values, they are normalised to one over the whole residuals, and they are smoothed by a median filter with a window of seven observations, as described in Equation 8 of Tewes et al. (2013):

$$s_i = \text{median}\left(\frac{|r_{j,\mathrm{obs}}|}{|r_{\mathrm{obs}}|}, j \in \{i-3, ..., i+3\}\right). \tag{5.3}$$

Note that the average amplitude of the synthetic noise is not affected by s_i, which, averaged over the whole lightcurve, tends to 1. It is instead dependent on A.
The optimisation of β and A is carried out by taking the residuals as a metric. In particular, their standard deviation σ and their number of "runs" r. This is defined as the sequence of consecutive residuals with common sign, i.e. either all > 0 or < 0. This statistic is used to test if the adjacent residuals are independent or not, based on (Wall, Jenkins, 2003, , Chapter 5). Take a large enough sample of N statistically independent data points, for which the analysis produces N_+ positive residuals and N_- negative ones. The number of runs r is expected to follow a normal distribution with $\mu_r = \frac{2N_+N_-}{N} + 1$ and $\sigma_r^2 = \frac{(\mu_r-1)(\mu_r-2)}{N-1}$ as its expectation value and variance, respectively.
It is then simple to verify if the obtained run distribution is in agreement with the theoretical one,

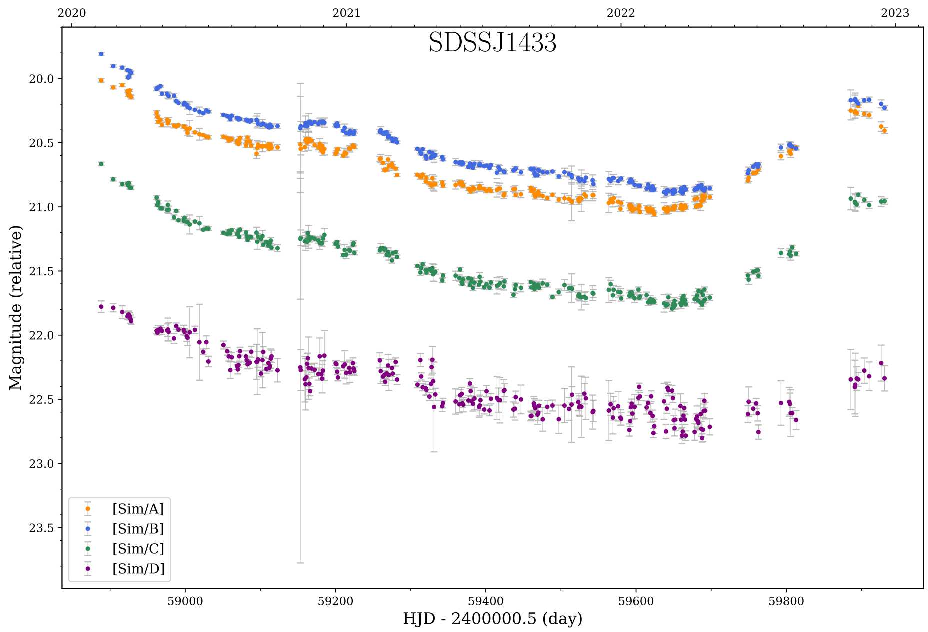

Figure 5.13: A random example of the simulated lightcurves for the QSO's images.

i.e. if the residuals are correlated, by taking $z_r = \frac{r-\mu_r}{\sigma_r}$. This value should be around 1 if such a hypothesis is correct. This is usually not the case by default, but can be achieved by fine-tuning β and A. Interestingly, such parameters are directly correlated with r and σ_r, respectively. This makes the procedure straightforward and therefore computationally fast, as only a few iterations are needed to converge.

Once the small-scale noise is added to the synthetic ligthcurves, these are ripe to be used for the uncertainty estimation. An example of such a lightcurve can be seen in Figure 5.13. Note how the lightcurve is indistinguishable from the real data, seen in Figure 5.7. This is only one of the 800 lightcurves produced using the generative model for a given set of parameter $\mathcal{S}$ (here I have taken the same as the previous examples, e.g. Figure 5.9: $\mathcal{S} = \{\eta = 45, \text{microlensing} = \text{spline}\}$). Finally, after the synthetic dataset is produced, the analysis is rerun for every $\mathcal{S}$. In this case, since the time delays and magnitude shifts that had been injected in the synthetic lightcurves are randomised around the measured values, it is not necessary to further add randomisation of the initial conditions, as previously done for the real data analysis in Section 5.3. The analysis is thus started around the expected values of time delay and magnitude shifts.

The results are then compared with the injected results, and their difference produces the "error distribution", shown in Figure 5.14 for the $\mathcal{S}$ usually taken for the examples. From such distribution, I measure the systematic error σ_{sys} and the random error $\sigma_{\Delta t,\text{rnd}}$ from the median and the width of the distribution. Specifically, the latter is measured as half the width of the 68% confidence interval; in practice, it is obtained by taking the average absolute difference between the 16% and 84% quantiles and the median. As it can be seen in Figure 5.14, the $\sigma_{\Delta t,\text{sys}}$ is in all cases very small and compatible with 0, while the random uncertainty dominates the error

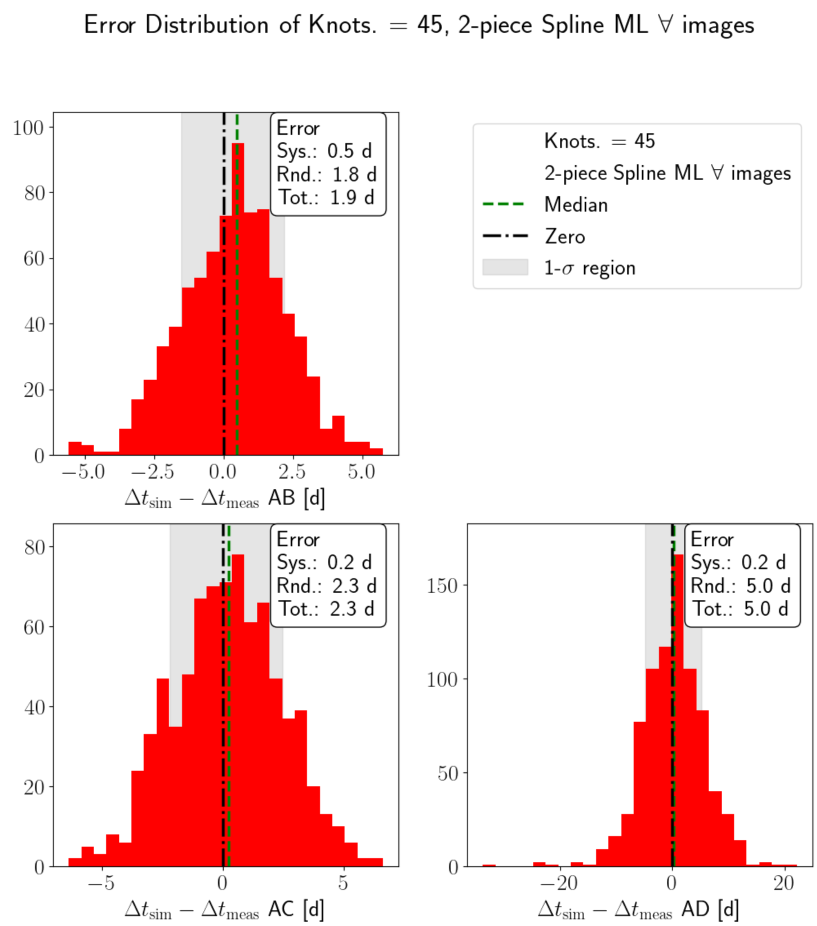

Figure 5.14: Error distribution of the time-delay estimate given the set of parameters $\mathcal{S} = \{\eta = 45, \text{microlensing} = \text{spline}\}$. See Figure 5.10 for the corresponding result distribution. Note that the error estimate is dominated by the random error, most notably for the time-delay AD, as expected due to the low magnification of image D and thus higher photometric uncertainty.

Image Pair	Intrinsic Spline	Microlensing Correction		$\sigma_{\Delta t,\text{an.}}$ [d]	$\sigma_{\Delta t,\text{sys}}$ [d]	$\sigma_{\Delta t,\text{rnd}}$ [d]	$\sigma_{\Delta t,\text{tot}}$ [d]
	Knot Steps η	ML Type	ML Degree				
AB	35	Polynom.	2	0.42	0.44	2.34	2.42
AC	35	Polynom.	2	0.40	0.43	2.49	2.56
AD	35	Polynom.	2	1.77	1.04	7.83	8.09
AB	35	Polynom.	3	0.33	0.26	2.34	2.38
AC	35	Polynom.	3	0.54	0.05	2.31	2.38
AD	35	Polynom.	3	1.43	0.77	6.86	7.05
AB	40	Polynom.	3	0.33	0.22	2.29	2.32
AC	40	Polynom.	3	0.52	0.12	2.29	2.36
AD	40	Polynom.	3	1.83	0.15	6.7	6.95
AB	40	Spline		0.54	0.67	1.89	2.08
AC	40	Spline		0.68	0.58	2.42	2.58
AD	40	Spline		1.68	0.54	5.24	5.53
AB	45	Polynom.	2	0.60	0.03	2.33	2.41
AC	45	Polynom.	2	0.52	0.24	2.58	2.64
AD	45	Polynom.	2	1.98	0.65	6.44	6.77
AB	45	Spline		0.45	0.47	1.84	1.95
AC	45	Spline		0.35	0.22	2.34	2.37
AD	45	Spline		1.47	0.25	5.04	5.25

Table 5.4: Resulting uncertainty for the time delay method for the various sets of parameters $\mathcal{S}$, divided into components and the total resulting error. Note how, with respect to the error components, the total error is dominated by the random uncertainty. With respect to the image pairs, AD carries the largest uncertainty, both from the analysis of the real data in $\sigma_{\Delta t,\text{an.}}$ and of the synthetic data in $\sigma_{\Delta t,\text{rnd}}$.

budget, especially for the image pair AD, as expected due to its low photometric accuracy, as previously mentioned. Also, note how this uncertainty is far larger than the scatter of the distribution obtained from the analysis of the real data, $\sigma_{\Delta\text{t},\text{an.}}$.
For each $\mathcal{S}$, the uncertainty for the time delay of each image pair is then computed as $\sigma_{\Delta t,ij} = \sqrt{\sigma^2_{\Delta t,ij,\text{rnd}} + \sigma^2_{\Delta t,ij,\text{sys}} + \sigma^2_{\Delta t,ij,\text{an.}}}$, where ij indicate the image pair. The resulting uncertainties, divided into the various components, are shown in Table 5.4. Note how, as expected, the random error is dominant in all cases. For each $\mathcal{S}$, we have now obtained the time delay estimates (see Section 5.3) and their relative uncertainties (see Table 5.4). I now have to combine the results, which are described in the following Section.

5.5 Time Delay Estimates Combination and Results

I combine the time delay estimates following the approach of Tewes et al. (2013), as described in detail in the introduction of the time delay analysis in Section 5.2, although with some small deviations.

The main takeaway is that this approach, rather than marginalising over all results or selecting only the best result, aims to combine a selection of results that minimises the tension for all the remaining ones. I previously described the process of selecting the results in Section 5.4, and it is further schematised in the flowchart in Figure 5.8. Her,e I will describe the implementation and how the results are combined in practice.

Firstly, I will refer to the set of parameters as $\mathcal{S}$, as previously, while the results of a given set $\mathcal{S}_{||}$ will be indicated by $\mathcal{G}_{||}$. This corresponds to an array of time delay estimates, $\mathbf{E}_{ij}$, for each given pair of images ij, thus $\mathcal{G}_{||} = [\mathbf{E}_{\mathrm{AB}}, \mathbf{E}_{\mathrm{AC}}, \mathbf{E}_{\mathrm{AD}}]$. $\mathbf{E}_{ij}$ is then the time delay estimate which contains the time delay and the corresponding uncertainty, $\mathbf{E}_{ij} = \Delta t_{ij,-\sigma_-}^{\ \ +\sigma_+}$. The choice for the best result is therefore simply the group $\mathcal{G}_{\mathrm{best}}$ which presents the lowest average error (averaging over the image pairs, the analytical, the systematic and the random error). This choice is made when selecting the best absolute group as a starting point, and, once defined, the subset of groups in tension with the given best group when selecting the best group within such "tension group".

Notice that, by considering the uncertainty separately for each $\mathbf{E}_{ij}$, i.e. for the estimate of the time delay of each lightcurve pair, all correlation between the lightcurves is neglected. Moreover, the uncertainty is treated as if it follows a normal distribution. This is a valid simplification for the choice of the best result, but might underestimate the error when combining the results. Instead, I opted for treating the uncertainty distribution as a proper posterior distribution. In order to do that, I record the median of the distribution as $\sigma_{\Delta t,\mathrm{sys}}$, and correct the distribution for this value, thus centring it on zero. The uncertainty obtained from the scatter of the analysis, i.e. $\sigma_{\Delta t,\mathrm{an.}}$, is also to be considered a systematic uncertainty, and is therefore added under quadrature to $\sigma_{\Delta t,\mathrm{sys}}$, from which results the corrected systematic uncertainty for the given group, indicated as σ_{sys}.

I then add the estimate of the time delay Δt to the distribution. The result is a distribution centred around the estimated result, whose width indicates the probability of the measurement. Net of the normalisation, this can be considered a probability distribution, comparable to the chains produced by an MCMC algorithm. Note that this is a three-dimensional distribution, given by the number of image pairs. The only caveat is that the systematic error is now not explicitly encoded in the distribution itself.

Given this approach, when two groups are combined, their distributions are stacked together. The resulting distribution is then considered the posterior distribution of the combined result, whose mean is the new result. The systematic errors of the two groups are added under quadrature, giving the new systematic error.

Once the iterative process of combining the best group in tension has converged, the resulting group is now in agreement with all other groups, while considering the smallest possible uncertainty. The result of the combination is shown in the Figure 5.15. The coloured results are the various groups for the remaining sets of parameters (after the quality cut-off of the analysis, see Figure 5.11). The selected groups for the combination are indicated by the (S) and the circle around their point. Given the fairly good agreement, the combination of only two sets is enough

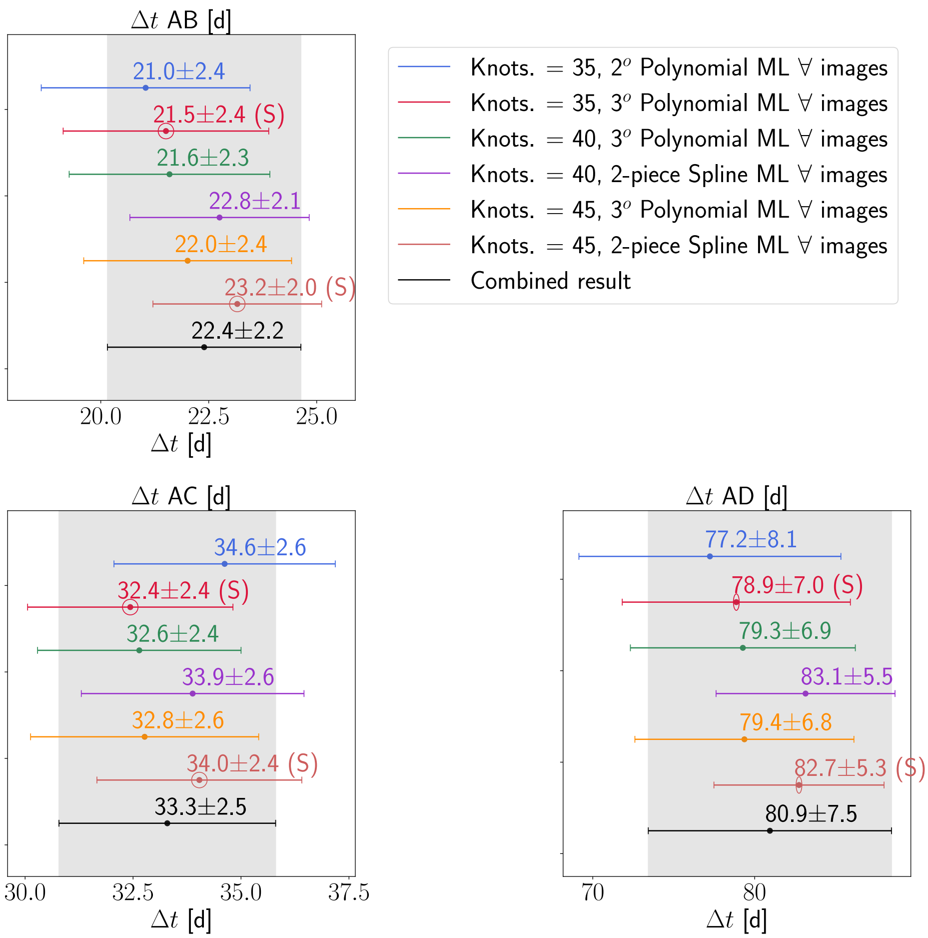

Figure 5.15: Results of the time-delay for the various light-curves couples and its combined final time-delay. The results iteratively selected for the "Combined result" are indicated with the (S).

to reach a combined result which is not in tension with any other result. The combined result is shown in black, and its 1-σ region is shaded in grey to underline the good agreement with all $\mathcal{G}$.

The numerical results are also reported in Table 5.5. Note, however, that the full covariances are not reported here, but are considered by taking the full distribution, which is shown in Figure 5.16.

Note how the resulting distribution is mostly Gaussian, as indicated by the black dotted line, which traces a Gaussian distribution centred in the result and which σ is obtained as the mean of the difference between the 84% quantile and the 16%. This can be more clearly seen in Figure 5.17, where I discarded any point further than 5 σ away from the median. However, the correlation between the time delays, as shown very clearly between AC and AB, should not be dismissed. This will be later taken into consideration in the following Chapter 6, where the results presented here will be summarised by a multivariate Gaussian.

Also note that in this analysis, we do not account for the microlensing time delay, introduced

Δt AB [d]	Δt AC [d]	Δt AD [d]
22.4 ± 2.2	33.3 ± 2.5	80.9 ±7.5

Table 5.5: Combined results for the time-delay between the pairs of images of J1433.

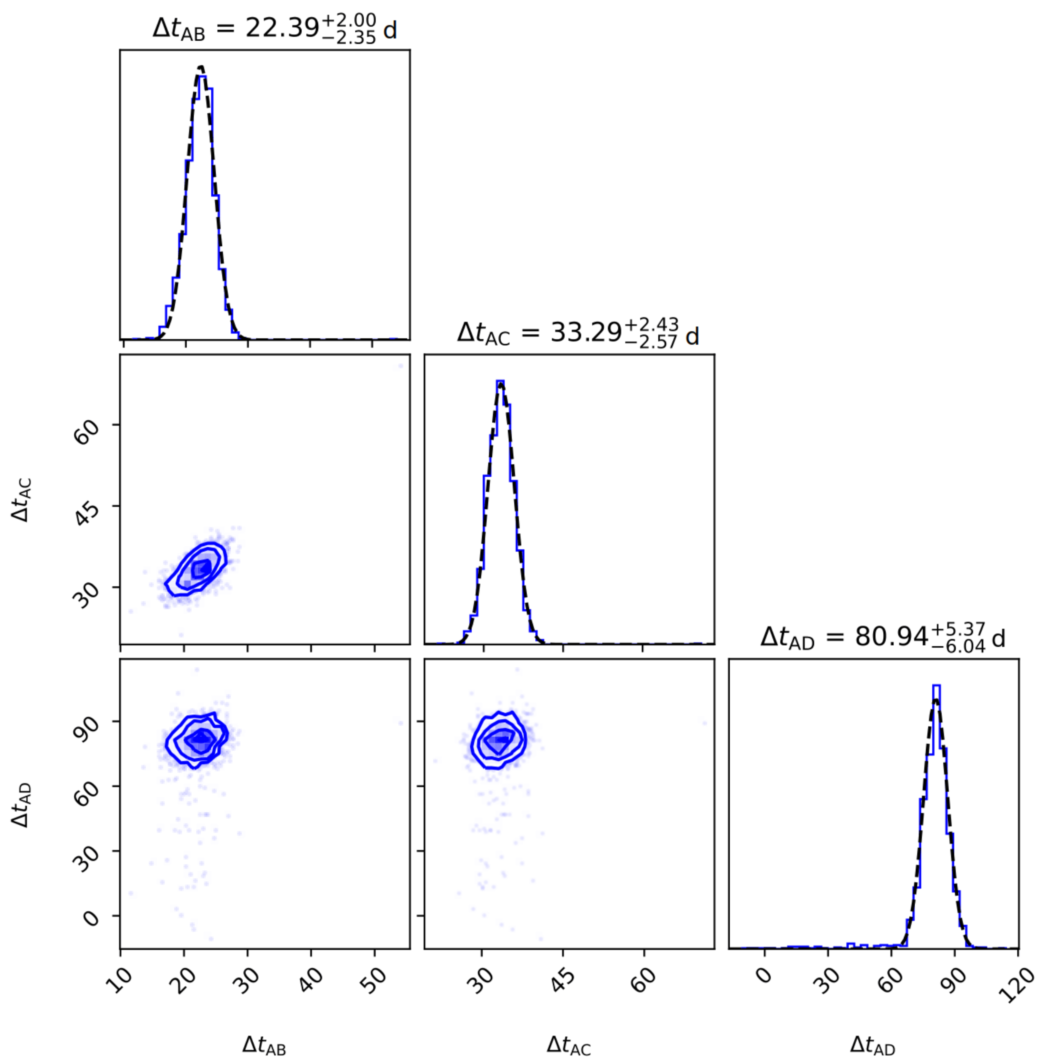

Figure 5.16: Corner plot of the distribution of the time delay result

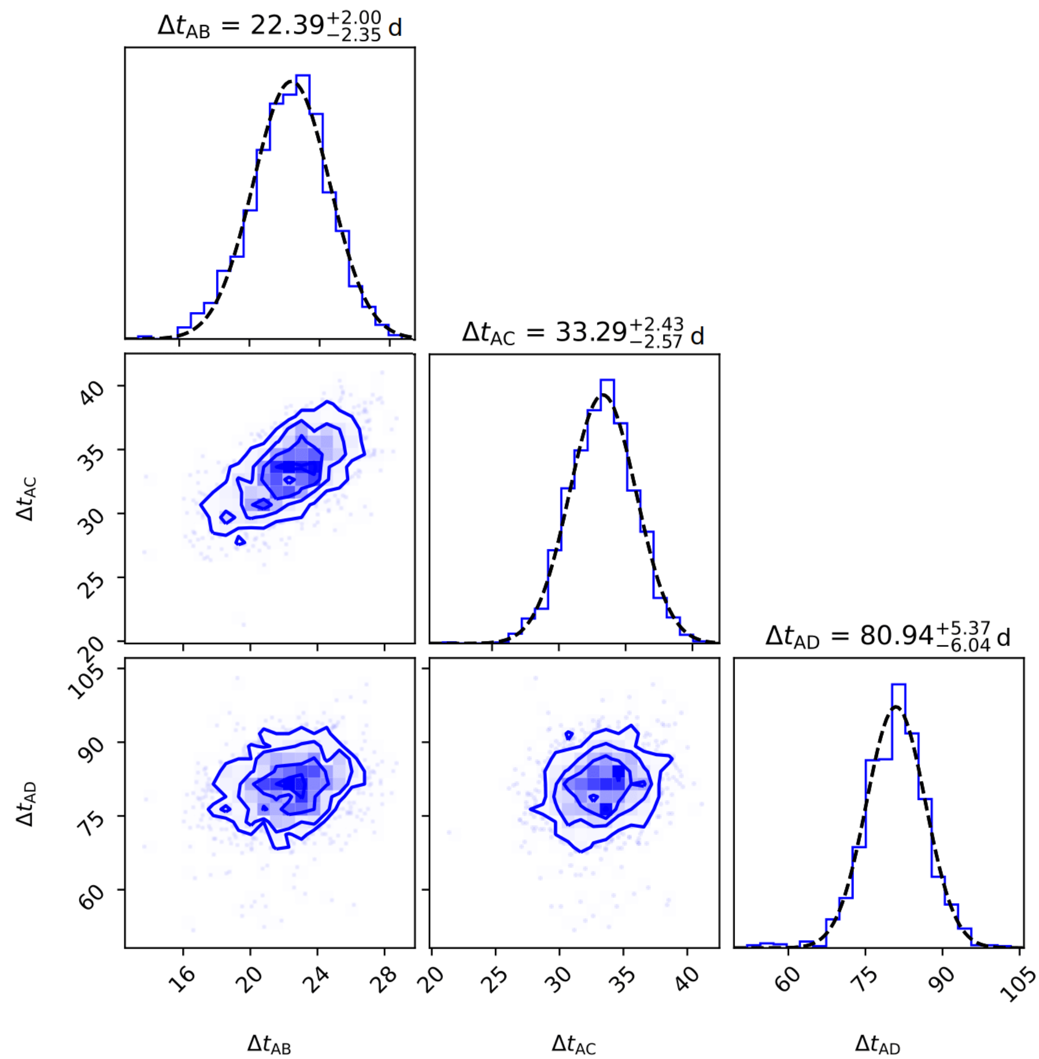

Figure 5.17: Close in on the corner plot of the time delay result distribution, discarding all points further than 5 σ from the median. This is for illustrative purposes only, to indicate how the 1-D distribution closely follows the Gaussian.

by Tie, Kochanek (2017). This phenomenon is due to the intrinsic extended nature of the AGN accretion disc. In the "lamp-post model" defined by Cackett et al. (2007), the disk's temperature variations are correlated with its luminosity variations. The temperature changes at the centre of the disk and propagate outwards, inducing correlated emission at different radii. This propagation, however, takes time, therefore creating a lag between the emission near the core and the one induced at the edges of the disk. In general, a larger disk will therefore have a longer lag. In the absence of microlensing, this phenomenon is barely observable, as the emission of the edges is far less luminous than the inner core. Moreover, they would similarly contribute to the lightcurves in each image. However, microlensing can magnify different regions of the accretion disk, and in different images, it could magnify the lagged signal differently. In practice, a microlensed image might appear further delayed and distorted with respect to the unperturbed image.

This phenomenon depends on the number of microlenses at the image position and the size and orientation of the accretion disk (Tie, Kochanek, 2017). The last two parameters are not observable directly, but we did somewhat constrain the microlensing intensity by fitting the lightcurves with a low degree of microlensing correction. This would indicate that J1433 is weakly affected by microlenses. For this reason, in this analysis, I do not explicitly account for microlensing time delay. This phenomenon may be nevertheless present in the lightcurves, but its systematic uncertainty would very likely be subdominant compared to the random uncertainty (see, e.g., Figure 6 of Bonvin et al., 2019, , the largest systematic effect would correspond to less than a day shift, well within the uncertainty of the measurement, see Figure 5.15). A more accurate study of microlensing time delay for J1433 is therefore left for a future expansion of this work.

Chapter 6

TDC@W: Constraint On H_0

Following the initial principle given by the equation 2.79, I have shown how to obtain constrain the mass profile of J1433 in Chapter 4, and obtain its time-delays in Chapter 5. In order to constrain the Hubble parameter H_0, I will then follow a Bayesian approach. This will be similar to what is described in most Time-Delay Cosmographic studies, such as (Suyu et al., 2013, , henceforth S13) and (Wong et al., 2017), although with some minor changes.
The reason for such deviation in the approach is only due to the implementation of the analysis, and I will show how they are indeed mathematically equivalent.

6.1 Joint Inference on H_0

Firstly, S13 does consider ulterior data to constrain the model which was not available for this analysis. These are the spectroscopic information on the lens, given by σ and r_{ani} (the lens velocity dispersion and the anisotropic radius for the stellar orbit, respectively) and properties of the lens environment, given by κ_{ext} and γ_{ext} (the external convergence and external shear, respectively). These are the constraints required to break the internal and external mass sheet degeneracy, as described in Section 2.3.3. Such constraints are independent of each other and of the constraints given by the lens model and time delay, as seen in equation 10 of S13. I can therefore constrain the cosmological parameter relying only on the available data, albeit without breaking the MSD. Thus, we can see that equations 8 to 11 can be adapted to the data available in this paper as follows:

$$P(H_0|\boldsymbol{D}_{HST}, \boldsymbol{D}_{\text{lc}}) = \int d\boldsymbol{\Delta\phi} \frac{P(\boldsymbol{D}_{HST}|\boldsymbol{\Delta\phi})}{P(\boldsymbol{D}_{HST})} \frac{P(\boldsymbol{D}_{\text{lc}}|\boldsymbol{\Delta\phi}, H_0)}{P(\boldsymbol{D}_{\text{lc}})} P(\boldsymbol{\Delta\phi}) P(H_0). \tag{6.1}$$

Here we defined the lightcurve data as $\boldsymbol{D}_{\text{lc}}$ and the HST data as $\boldsymbol{D}_{HST}$. I also limit the cosmological constraints to H_0, assuming all other cosmological parameters as constant. Furthermore, the integral is over the Fermat potential differences $\boldsymbol{\Delta\phi}$, instead of all lensing parameters. This is equivalent, as the former can be defined as the marginalisation of the latter.

Equation 6.1 assumes various relations of independence: the datasets being independent one from another (i.e. $P(\boldsymbol{D}_{HST}, \boldsymbol{D}_{\rm lc}|\boldsymbol{\Delta\phi}, H_0) = P(\boldsymbol{D}_{HST}|\boldsymbol{\Delta\phi}, H_0)P(\boldsymbol{D}_{\rm lc}|\boldsymbol{\Delta\phi}, H_0)$), the imaging data being independent on the cosmological parameter H_0 (i.e. $P(\boldsymbol{D}_{HST}|\boldsymbol{\Delta\phi}, H_0) = P(\boldsymbol{D}_{HST}|\boldsymbol{\Delta\phi})$) and the priors being independent one from another (i.e. $P(\boldsymbol{\Delta\phi}, H_0) = P(\boldsymbol{\Delta\phi})P(H_0)$).
Note, however, that the modular approach in the analysis carried out until now resulted in posterior distributions over $\boldsymbol{\Delta t}$ and $\boldsymbol{\Delta\phi}$, whereas equation 6.1 requires their corresponding likelihoods. Those can be inverted following Bayes' theorem, similarly to equation 4.17, as follows:

$$P(\boldsymbol{D}_{HST}|\boldsymbol{\Delta\phi}) = \frac{P(\boldsymbol{\Delta\phi}|\boldsymbol{D}_{HST})P(\boldsymbol{D}_{HST})}{P(\boldsymbol{\Delta\phi})} \tag{6.2}$$

$$P(\boldsymbol{D}_{\rm lc}|\boldsymbol{\Delta\phi}, H_0) = \frac{P(\boldsymbol{\Delta\phi}, H_0|\boldsymbol{D}_{\rm lc})P(\boldsymbol{D}_{\rm lc})}{P(\boldsymbol{\Delta\phi}, H_0)} = \frac{P(\boldsymbol{\Delta\phi}, H_0|\boldsymbol{D}_{\rm lc})P(\boldsymbol{D}_{\rm lc})}{P(\boldsymbol{\Delta\phi})P(H_0)}. \tag{6.3}$$

Note that the last passage is given by the aforementioned independence of the prior of $\boldsymbol{\Delta\phi}$ and H_0. The second line has to be further converted in order to be in agreement with equation 6.1 by considering $P(\boldsymbol{\Delta\phi}, H_0|\boldsymbol{D}_{\rm lc}) = P(\boldsymbol{\Delta\phi}|H_0, \boldsymbol{D}_{\rm lc})P(H_0|\boldsymbol{D}_{\rm lc}) = P(\boldsymbol{\Delta\phi}|H_0, \boldsymbol{D}_{\rm lc})P(H_0)$. This is given by the independence of H_0 from $\boldsymbol{D}_{\rm lc}$ alone.
Finally, equation 6.1 can be rewritten as:

$$P(H_0|\boldsymbol{D}_{HST}, \boldsymbol{D}_{\rm lc}) = \int d\boldsymbol{\Delta\phi} \frac{P(\boldsymbol{\Delta\phi}|\boldsymbol{D}_{HST})P(\boldsymbol{\Delta\phi}|H_0, \boldsymbol{D}_{\rm lc})P(H_0)}{P(\boldsymbol{\Delta\phi})}. \tag{6.4}$$

In this formulation, $\boldsymbol{\Delta\phi}$ is three-dimensional, and so is the integral on the right-hand side of the equation. Interestingly, given the reformulation detailed in the equations 6.2, it can be seen that the evidence in the denominator of equation 6.1 cancel out, while the prior $P(\boldsymbol{\Delta\phi})$ appears now in the denominator. The components of equation 6.4 are the following:

1. $P(\boldsymbol{\Delta\phi}|\boldsymbol{D}_{HST})$, which is the combined posterior of $\boldsymbol{\Delta\phi}$ obtained from the modelling of the *HST* images $\boldsymbol{D}_{HST}$ described in Section 4.5;

2. $P(H_0)$ and $P(\boldsymbol{\Delta\phi})$, the priors on H_0 and $\boldsymbol{\Delta\phi}$, respectively. The first is defined by taking a uniform prior between 50 and 100 $\frac{\rm km}{\rm s\,Mpc}$, whereas the second is obtained from the prior of the lens parameters (see Section 4.3.4 for its definition and Section 4.5 for its computation) ;

3. $P(\boldsymbol{\Delta\phi}|H_0, \boldsymbol{D}_{\rm lc})$, which is the posterior of $\boldsymbol{\Delta\phi}$ given the lightcurve data and H_0.

The last component, $P(\boldsymbol{\Delta\phi}|H_0, \boldsymbol{D}_{\rm lc})$, is obtained by converting the posterior on the time delay given the lightcurve data, $P(\boldsymbol{\Delta t}|\boldsymbol{D}_{\rm lc})$, using equation 2.79. Formally, this is a change of variable of the posterior distribution. In practice, this can be easily implemented given that $P(\boldsymbol{\Delta t}|\boldsymbol{D}_{\rm lc})$ is defined as a multivariate Gaussian centred around the measured $\boldsymbol{\Delta t}$ and which covariance matrix is obtained from the uncertainty estimation. These are obtained from the combination of time delay constraints described in Section 5.5. Equation 2.79 is linear, and thus the change of variable is straightforward:

$$\boldsymbol{\Delta\phi}(\boldsymbol{\Delta t}, H_0) = \boldsymbol{\Delta t}\frac{H_0}{k} \tag{6.5}$$

$$\mathrm{cov}(\boldsymbol{\Delta\phi}(\boldsymbol{\Delta t}, H_0)) = \mathrm{cov}(\boldsymbol{\Delta t})\left(\frac{H_0}{k}\right)^2. \tag{6.6}$$

Thus, for each value of the prior $P(H_0)$, $P(\boldsymbol{\Delta\phi}|H_0, \boldsymbol{D}_{\mathrm{lc}})$ can be computed.
In practice, the computation is implemented in several steps. Given the uniform prior $P(H_0)$, the sampling over H_0 is simply a range of equidistant points between the minimum ($H_{0,\,\mathrm{min}}$ =50 $\frac{\mathrm{km}}{\mathrm{Mpc\,s}}$) and maximum ($H_{0,\,\mathrm{max}}$ =100 $\frac{\mathrm{km}}{\mathrm{Mpc\,s}}$). For each of these points $H_{0,\,i}$, I compute the analytical posterior density distribution $P(\boldsymbol{\Delta\phi}|H_{0,\,i}, \boldsymbol{D}_{\mathrm{lc}})$ following the equation 6.5. Recalling that $P(\boldsymbol{\Delta\phi}|\boldsymbol{D}_{HST})$ is defined as a histogram over a set of bins, as described in Section 4.5, I then compute the value of $P(\boldsymbol{\Delta\phi}|H_{0,\,i}, \boldsymbol{D}_{\mathrm{lc}})$ at the centre of these bins. Similarly, the prior density distribution $P(\boldsymbol{\Delta\phi})$ was computed over the same set of bins; it is therefore possible to compute $\frac{P(\boldsymbol{\Delta\phi}|H_{0,\,i},\boldsymbol{D}_{\mathrm{lc}})_{\mathrm{bin\,j}}P(\boldsymbol{\Delta\phi}|\boldsymbol{D}_{HST})_{\mathrm{bin\,j}}}{P(\boldsymbol{\Delta\phi})_{\mathrm{bin\,j}}}d\boldsymbol{\Delta\phi}_{\mathrm{bin\,j}}$ for each i-th bin, where $d\boldsymbol{\Delta\phi}_{\mathrm{bin\,j}}$ is the volume of the bin. Note that the $d\boldsymbol{\Delta\phi}_{\mathrm{bin\,j}}$ is in this case constant, but this might be different given different implementations. The posterior $P(H_{0,\,i}|\boldsymbol{D}_{HST}, \boldsymbol{D}_{\mathrm{lc}})$ is then the sum over all the bins. Finally, the distribution $P(H_0|\boldsymbol{D}_{HST}, \boldsymbol{D}_{\mathrm{lc}}))$ is obtained by normalising over each $H_{0,\,i}$. The result is shown in Figure 6.1.

The resulting median and 1σ deviation are

$$H_0 = 71.3^{+5.0}_{-4.5}\,\frac{\mathrm{km}}{\mathrm{Mpc\ s}}. \tag{6.7}$$

This measurement has a precision of ~ 6.7 %, which is dominated by the time-delay uncertainty. Its relative error contribution is 11.2 %, which is unevenly divided between the image couples due to their accuracy (ranging between 8 % for AC and 15 % for AD). The difference of the Fermat potential relative error contribution is ~ 3.7 %, and is instead divided evenly to ~ 4 % for every image couple.

Note that the combined error of H_0 is based on the combination of three measurements, given the three combinations of images, which are considered independent.

Finally, it is important to consider the bias that this analysis has not taken into account. While the random error of both the time delay measurements and the Fermat potential measurements is carefully measured and integrated into the final result, I do not consider systematic errors. Most importantly, the mass-sheet degeneracy would directly bias the final constraint on H_0. I discuss the uncertainty over the redshift of the perturber in the Appendix B, but given the lack of suitable data (namely, deep spectroscopic observations of the perturber), the constraints are limited. I therefore do not integrate this source of uncertainty in the lens model. Finally, the results were never blinded at any point, which has become a customary approach for analysis which might be affected by confirmation bias or similar human-driven sources of bias. This is probably the only aspect which can not be improved upon, as data can not be formally re-blinded once observed. A future upgrade of the current work would require either the implementation of a new analysis,

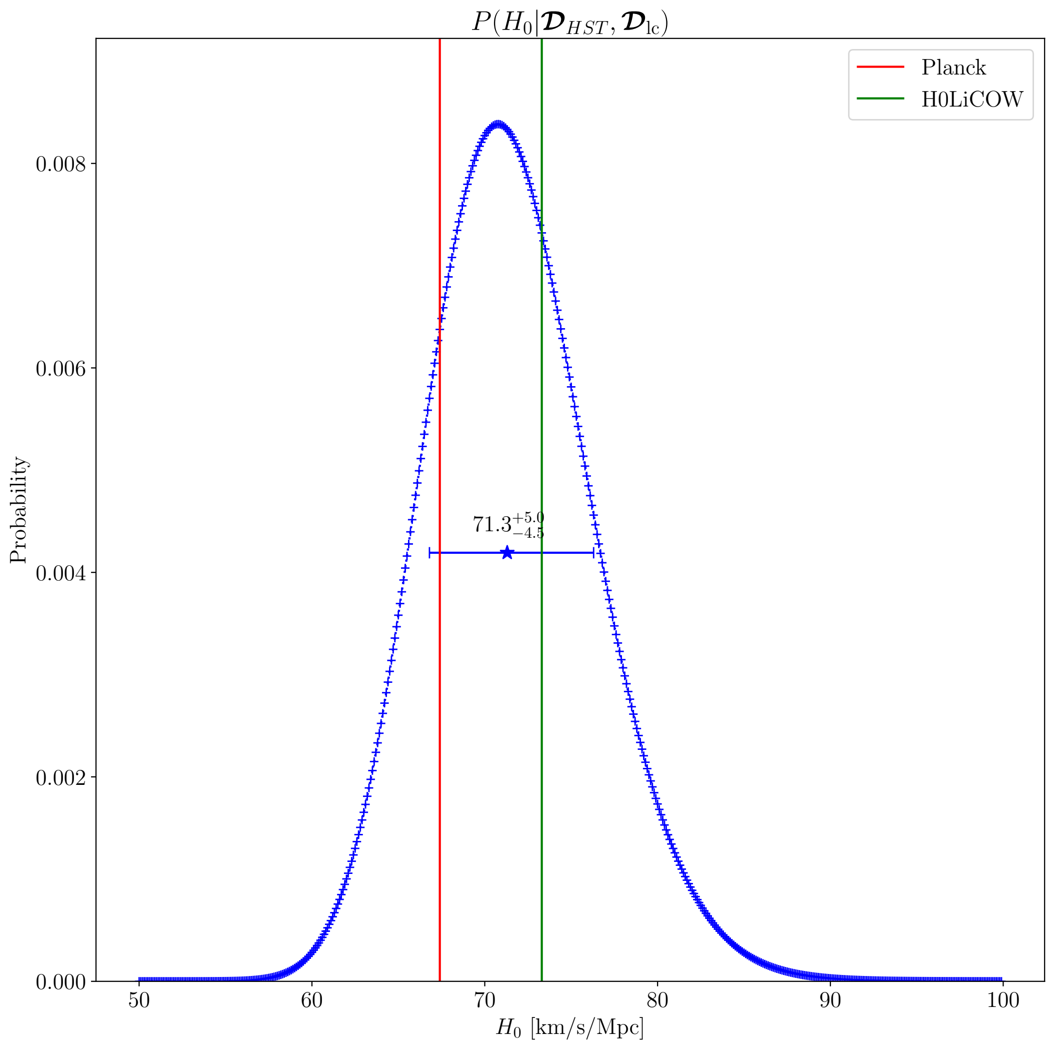

Figure 6.1: Resulting posterior probability $P(H_0|\boldsymbol{D}_{HST}, \boldsymbol{D}_{\mathrm{lc}})$ constrained by light-curves A, B, C, and D and mass model of J1433. I compare it with the results from the H0LiCOW Collaboration Wong et al. (2020) and the Planck Collaboration Aghanim et al. (2020).

which could be blinded from the beginning, or the finding of a way to verify every step of the analysis by blinding the required data. I will further discuss the result and the complete analysis in the following Chapter 7.

Chapter 7

Discussion and Conclusion

This work presents the first integral Time Delay Cosmographic analysis of the quadruply lensed QSO system J1433. It relies on two main datasets, the archival *HST* observations in five filters (F475X, F814W, F105W, F140W and F160W) and on the dedicated observational campaign carried out from the 2.1-meter optical telescope at the Wendelstein observatory in the *g'* filter. The analysis can be divided into three, mostly independent sections: the lens modelling and the constraining of the Fermat potential difference $\Delta\phi$ in Chapter 4, the lightcurves analysis and the constraining of the time delay Δt in Chapter 5 and the combination of the two results to constrain the Hubble constant H_0 in Chapter 6. This analysis followed a modular structure, thus each section is largely independent of the results of the others. This allowed for an iterative improvement of the analysis, where each part could be improved or modified independently without requiring the restructuring of the whole study.
More specifically, for the lens analysis, I followed a similar approach to what was developed by the TDCOSMO collaboration Millon et al. (2020b) using the public `lenstronomy` modelling tool. I employed all available *HST* filters to produce an isophotal light model of the lens galaxy (see Section 4.2.4), which could be subtracted a priori of the lens modelling and allowed to inform the likelihood of the mass distribution (see Section 4.3.5). I then used the optical filters to strictly constrain the positions of the QSO lensed images (see Section 4.3.4). The filter F105W was instead ignored due to its low SNR, low resolution and the absence of a visible lensed host galaxy. Then, I run two independent mass models, the two NIR filters F140W and F160W (see Section 4.4). These results were initially strongly in tension, leading me to the discovery of an excess of light appearing in F160W, which I indicate as the "contaminants" (see Section 4.4.1). Since it was impossible to definitively determine the nature of the source, I instead opted to mask it. The results of the modelling run with the updated mask were then in full agreement between the two NIR filters. I then described how I implemented the combination of the posterior in Section 4.5, following a Bayesian approach. This resulted in constraints on the Fermat potential difference with 3% precision, accounting for statistical uncertainty. Regarding the lightcurve analysis, I followed the collection of the data and handled the data reduction of the observations. These resulted in almost 300 nights over three years of observations, with a median PSF of 1.16". I oversaw the development of the pipeline for the lightcurve measurement described in Section 5.1, which I

employed to obtain the lightcurves shown in Figure 5.7. The resulting photometric precision for the QSO fluxes ranges from 1.6 % to 5.7 %. This result, besides being the cornerstone for the time delay analysis of this work, also proves the capabilities of the Wendelstein observatory and the 2.1-meter *WWFI*. These instruments are suitable for light-curve monitoring of multiple imaged QSO-galaxy lensing systems for time-delay cosmography and similar monitoring observational campaigns. I then analysed the obtained lightcurves with the public `PyCS3` python library (Tewes et al., 2013), fitting the lightcurve for time-delay, magnitude shift, microlensing perturbations and intrinsic variability of the source. I then considered a range of reasonable parametric fits and obtained for each an individual time delay constraint (see Section 5.3) and an estimation of the relative uncertainty (see Section 5.4). Instead of averaging the results or selecting only the best, I implemented the approach described in Millon et al. (2020a), through which a combined result is defined such that both the final uncertainty and the tension with the combined result are minimised. This yielded a time delay result with precision ranging from 8 % and 15 %. This scatter is due to the difference in the observed brightness of the images, which greatly affects the photometric precision, especially for image D.

Once I obtained the constraint for both $\Delta\phi$ and Δt, I adapted the Bayesian framework for the joint analysis described in Suyu et al. (2013) to the system at hand (see Section 6.1). The analysis yields a constraint on H_0 with a $\sim 6.7\%$ precision, accounting for random uncertainties. This result reaches precision comparable to similar studies, e.g. Chen et al. (2019). Additional results, which emerged as byproducts of the main analysis, were the colour profile of the main lens galaxy (see Section 4.6.3) and its mass-to-light ratio (see Section 4.6.2) as well as the study of the flux ratio anomaly in the Appendix A.

This study showed the importance of preliminary individual modelling of the single filter when reconstructing the mass profile of the lenses, underlined by the discovery of the "contaminants". Such excesses of light were not considered in the previous models of this system, and therefore, the results are severely in tension with those obtained in this work. Clearly, the technique of multifilter analysis is not to be fully discarded, but it should require an a priori study of the individual filters to gauge the information carried by each of them.

Another particular feature of this work is that the whole analysis has been carried out independently from larger collaborations, such as TDCOSMO. This is unusual in time delay cosmography due to the extended amount of different analyses required. This, therefore, allowed me the freedom to explore new approaches to the problems at hand, while necessarily requiring extensive work and time to complete. In particular, this prevented me from completing further improvements on the analysis, especially regarding the study of systematic errors. Firstly, the Mass-Sheet Degeneracy (MSD) has not been considered in this work, due to the lack of spectroscopic data which is necessary to constrain both the inner structure of the lens system and the effect of neighbouring structures (for details, refer to Section 2.3.3). Then the mass profile considered in the lens modelling was limited to the PEMD, i.e. the power-law profile. Similar studies can also consider a combined profile, i.e. a combination of a dark matter profile (e.g. a Navarro-Frenk-White profile, Navarro et al., 1997) and a baryonic profile (usually obtained by scaling the light profile with a constant mass-to-light ratio). Another source of systematic uncertainty would be the redshift of the perturber, which in this analysis is assumed to be at the same redshift as the main lens. While this is a reasonable assumption, due to a lack of spectroscopic data, I only obtained weak observational proof that this was true, as discussed in the Appendix B. Other possible future improvements in the lens modelling could encompass the use of iterative PSF fitting processes (Schmidt et al., 2022, as used in e.g.) and the use of multifilter modelling, which is now a viable option given the knowledge of the possible pitfalls, such as the presence of the "contaminants". Regarding the time delay analysis, it would be possible to follow a different approach to the photometric measurement by deconvolving the exposures and measuring the luminosities of the QSO images within small apertures. Furthermore, it would be possible to implement the analysis of the lightcurve with the second available method in `PyCS3`, which fits the lightcurves with a Gaussian process regression (Tewes et al., 2013).

Appendix A

Flux Ratio Anomaly

A further analysis, possible with the data and results obtained, is to investigate the presence of a flux ratio (FR) anomaly. This refers to the possible tension between the ratio of flux measured from two lensed images of the same source and the equivalent value predicted from the magnification of the lens modelling. More specifically, let's assume a constant (i.e. not variable over time) source with luminosity L in the background and a foreground transparent strong lens. We indicate the magnification map of the lens for the given source to be $\mu(\vec{\theta})$ and the time delay map to be $\tau(\vec{\theta})$. Then we should observe the lensed images i of the source at position θ_i to have the luminosity $L_\mathrm{i} = \mu(\vec{\theta}_\mathrm{i}) \cdot L$. Considering the two images A and B, their FR should then be $\mathrm{FR}_\mathrm{AB} = \frac{L_\mathrm{A}}{L_\mathrm{B}} = \frac{\mu(\vec{\theta}_\mathrm{A})}{\mu(\vec{\theta}_\mathrm{B})}$. In this case, however, the luminosity of the source (here, a QSO) varies over time, as it can be seen from the light-curves in Figure 5.7. We therefore must consider $L = L(t)$ and $L_\mathrm{i} = \mu(\vec{\theta}_\mathrm{i}) \cdot L(t+\tau(\vec{\theta}_\mathrm{i}))$, where now the luminosity is also affected by the time delay. Furthermore, each image could be temporarily microlensed by an intervening massive object, introduced in Section 2.3.5 and discussed in Section 5.3. This would further affect the FR, as $L_\mathrm{i}(t) = (\mu(\vec{\theta}_\mathrm{i}) + \delta\mu(\vec{\theta}, t)) \cdot L(t + \tau(\vec{\theta}_\mathrm{i}))$, where $\delta\mu(\vec{\theta}, t)$ is the microlensing effect at $\vec{\theta}$ and time t. Given these effects, the FR can not be measured directly from the *HST* images, as it would be biased by the time-delay and possibly by microlensing events. Instead, I can measure it from the resulting magnitude shift obtained from the time-delay analysis of the *WST* lightcurves. In this analysis, I take into account the time-delay and correct for it. Moreover, the analysis also takes into account the microlensing distortion; however, this is limited to the microlensing shifts whose time scale is shorter than the observational campaign. Otherwise, the effect would be fully degenerate with a constant magnitude shift of the lighcurve, as its effect would appear as a constant magnification of one image over the observation campaign. This would not affect the time-delay analysis but would bias the measured FR. Such events should become less likely to happen the longer the observation, as they would require a relatively stable alignment of microlens(es) along the line of sight.

In the time-delay analysis, I measure the relative magnitude shift between lighcurves as the average of the microlensing correction, being a polynomial or a spline, $\langle\mu\rangle$ added to the initial magnitude shift $\Delta\mathrm{mag}_\mathrm{in}$. The latter is obtained a priori by taking the difference of the weighted mean for the magnitude of the light-curves, as described in Section 5.3. As previously stated,

FR	LCs Analysis	Lens Model	Tension
FR_{AB}	1.28 ± 0.05	$1.54^{+0.01}_{-0.02}$	4.8
FR_{AC}	0.9 ± 0.1	0.78 ± 0.01	1.2
FR_{AD}	0.22 ± 0.05	0.41 ± 0.01	3.5

Table A.1: Comparison between FR obtained from lens modelling and light-curve analysis, and the corresponding tension τ (see equation 5.2).

the reason for the consideration of $\langle\mu\rangle$ in the magnitude shift is that a constant microlensing correction and a magnitude shift are indistinguishable in the lightcurves, thus, in the analysis, the latter is formally defined as a polynomial microlensing of degree 0. Moreover, when considering microlensing correction with a higher degree of freedom, the first-order shift $\langle\mu\rangle$ would also fit the magnitude shift. These magnitude shifts were computed in parallel with the time-delay, following the same procedure. The results to be combined were selected by taking only the groups that were combined for the time-delay analysis, i.e. the selected groups indicated by (S) shown in Figure 5.15. The combined result $\boldsymbol{\Delta mag}$ is shown in Figure A.1.

Note that, as for the time delays obtained in Chapter 5, it is not interesting to have the absolute magnification of the single lightcurve, but rather the relative one; following the convention of the current work, I took the lightcurve of image A as reference. The magnitude shifts $\Delta \mathrm{mag}_{\mathrm{Ai}}$ are then computed in the time-delay analysis as magnitude shifts between images i and A, where i=B, C, D: $\Delta \mathrm{mag}_{\mathrm{Ai}} = \mathrm{mag}_{\mathrm{i}} - \mathrm{mag}_{\mathrm{A}}$. This is than be converted as $\mathrm{FR}_{\mathrm{Ai}} = \frac{F_{\mathrm{i}}}{F_{\mathrm{A}}} = 10^{\frac{-2\cdot\Delta \mathrm{mag}_{\mathrm{Ai}}}{5}}$. The resulting FR are reported in Table A.1.

For what concerns the FR measured from the lens model, this can be obtained by computing the magnification μ at the image's position. This is by definition the ratio between the lensed and unlesed flux for a given image i : $|\mu_{\mathrm{i}}| = \frac{\mathrm{F_i}}{\mathrm{F_{i,\,unlensed}}}$; note that we are interested in the absolute value of μ, as the sign indicates the chirality of the lensed image, which is here negligible.

μ can then be computed in a similar way as for $\Delta\phi$: for each point of any given MCMC chain sampling $P(\boldsymbol{\nu}|\mathcal{D}_\rangle)$, the posterior of the lens parameters $\boldsymbol{\nu}$ for a given image $\mathcal{D}_i$, I can compute μ at the image positions and take their ratio with respect to image A, obtaining a distribution over $\boldsymbol{\mu}/\mu_{\mathrm{A}}$ which samples the posterior $P(\boldsymbol{\mu}/\mu_{\mathrm{A}}|\mathcal{D}_i)$. Following the same Bayesian procedure as described for $\boldsymbol{\Delta\phi}$ in Section 4.5 with the Equation 4.17, I combine the posterior obtained from the modelling of F140W and F160W, obtaining a sample of $P(\boldsymbol{\mu}/\mu_{\mathrm{A}}|\mathcal{D}_{\mathrm{F140W}}, \mathcal{D}_{\mathrm{F160W}})$. The resulting distribution is shown in Figure A.2. The comparison between the two results is shown in Table A.1. This result shows varying degrees of tension, all larger than 1σ. There are multiple plausible explanations for it. For what concerns the observed FR from the lightcurves, there might be two explanations: unaccounted absorption along the different lines of sight and very long microlensing. Both seem improbable, as the first would affect the colour of the QSO images, while the second option would require a very stable microlensing configuration, as previously mentioned. A more likely scenario would be the presence of substructures (e.g. subhalos, Dalal, Kochanek, 2002) of the lens which are not fitted by the smooth PEMD lens model, or more

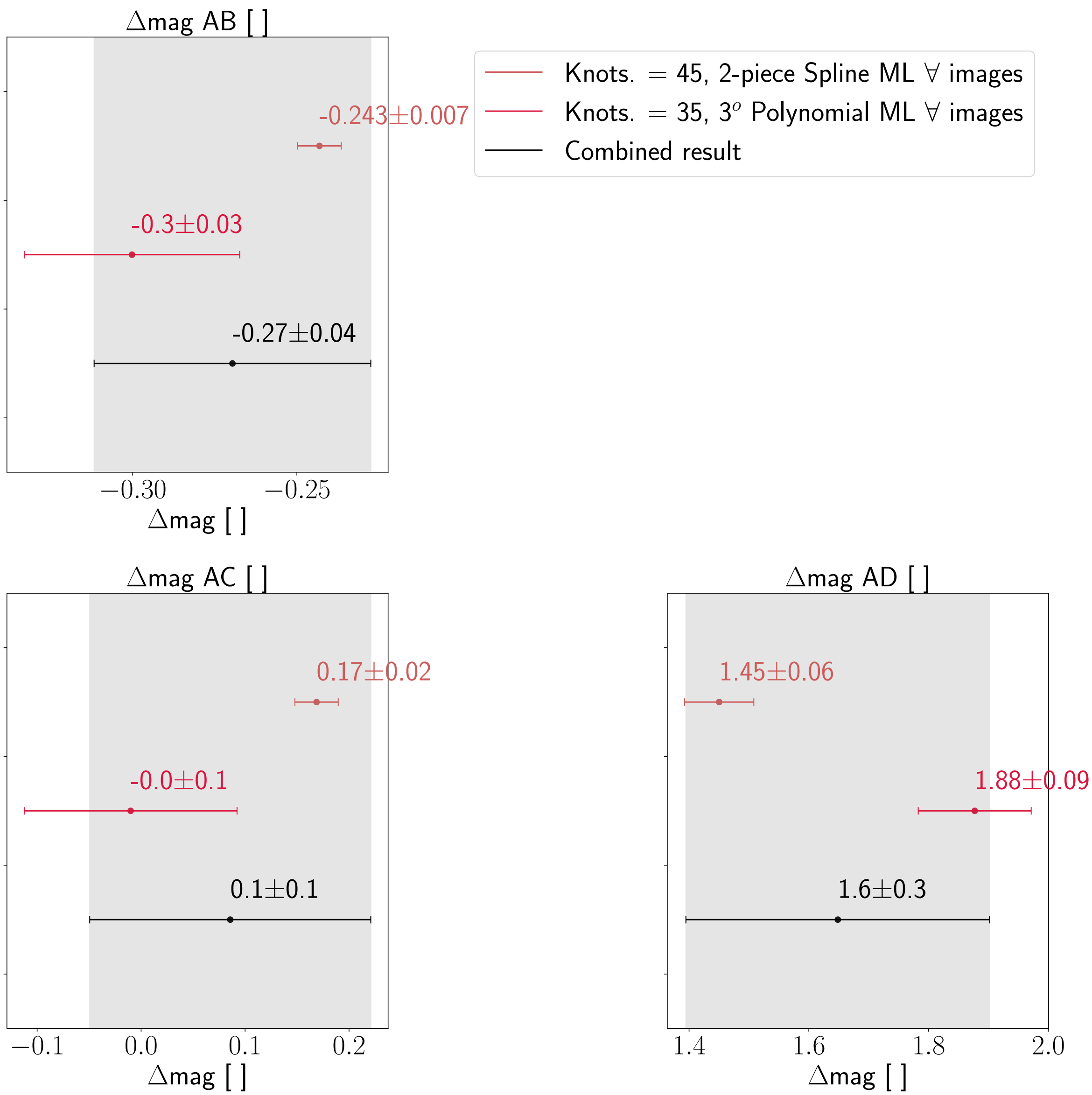

Figure A.1: Results of the magnitude shift measurements $\boldsymbol{\Delta mag}$ from the light-curves analysis (see Chapter 5), from which the flux ratio is computed as shown in Table A.1. "Knots." indicate the knot steps of the intrinsic spline, inversely proportional to the flexibility of the fitting. The ML refers to microlensing, which is either accounted for by a polynomial or with a 2-degree spline.

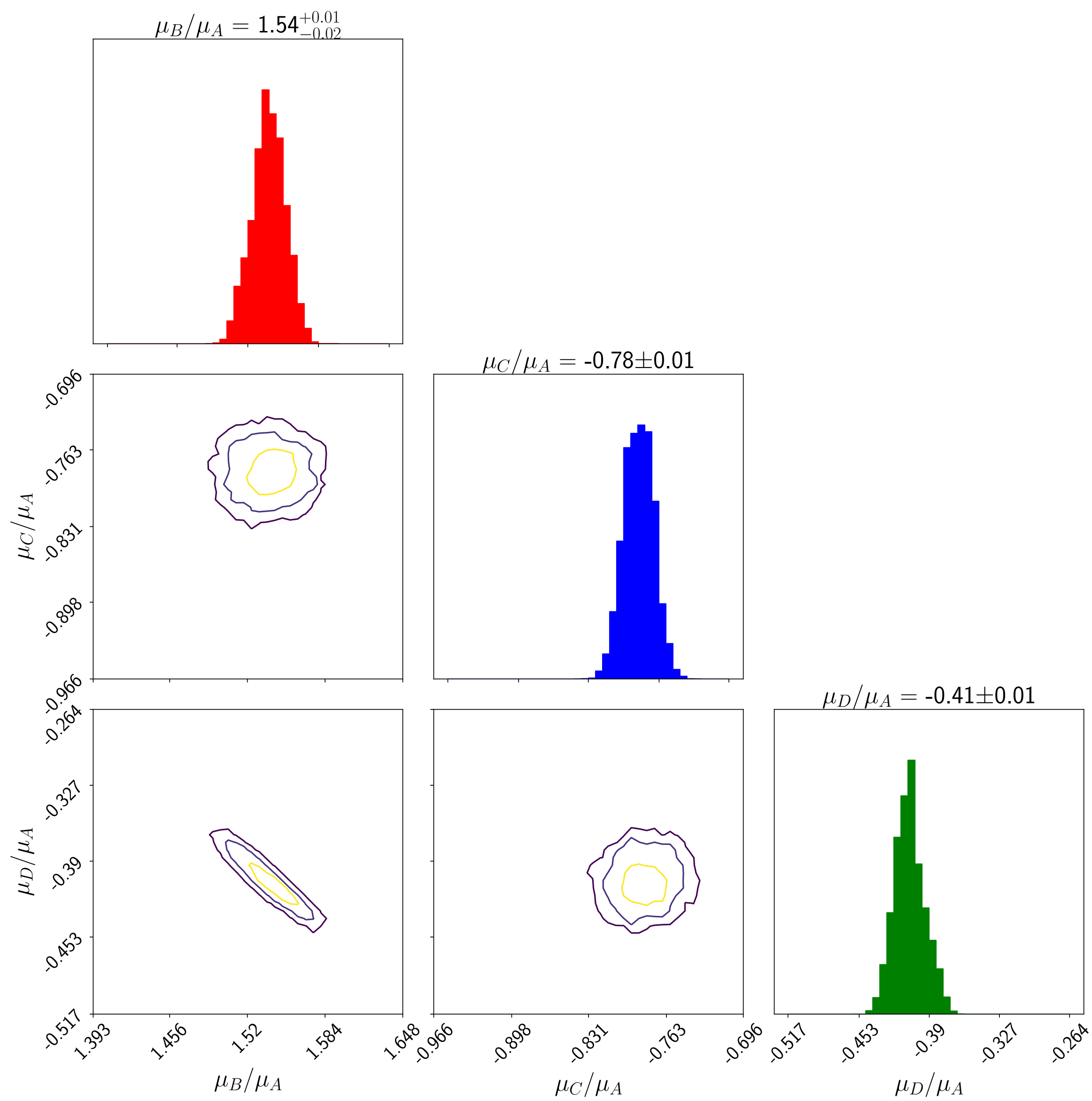

Figure A.2: Combined posterior of the magnification ratio $\boldsymbol{\mu}/\mu_{\rm A}$ at the positions of the images B, C, and D from the modelling in Section 4.3 with respect to image A. The 2D contour levels indicate the 68, 95 and 99.7 % confidence level, while the reported values indicate the median and 1-σ confidence interval.

generally, a lens mass profile which differs significantly from the PEMD model.

This latter case could significantly bias the present measurement of H_0 and should be taken into account in a future expansion of this work.

Appendix B

Perturber Photometric Redshift

One of the first assumptions of the lens modelling analysis carried out in this work was the assumption that the perturber galaxy, which can be seen near image C in the colour Figure 4.1, is located at the same redshift as the main lens galaxy. This assumption might be reasonable, and it is certainly appealing, for it simplifies the modelling; thus, it is a common assumption in the previous models of the system (see Shajib et al. (2019) and S22). However, without an estimation of the redshift, there is no basis for such an assumption. If this were proven to be wrong, it would strongly bias the lens model and thus the constraints on H_0. Therefore, I tried to constrain the redshift of the perturber, z_{pert}. No spectra were taken where the perturber is visible: a slitless spectroscopic image was taken from WFC3 from *HST*, but the perturber was too dim to be detected. Similarly,Mozumdar et al. (2023) reports a slit spectrum of the system (see Figure 2) where the perturber should be within the aperture but it is too dim to be detected.

I instead tried to constrain its redshift by photometric redshift estimation. For this, I used the public Fortran library LEPHARE, (Arnouts et al., 1999) and (Ilbert et al., 2006). I then measured the magnitude of the perturber in all 5 *HST* exposures available: F475X, F814W, F105W, F140W and F160W. Due to the blending of the perturber's light with nearby sources (image C, main lens and lensed arclets) I do not measure its brightness directly from the image. I instead use the corresponding Sérsic fit obtained from the lens modelling (see Chapter 4). However, I did not model F105W due to its low SNR. I therefore obtained a fit by running a very constrained model, where all lens parameters were fixed to the results obtained from the modelling of F140W. I chose this filter due to its similarity with F105W: similar wavelength, pixel resolution (although F140W had a higher resolution after drizzling) and similar exposure time (with respect to all other filters, see Table 4.2). I allowed for a small freedom in the QSO images' positions, which could vary within a range of $\pm 0.012''$, while still considering the same Gaussian likelihood as the other model of the NIR filters (see Section 4.3.4). The obtained model, which was only used to fit the light of the perturber, results in a $\chi^2_{\mathrm{red}} = 0.88$.

To measure the luminosity of the perturber, I took a circular aperture of 10" of radius, centred on the centre of the modelled perturber. The conversion to magnitude followed the same approach as described in Section 4.6.2. I considered the uncertainty by taking the standard deviation of the full residual map (i.e. the residual after the subtraction of the whole lens light profile, not only

F475X	F814W	F105W	F140W	F160W
23.0 ± 0.5	21.3 ± 0.2	22.1 ± 0.9	21.7 ± 0.3	20.7 ± 0.3

Table B.1: Apparent magnitudes of the perturber. These were obtained from the lens light model with a 10" aperture. The uncertainty is estimated from the standard deviation of the residual map.

the perturber). The resulting apparent magnitudes are shown in Table B.1.

These magnitudes were used as input for the photometric redshift estimation with LEPHARE. I consider the perturber to be a galaxy (i.e. not a star or a QSO), and marginalise over all available libraries of galaxy spectral energy distributions (SEDs). The output produces a maximum likelihood (ML) distribution for each of these SEDs, shown in Figure B.1. I here implemented a cutout, discarding the results of SEDs which had too large an uncertainty, as in $\sigma_z > 0.5$, as I deemed that these SEDs were not able to fit the data.

The resulting precision is nevertheless very poor, with an uncertainty of $\sigma_z \sim 0.3$, mostly due to the low brightness of the object and correspondingly large uncertainty, and partly due to the intrinsic lack of precision of the method. Thus, it can be said that the result is not in disagreement with the starting assumption, but without proving it with reasonable constraints. Interestingly, when repeating the same analysis for the main lens, I obtained a more constrained set of distributions, which, however, were not all centred on the spectroscopic redshift. This is shown in Figure B.2. This is a further indication that this method, while being the only option available, remains far from accurate.

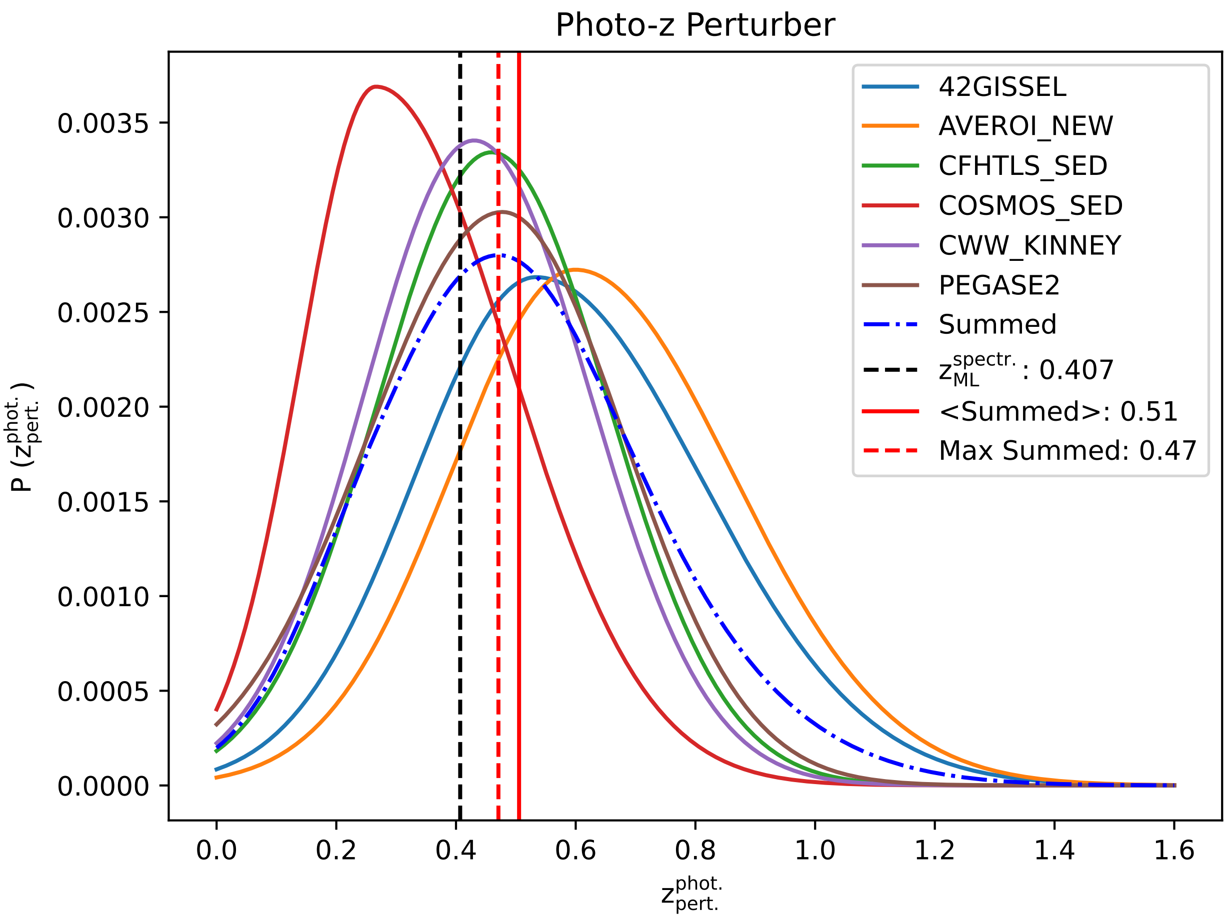

Figure B.1: Maximum likelihood posterior distributions for the photo-z estimate of the perturber for various SED libraries. The black dotted line indicates the Main Lens (ML) redshift known from spectroscopy. The blue dotted distribution is the normalised sum of all other posterior distributions, and the red vertical line indicates its expected value, while the dotted vertical line indicates the most likely value.

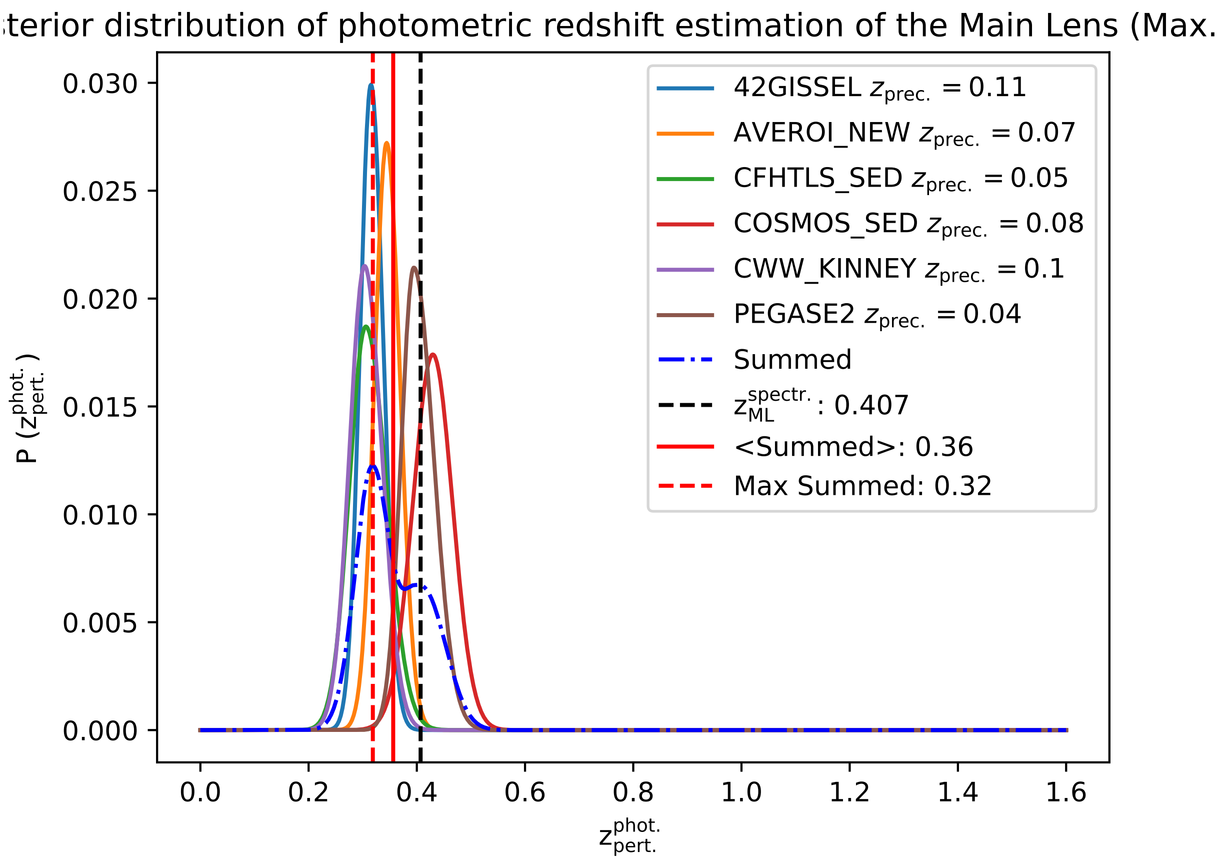

Figure B.2: Maximum likelihood posterior distributions for the photo-z estimate of the main lens galaxy for various SED libraries. The black dotted line indicates the Main Lens (ML) redshift known from spectroscopy. As for Figure B.1, the blue distribution is the result of the sum and normalisation of all other distributions. The solid/dashed red line indicates the expected/maximum value of the summed distribution.

Bibliography

Abdalla Elcio, Abellán Guillermo Franco, Aboubrahim Amin, Agnello Adriano, Akarsu Özgür, Akrami Yashar, Alestas George, Aloni Daniel, Amendola Luca, Anchordoqui Luis A., Anderson Richard I., Arendse Nikki, Asgari Marika, Ballardini Mario, Barger Vernon, Basilakos Spyros, Batista Ronaldo C., Battistelli Elia S., Battye Richard, Benetti Micol, Benisty David, Berlin Asher, de Bernardis Paolo, Berti Emanuele, Bidenko Bohdan, Birrer Simon, Blakeslee John P., Boddy Kimberly K., Bom Clecio R., Bonilla Alexander, Borghi Nicola, Bouchet François R., Braglia Matteo, Buchert Thomas, Buckley-Geer Elizabeth, Calabrese Erminia, Caldwell Robert R., Camarena David, Capozziello Salvatore, Casertano Stefano, Chen Geoff C. F., Chluba Jens, Chen Angela, Chen Hsin-Yu, Chudaykin Anton, Cicoli Michele, Copi Craig J., Courbin Fred, Cyr-Racine Francis-Yan, Czerny Bożena, Dainotti Maria, D'Amico Guido, Davis Anne-Christine, de Cruz Pérez Javier, de Haro Jaume, Delabrouille Jacques, Denton Peter B., Dhawan Suhail, Dienes Keith R., Di Valentino Eleonora, Du Pu, Eckert Dominique, Escamilla-Rivera Celia, Ferté Agnès, Finelli Fabio, Fosalba Pablo, Freedman Wendy L., Frusciante Noemi, Gaztañaga Enrique, Giarè William, Giusarma Elena, Gómez-Valent Adrià, Handley Will, Harrison Ian, Hart Luke, Hazra Dhiraj Kumar, Heavens Alan, Heinesen Asta, Hildebrandt Hendrik, Hill J. Colin, Hogg Natalie B., Holz Daniel E., Hooper Deanna C., Hosseininejad Nikoo, Huterer Dragan, Ishak Mustapha, Ivanov Mikhail M., Jaffe Andrew H., Jang In Sung, Jedamzik Karsten, Jimenez Raul, Joseph Melissa, Joudaki Shahab, Kamionkowski Marc, Karwal Tanvi, Kazantzidis Lavrentios, Keeley Ryan E., Klasen Michael, Komatsu Eiichiro, Koopmans Léon V. E., Kumar Suresh, Lamagna Luca, Lazkoz Ruth, Lee Chung-Chi, Lesgourgues Julien, Levi Said Jackson, Lewis Tiffany R., L'Huillier Benjamin, Lucca Matteo, Maartens Roy, Macri Lucas M., Marfatia Danny, Marra Valerio, Martins Carlos J. A. P., Masi Silvia, Matarrese Sabino, Mazumdar Arindam, Melchiorri Alessandro, Mena Olga, Mersini-Houghton Laura, Mertens James, Milaković Dinko, Minami Yuto, Miranda Vivian, Moreno-Pulido Cristian, Moresco Michele, Mota David F., Mottola Emil, Mozzon Simone, Muir Jessica, Mukherjee Ankan, Mukherjee Suvodip, Naselsky Pavel, Nath Pran, Nesseris Savvas, Niedermann Florian, Notari Alessio, Nunes Rafael C., Ó Colgáin Eoin, Owens Kayla A., Özülker Emre, Pace Francesco, Paliathanasis Andronikos, Palmese Antonella, Pan Supriya, Paoletti Daniela, Perez Bergliaffa Santiago E., Perivolaropoulos Leandros, Pesce Dominic W., Pettorino Valeria, Philcox Oliver H. E., Pogosian Levon, Poulin

Vivian, Poulot Gaspard, Raveri Marco, Reid Mark J., Renzi Fabrizio, Riess Adam G., Sabla Vivian I., Salucci Paolo, Salzano Vincenzo, Saridakis Emmanuel N., Sathyaprakash Bangalore S., Schmaltz Martin, Schöneberg Nils, Scolnic Dan, Sen Anjan A., Sehgal Neelima, Shafieloo Arman, Sheikh-Jabbari M. M., Silk Joseph, Silvestri Alessandra, Skara Foteini, Sloth Martin S., Soares-Santos Marcelle, Solà Peracaula Joan, Songsheng Yu-Yang, Soriano Jorge F., Staicova Denitsa, Starkman Glenn D., Szapudi István, Teixeira Elsa M., Thomas Brooks, Treu Tommaso, Trott Emery, van de Bruck Carsten, Vazquez J. Alberto, Verde Licia, Visinelli Luca, Wang Deng, Wang Jian-Min, Wang Shao-Jiang, Watkins Richard, Watson Scott, Webb John K., Weiner Neal, Weltman Amanda, Witte Samuel J., Wojtak Rados law, Yadav Anil Kumar, Yang Weiqiang, Zhao Gong-Bo, Zumalacárregui Miguel. Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies // Journal of High Energy Astrophysics. VI 2022. 34. 49–211.

Addison Graeme E, Hinshaw Gary, Halpern Mark. Cosmological constraints from baryon acoustic oscillations and clustering of large-scale structure // Monthly Notices of the Royal Astronomical Society. 2013. 436, 2. 1674–1683.

Addison-Weasley Pearson. Structure of the quasar's accretion disk. 2019. Accessed: 2024-05-31.

Aghanim Nabila, Akrami Yashar, Ashdown Mark, Aumont J, Baccigalupi C, Ballardini M, Banday AJ, Barreiro RB, Bartolo N, Basak S, others . Planck 2018 results-VI. Cosmological parameters // Astronomy & Astrophysics. 2020. 641. A6.

Agnello Adriano, Grillo Claudio, Jones Tucker, Treu Tommaso, Bonamigo Mario, Suyu Sherry H. Discovery and first models of the quadruply lensed quasar SDSS J1433+6007 // . III 2018. 474, 3. 3391–3396.

Amendola Luca, Tsujikawa Shinji. Dark energy: theory and observations. 2010.

Amvrosiadis Aristeidis, Lange Samuel, Nightingale James, He Qiuhan, Frenk Carlos S, Oman Kyle A, Smail Ian, Swinbank Mark A, Fragkoudi Francesca, Gadotti Dimitri A, others . The onset of bar formation in a massive galaxy at $z\backslash sim$ 3.8 $//arXivpreprintarXiv$: 2404.01918.2024.

Anchordoqui Luis A., Di Valentino Eleonora, Pan Supriya, Yang Weiqiang. Dissecting the H0 and S8 tensions with Planck + BAO + supernova type Ia in multi-parameter cosmologies // Journal of High Energy Astrophysics. 2021. 32. 28–64.

Arkani-Hamed Nima, Hall Lawrence J., Kolda Christopher, Murayama Hitoshi. New Perspective on Cosmic Coincidence Problems // Phys. Rev. Lett. Nov 2000. 85. 4434–4437.

Arnouts Stephane, Cristiani Stefano, Moscardini Lauro, Matarrese Sabino, Lucchin Francesco, Fontana Adriano, Giallongo Emanuele. Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North // Monthly Notices of the Royal Astronomical Society. 1999. 310, 2. 540–556.

Astrophysics Centre for, Technology Supercomputing Swinburne University of. GIMLET: The GELUMPH Interactive Microlensing Lightcurve Extraction Tool. 2020. Accessed: 2024-05-31.

Auger MW, Treu T, Gavazzi R, Bolton AS, Koopmans LVE, Marshall PJ. Dark matter contraction and the stellar content of massive early-type galaxies: disfavoring "light" initial mass functions // The Astrophysical Journal Letters. 2010. 721, 2. L163.

Ballard Daniel J, Enzi Wolfgang JR, Collett Thomas E, Turner Hannah C, Smith Russell J. Gravitational imaging through a triple source plane lens: revisiting the ΛCDM–defying dark subhalo in SDSSJ0946+ 1006 // Monthly Notices of the Royal Astronomical Society. 2024. stae514.

Beckmann Volker, Shrader Chris. Active galactic nuclei. 2012.

Beichman Charles A, Rieke Marcia, Eisenstein Daniel, Greene Thomas P, Krist John, McCarthy Don, Meyer Michael, Stansberry John. Science opportunities with the near-IR camera (NIRCam) on the James Webb Space Telescope (JWST) // Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. 8442. 2012. 973–983.

Bertin Emmanuel, Arnouts Stephane. SExtractor: Software for source extraction // Astronomy and astrophysics supplement series. 1996. 117, 2. 393–404.

Bertone Gianfranco, Hooper Dan. History of dark matter // Rev. Mod. Phys. Oct 2018. 90. 045002.

Birrer S, Shajib AJ, Galan A, Millon M, Treu T, Agnello A, Auger M, Chen GC-F, Christensen L, Collett T, others . TDCOSMO-IV. Hierarchical time-delay cosmography–joint inference of the Hubble constant and galaxy density profiles // Astronomy & Astrophysics. 2020. 643. A165.

Birrer Simon, Amara Adam, Refregier Alexandre. Gravitational lens modeling with basis sets // The Astrophysical Journal. 2015. 813, 2. 102.

Birrer Simon, Treu Tommaso. Astrometric requirements for strong lensing time-delay cosmography // Monthly Notices of the Royal Astronomical Society. 08 2019. 489, 2. 2097–2103.

Blandford Roger, Meier David, Readhead Anthony. Relativistic jets from active galactic nuclei // Annual Review of Astronomy and Astrophysics. 2019. 57, 1. 467–509.

Bolton Adam S, Burles Scott, Koopmans Léon VE, Treu Tommaso, Gavazzi Raphaël, Moustakas Leonidas A, Wayth Randall, Schlegel David J. The sloan lens ACS survey. V. The full ACS strong-lens sample // The Astrophysical Journal. 2008. 682, 2. 964.

Bolton Adam S, Burles Scott, Koopmans Léon VE, Treu Tommaso, Moustakas Leonidas A. The Sloan Lens ACS Survey. I. A large spectroscopically selected sample of massive early-type lens galaxies // The Astrophysical Journal. 2006. 638, 2. 703.

Bonvin V, Millon M, Chan JH-H, Courbin F, Rusu CE, Sluse Dominique, Suyu SH, Wong KC, Fassnacht CD, Marshall PJ, others . COSMOGRAIL-XVIII. time delays of the quadruply lensed quasar WFI2033- 4723 // Astronomy & Astrophysics. 2019. 629. A97.

Born Max. Quantenmechanik der stoßvorgänge // Zeitschrift für physik. 1926. 38, 11. 803–827.

Bradley Larry, Sipőcz Brigitta, Robitaille Thomas, Tollerud Erik, Vinícius Zé, Deil Christoph, Barbary Kyle, Wilson Tom J, Busko Ivo, Donath Axel, Günther Hans Moritz, Cara Mihai, Lim P. L., Meßlinger Sebastian, Conseil Simon, Bostroem Azalee, Droettboom Michael, Bray E. M., Bratholm Lars Andersen, Barentsen Geert, Craig Matt, Rathi Shivangee, Pascual Sergio, Perren Gabriel, Georgiev Iskren Y., Val-Borro Miguel de, Kerzendorf Wolfgang, Bach Yoonsoo P., Quint Bruno, Souchereau Harrison. astropy/photutils: 1.5.0. VII 2022.

Burbidge Geoffrey. The quasi-stellar objects // American Scientist. 1967. 55, 3. 282–295.

Burgess CP. The cosmological constant problem: why it's hard to get dark energy from micro-physics // 100e Ecole d'Ete de Physique: Post-Planck Cosmology. 2015. 149–197.

Cackett Edward M., Horne Keith, Winkler Hartmut. Testing thermal reprocessing in active galactic nuclei accretion discs // . IX 2007. 380, 2. 669–682.

Caminha GB, Suyu SH, Grillo C, Rosati P. Galaxy cluster strong lensing cosmography-Cosmological constraints from a sample of regular galaxy clusters // Astronomy & Astrophysics. 2022. 657. A83.

Cañameras R, Schuldt S, Shu Y, Suyu SH, Taubenberger S, Meinhardt T, Leal-Taixé L, Chao DC-Y, Inoue KT, Jaelani AT, others . HOLISMOKES-VI. New galaxy-scale strong lens candidates from the HSC-SSP imaging survey // Astronomy & Astrophysics. 2021. 653. L6.

Capozziello Salvatore, Francaviglia Mauro. Extended theories of gravity and their cosmological and astrophysical applications // General Relativity and Gravitation. 2008. 40. 357–420.

Carroll Sean M. The cosmological constant // Living reviews in relativity. 2001. 4, 1. 1–56.

Carroll Sean M. Spacetime and geometry. 2019.

Chen Geoff CF, Fassnacht Christopher D, Suyu Sherry H, Rusu Cristian E, Chan James HH, Wong Kenneth C, Auger Matthew W, Hilbert Stefan, Bonvin Vivien, Birrer Simon, others . A SHARP view of H0LiCOW: H0 from three time-delay gravitational lens systems with adaptive optics imaging // Monthly Notices of the Royal Astronomical Society. 2019. 490, 2. 1743–1773.

Chevallier Michel, Polarski David. Accelerating universes with scaling dark matter // International Journal of Modern Physics D. 2001. 10, 02. 213–223.

Chilingarian Igor V., Melchior Anne-Laure, Zolotukhin Ivan Yu. Analytical approximations of K-corrections in optical and near-infrared bands // . VII 2010. 405, 3. 1409–1420.

Chilingarian Igor V., Zolotukhin Ivan Yu. A universal ultraviolet-optical colour-colour-magnitude relation of galaxies // . I 2012. 419, 2. 1727–1739.

Colgate Stirling A. Supernovae as a standard candle for cosmology // Astrophysical Journal, Part 1, vol. 232, Sept. 1, 1979, p. 404-408. Research supported by the US Department of Energy and NSF. 1979. 232. 404–408.

Collett Thomas E. The population of galaxy–galaxy strong lenses in forthcoming optical imaging surveys // The Astrophysical Journal. 2015. 811, 1. 20.

Collett Thomas E, Oldham Lindsay J, Smith Russell J, Auger Matthew W, Westfall Kyle B, Bacon David, Nichol Robert C, Masters Karen L, Koyama Kazuya, Bosch Remco van den. A precise extragalactic test of General Relativity // Science. 2018. 360, 6395. 1342–1346.

Courbin Frederic, Bonvin V, Buckley-Geer E, Fassnacht CD, Frieman J, Lin H, Marshall PJ, Suyu SH, Treu T, Anguita T, others . COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses-XVI. Time delays for the quadruply imaged quasar DES J0408- 5354 with high-cadence photometric monitoring // Astronomy & Astrophysics. 2018. 609. A71.

Dalal N, Kochanek CS. Direct detection of cold dark matter substructure // The Astrophysical Journal. 2002. 572, 1. 25.

Dawson Kyle S, Schlegel David J, Ahn Christopher P, Anderson Scott F, Aubourg Éric, Bailey Stephen, Barkhouser Robert H, Bautista Julian E, Beifiori Alessandra, Berlind Andreas A, others . The baryon oscillation spectroscopic survey of SDSS-III // The Astronomical Journal. 2012. 145, 1. 10.

De Jaeger Thomas, Galbany L, González-Gaitán S, Kessler R, Filippenko AV, Förster F, Hamuy M, Brown PJ, Davis TM, Gutiérrez CP, others . Studying Type II supernovae as cosmological standard candles using the Dark Energy Survey // Monthly Notices of the Royal Astronomical Society. 2020. 495, 4. 4860–4892.

Despali Giulia, Sparre Martin, Vegetti Simona, Vogelsberger Mark, Zavala Jesús, Marinacci Federico. The interplay of self-interacting dark matter and baryons in shaping the halo evolution // Monthly Notices of the Royal Astronomical Society. 01 2019. 484, 4. 4563–4573.

Dhanasingham Birendra, Cyr-Racine Francis-Yan, Peter Annika HG, Benson Andrew, Gilman Daniel. Interlopers speak out: studying the dark universe using small-scale lensing anisotropies // Monthly Notices of the Royal Astronomical Society. 2023. 518, 4. 5843–5861.

Di Valentino Eleonora, Mena Olga, Pan Supriya, Visinelli Luca, Yang Weiqiang, Melchiorri Alessandro, Mota David F, Riess Adam G, Silk Joseph. In the realm of the Hubble tension—a review of solutions // Classical and Quantum Gravity. 2021a. 38, 15. 153001.

Di Valentino Eleonora, Mena Olga, Pan Supriya, Visinelli Luca, Yang Weiqiang, Melchiorri Alessandro, Mota David F, Riess Adam G, Silk Joseph. In the realm of the Hubble tension—a review of solutions // Classical and Quantum Gravity. 2021b. 38, 15. 153001.

Dodelson Scott, Schmidt Fabian. Modern cosmology. 2020.

Dutton Aaron A, Brewer Brendon J, Marshall Philip J, Auger Matthew W, Treu Tommaso, Koo David C, Bolton Adam S, Holden Bradford P, Koopmans Leon VE. The SWELLS survey–II. Breaking the disc–halo degeneracy in the spiral galaxy gravitational lens SDSS J2141- 0001 // Monthly Notices of the Royal Astronomical Society. 2011. 417, 3. 1621–1642.

Einstein Albert. Die Grundlagen der allgemeinen // Relativitats- teorie, Annale der Physic. 1916. 49. 769.

Einstein Albert. Lens-like action of a star by the deviation of light in the gravitational field // Science. 1936. 84, 2188. 506–507.

Eisenstein Daniel J, Zehavi Idit, Hogg David W, Scoccimarro Roman, Blanton Michael R, Nichol Robert C, Scranton Ryan, Seo Hee-Jong, Tegmark Max, Zheng Zheng, others . Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies // The Astrophysical Journal. 2005. 633, 2. 560.

Ertl S, Schuldt S, Suyu SH, Schmidt T, Treu T, Birrer S, Shajib AJ, Sluse Dominique. TDCOSMO. X. Automated modeling of nine strongly lensed quasars and comparison between lens-modeling software // Astronomy and Astrophysics. 2023. 672.

Falco EE, Gorenstein MV, Shapiro II. On model-dependent bounds on H (0) from gravitational images Application of Q0957+ 561A, B // The Astrophysical Journal. 1985. 289. L1–L4.

Flewelling HA, Magnier EA, Chambers KC, Heasley JN, Holmberg C, Huber ME, Sweeney W, Waters CZ, Calamida A, Casertano S, others . The Pan-STARRS1 database and data products // arXiv preprint arXiv:1612.05243. 2016.

Foreman-Mackey Daniel, Hogg David W, Lang Dustin, Goodman Jonathan. emcee: the MCMC hammer // Publications of the Astronomical Society of the Pacific. 2013. 125, 925. 306.

Frank Juhan, King Andrew R, Raine Derek. Accretion power in astrophysics. 2002.

Freedman Wendy L, Madore Barry F, Gibson Brad K, Ferrarese Laura, Kelson Daniel D, Sakai Shoko, Mould Jeremy R, Kennicutt Jr Robert C, Ford Holland C, Graham John A, others . Final results from the Hubble Space Telescope key project to measure the Hubble constant // The Astrophysical Journal. 2001. 553, 1. 47.

Freeman Kenneth C. On the disks of spiral and S0 galaxies // Astrophysical Journal, vol. 160, p. 811. 1970. 160. 811.

Fruchter A. S., Hook R. N. Drizzle: A Method for the Linear Reconstruction of Undersampled Images // Publications of the Astronomical Society of the Pacific. feb 2002. 114, 792. 144.

Fruchter A. S., et al. . BetaDrizzle: A Redesign of the MultiDrizzle Package // 2010 Space Telescope Science Institute Calibration Workshop. VII 2010. 382–387.

Fruchter Andrew, Hook Richard N. Novel image reconstruction method applied to deep Hubble space telescope images // Applications of Digital Image Processing XX. 3164. 1997. 120–125.

Frye Brenda L, Pascale Massimo, Pierel Justin, Chen Wenlei, Foo Nicholas, Leimbach Reagen, Garuda Nikhil, Cohen Seth H, Kamieneski Patrick S, Windhorst Rogier A, others . The JWST discovery of the triply imaged type Ia "Supernova H0pe" and observations of the galaxy cluster PLCK G165. 7+ 67.0 // The Astrophysical Journal. 2024. 961, 2. 171.

Gardner Jonathan P, Mather John C, Clampin Mark, Doyon Rene, Greenhouse Matthew A, Hammel Heidi B, Hutchings John B, Jakobsen Peter, Lilly Simon J, Long Knox S, others . The james webb space telescope // Space Science Reviews. 2006. 123. 485–606.

Gelman Andrew, Rubin Donald B. A single series from the Gibbs sampler provides a false sense of security // Bayesian statistics. 1992. 4, 1. 625–631.

Ghirardini V, Bulbul E, Artis E, Clerc N, Garrel C, Grandis S, Kluge M, Liu A, Bahar YE, Balzer F, others . The SRG/eROSITA All-Sky Survey: Cosmology Constraints from Cluster Abundances in the Western Galactic Hemisphere // arXiv preprint arXiv:2402.08458. 2024.

Ghosh Subhajit, Khatri Rishi, Roy Tuhin S. Dark neutrino interactions make gravitational waves blue // Physical Review D. 2018. 97, 6. 063529.

Gilman Daniel, Birrer Simon, Nierenberg Anna, Treu Tommaso, Du Xiaolong, Benson Andrew. Warm dark matter chills out: constraints on the halo mass function and the free-streaming length of dark matter with eight quadruple-image strong gravitational lenses // Monthly Notices of the Royal Astronomical Society. 12 2019. 491, 4. 6077–6101.

Gilman Daniel, Bovy Jo, Treu Tommaso, Nierenberg Anna, Birrer Simon, Benson Andrew, Sameie Omid. Strong lensing signatures of self-interacting dark matter in low-mass haloes // Monthly Notices of the Royal Astronomical Society. 08 2021. 507, 2. 2432–2447.

Giovanelli Riccardo, Haynes Martha P, Costa Luiz N da, Freudling Wolfram, Salzer John J, Wegner Gary. The Tully-Fisher Relation and H0 // The Astrophysical Journal. 1997. 477, 1. L1.

Goodman Jonathan, Weare Jonathan. Ensemble samplers with affine invariance // Communications in applied mathematics and computational science. 2010. 5, 1. 65–80.

Greene Zach S, Suyu Sherry H, Treu Tommaso, Hilbert Stefan, Auger Matthew W, Collett Thomas E, Marshall Philip J, Fassnacht Christopher D, Blandford Roger D, Bradač Maruša, others . Improving the Precision of Time-delay Cosmography with Observations of Galaxies along the Line of Sight // The Astrophysical Journal. 2013. 768, 1. 39.

Hamuy Mario, Phillips MM, Suntzeff Nicholas B, Schommer Robert A, Maza Jose, Smith RC, Lira P, Aviles R. The morphology of type Ia supernovae light curves // arXiv preprint astro-ph/9609063. 1996.

Hoffmann Samantha L., Mack Jason. The DrizzlePac Handbook. 2021. Version 2.

Hogg Natalie B, Fleury Pierre, Larena Julien, Martinelli Matteo. Measuring line-of-sight shear with Einstein rings: a proof of concept // Monthly Notices of the Royal Astronomical Society. 2023. 520, 4. 5982–6000.

Hopp Ulrich, Bender Ralf, Grupp Frank, Goessl Claus, Lang-Bardl Florian, Mitsch Wolfgang, Riffeser Arno, Ageorges Nancy. Commissioning and science verification of the 2m-Fraunhofer Wendelstein Telescope // Ground-based and Airborne Telescopes V. 9145. Jul 2014. 91452D. (Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series).

Howell Steve B, Koehn Bruce, Bowell Edward, Hoffman Mark. Detection and measurement of poorly sampled point sources imaged with 2-D array // Astronomical Journal v. 112, p. 1302. 1996. 112. 1302.

Hsiao Tiger Yu-Yang, Coe Dan, Larson Rebecca L, Jung Intae, Mingozzi Matilde, Dayal Pratika, Kumari Nimisha, Kokorev Vasily, Vikaeus Anton, Brammer Gabriel, others . JWST NIRSpec spectroscopy of the triply-lensed $z = 10.17$ galaxy MACS0647 − JD // arXiv preprint arXiv:2305.03042. 2023.

Hubble Edwin. A relation between distance and radial velocity among extra-galactic nebulae // Proceedings of the national academy of sciences. 1929. 15, 3. 168–173.

Ilbert Olivier, Arnouts S, Mccracken Henry J, Bolzonella M, Bertin Emmanuel, Le Fèvre Olivier, Mellier Yannick, Zamorani G, Pello R, Iovino Angela, others . Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey // Astronomy & Astrophysics. 2006. 457, 3. 841–856.

Institute Space Telescope Science. UVIS Encircled Energy. 01/31/2024. Accessed: 30/07/2024.

Jedrzejewski Robert I. CCD surface photometry of elliptical galaxies–I. Observations, reduction and results // Monthly Notices of the Royal Astronomical Society. 1987. 226, 4. 747–768.

Kabsch Wolfgang. A solution for the best rotation to relate two sets of vectors // Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography. 1976. 32, 5. 922–923.

Karwal Tanvi, Kamionkowski Marc. Dark energy at early times, the Hubble parameter, and the string axiverse // Physical Review D. 2016. 94, 10. 103523.

Kassiola Aggeliki, Kovner Israel. Elliptic mass distributions versus elliptic potentials in gravitational lenses // Astrophysical Journal v. 417, p. 450. 1993. 417. 450.

Kawaguchi Toshihiro, Mineshige Shin, Machida Mami, Matsumoto Ryoji, Shibata Kazunari. Temporal 1/f α fluctuations from fractal magnetic fields in black-hole accretion flow // Publications of the Astronomical Society of Japan. 2000. 52, 1. L1–L4.

Keeton Charles R. A catalog of mass models for gravitational lensing // arXiv preprint astro-ph/0102341. 2001.

Kelly PL, Brammer G, Selsing J, Foley RJ, Hjorth J, Rodney SA, Christensen L, Strolger L-G, Filippenko AV, Treu T, others . SN Refsdal: classification as a luminous and blue SN 1987A-like type II supernova // The Astrophysical Journal. 2016. 831, 2. 205.

Kennedy J., Eberhart R. Particle swarm optimization // Proceedings of ICNN'95 - International Conference on Neural Networks. 4. 1995. 1942–1948 vol.4.

Kluge Matthias. Structure of brightest cluster galaxies and intracluster light. I 2020.

Kluge Matthias, Bender Ralf. Minor Mergers Are Not Enough: The Importance of Major Mergers during Brightest Cluster Galaxy Assembly // . VIII 2023. 267, 2. 41.

Kluge Matthias, Neureiter B, Riffeser A, Bender R, Goessl C, Hopp U, Schmidt M, Ries C, Brosch N. Structure of brightest cluster galaxies and intracluster light // The Astrophysical Journal Supplement Series. 2020. 247, 2. 43.

Kluge Matthias, Remus Rhea-Silvia, Babyk Iurii V., Forbes Duncan A., Dolfi Arianna. A trail of the invisible: blue globular clusters trace the radial density distribution of the dark matter - case study of NGC 4278 // . VI 2023. 521, 4. 4852–4862.

Knox L., Millea M. Hubble constant hunter's guide // Phys. Rev. D. Feb 2020. 101. 043533.

Kochanek Christopher S. Gravitational lensing limits on cold dark matter and its variants // Arxiv preprint astro-ph/9411082. 1994.

Kosyra Ralf, Gössl Claus, Hopp Ulrich, Lang-Bardl Florian, Riffeser Arno, Bender Ralf, Seitz Stella. The 64 Mpixel wide field imager for the Wendelstein 2m telescope: design and calibration // Experimental Astronomy. 2014. 38. 213–248.

Laplace Pierre Simon marquis de. Exposition du système du monde. 1. 1813.

Laroche Alexander, Gilman Daniel, Li Xinyu, Bovy Jo, Du Xiaolong. Quantum fluctuations masquerade as haloes: bounds on ultra-light dark matter from quadruply imaged quasars // Monthly Notices of the Royal Astronomical Society. 09 2022. 517, 2. 1867–1883.

Leavitt Henrietta S. 1777 variables in the Magellanic Clouds // Annals of Harvard College Observatory, vol. 60, pp. 87-108.3. 1908. 60. 87–108.

Lemaître Georges. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques // Annales de la Société Scientifique de Bruxelles, A47, p. 49-59. 1927. 47. 49–59.

Lemon C, Anguita T, Auger-Williams MW, Courbin F, Galan A, McMahon R, Neira F, Oguri M, Schechter P, Shajib A, others . Gravitationally lensed quasars in Gaia–IV. 150 new lenses, quasar pairs, and projected quasars // Monthly Notices of the Royal Astronomical Society. 2023. 520, 3. 3305–3328.

Lemon Cameron A, Auger Matthew W, McMahon Richard G. Gravitationally lensed quasars in Gaia–III. 22 new lensed quasars from Gaia data release 2 // Monthly Notices of the Royal Astronomical Society. 2019. 483, 3. 4242–4258.

Linder Eric V. Exploring the expansion history of the universe // Physical review letters. 2003. 90, 9. 091301.

Lodge Oliver J. Gravitation and light // Nature. 1919. 104, 2614. 354–354.

MAST . Search MAST for Hubble. 09/2023. Accessed: 9/11/2023.

Martin Jerome. Everything you always wanted to know about the cosmological constant problem (but were afraid to ask) // Comptes Rendus Physique. 2012. 13, 6-7. 566–665.

McCully Curtis, Keeton Charles R, Wong Kenneth C, Zabludoff Ann I. Quantifying environmental and line-of-sight effects in models of strong gravitational lens systems // The Astrophysical Journal. 2017. 836, 1. 141.

Meneghetti Massimo. Introduction to gravitational lensing: with Python examples. 956. 2021.

Meylan Georges, Jetzer Philippe, North Pierre, Schneider Peter, Kochanek Christopher S, Wambsganss Joachim. Gravitational Lensing: Strong, Weak and Micro // Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro. 2006.

Millon M, Courbin F, Bonvin V, Paic E, Meylan G, Tewes M, Sluse Dominique, Magain Pierre, Chan JHH, Galan A, others . COSMOGRAIL-XIX. Time delays in 18 strongly lensed quasars from 15 years of optical monitoring // Astronomy & Astrophysics. 2020a. 640. A105.

Millon M, Galan A, Courbin F, Treu T, Suyu SH, Ding X, Birrer S, Chen GC-F, Shajib AJ, Sluse Dominique, others . TDCOSMO-I. An exploration of systematic uncertainties in the inference of H0 from time-delay cosmography // Astronomy & Astrophysics. 2020b. 639. A101.

Molinari Nicolas, Durand Jean-François, Sabatier Robert. Bounded optimal knots for regression splines // Computational statistics & data analysis. 2004. 45, 2. 159–178.

Mozumdar P, Fassnacht CD, Treu T, Spiniello C, Shajib AJ. TDCOSMO-XI. New lensing galaxy redshift and velocity dispersion measurements from Keck spectroscopy of eight lensed quasar systems // Astronomy & Astrophysics. 2023. 672. A20.

N Aghanim YA, Arroja F, Aumont J, Baccigalupi C, Ballardini M, Banday AJ, RB Barreiro NB, Basak S, Battye R, Benabed K, others . Planck 2018 results. I. Overview and the cosmological legacy of Planck // ASTRONOMY & ASTROPHYSICS. 2020. 641.

NASA/IPAC . NED - NASA/IPAC Extragalactic Database. 2023. Accessed:9/11/2023.

Narayan Ramesh, Bartelmann Matthias. Lectures on gravitational lensing // arXiv preprint astro-ph/9606001. 1996.

Narayan Ramesh, White Simon D. M. Gravitational lensing in a cold dark matter universe // Monthly Notices of the Royal Astronomical Society. 03 1988. 231, 1. 97P–103P.

Navarro Julio F, Frenk Carlos S, White Simon DM. A universal density profile from hierarchical clustering // The Astrophysical Journal. 1997. 490, 2. 493.

Newton Isaac. Opticks: or, A Treatise of the Reflections, Refractions, Inflections and Colours of Light. London: Royal Society, 1704.

Observatory Spanish Virtual. SVO Filter Profile Service. 08/05/2024. Accessed: 08/05/2024.

O'Riordan Conor M, Vegetti Simona. Angular complexity in strong lens substructure detection // Monthly Notices of the Royal Astronomical Society. 01 2024. 528, 2. 1757–1768.

Padovani Paolo, Alexander DM, Assef RJ, De Marco B, Giommi P, Hickox RC, Richards GT, Smolčić Vernesa, Hatziminaoglou E, Mainieri V, others . Active galactic nuclei: what's in a name? // The Astronomy and Astrophysics Review. 2017. 25. 1–91.

Pedregosa Fabian, Varoquaux Gaël, Gramfort Alexandre, Michel Vincent, Thirion Bertrand, Grisel Olivier, Blondel Mathieu, Prettenhofer Peter, Weiss Ron, Dubourg Vincent, Vanderplas Jake, Passos Alexandre, Cournapeau David, Brucher Matthieu, Perrot Matthieu, Duchesnay Édouard. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. 2011. 12, 85. 2825–2830.

Peebles P James E, Ratra Bharat. The cosmological constant and dark energy // Reviews of modern physics. 2003. 75, 2. 559.

Peebles PJE. Tests of cosmological models constrained by inflation // This page is intentionally left blank. 1984. 84.

Peebles Phillip James Edwin. The large-scale structure of the universe. 96. 1980.

Peebles Phillip James Edwin. Principles of physical cosmology. 27. 1993.

Penzias A.A. Measurement of Cosmic Microwave Background Radiation // IEEE Transactions on Microwave Theory and Techniques. 1968. 16, 9. 608–611.

Perivolaropoulos L., Skara F. Challenges for CDM: An update // New Astronomy Reviews. 2022. 95. 101659.

Perlick Volker. On Fermat's principle in general relativity. I. The general case // Classical and Quantum Gravity. 1990. 7, 8. 1319.

Peterson Bradley M. An introduction to active galactic nuclei. 1997.

Peterson Bradley M. Variability of active galactic nuclei // Advanced Lectures on the Starburst-AGN Connection. 2001. 3–68.

Refregier Alexandre. Shapelets—I. A method for image analysis // Monthly Notices of the Royal Astronomical Society. 2003. 338, 1. 35–47.

Refsdal Sjur. On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect // Monthly Notices of the Royal Astronomical Society. 1964. 128, 4. 307–310.

Riess Adam G., Anand Gagandeep S., Yuan Wenlong, Casertano Stefano, Dolphin Andrew, Macri Lucas M., Breuval Louise, Scolnic Dan, Perrin Marshall, Anderson Richard I. Crowded No More: The Accuracy of the Hubble Constant Tested with High-resolution Observations of Cepheids by JWST // The Astrophysical Journal Letters. oct 2023. 956, 1. L18.

Riess Adam G, Casertano Stefano, Yuan Wenlong, Bowers J Bradley, Macri Lucas, Zinn Joel C, Scolnic Dan. Cosmic distances calibrated to 1% precision with Gaia EDR3 parallaxes and Hubble Space Telescope photometry of 75 Milky Way Cepheids confirm tension with ΛCDM // The Astrophysical Journal Letters. 2021. 908, 1. L6.

Riess Adam G., Casertano Stefano, Yuan Wenlong, Macri Lucas M., Scolnic Dan. Large Magellanic Cloud Cepheid Standards Provide a 1

Riess Adam G, Filippenko Alexei V, Challis Peter, Clocchiatti Alejandro, Diercks Alan, Garnavich Peter M, Gilliland Ron L, Hogan Craig J, Jha Saurabh, Kirshner Robert P, others . Observational evidence from supernovae for an accelerating universe and a cosmological constant // The astronomical journal. 1998. 116, 3. 1009.

Riess Adam G., Macri Lucas M., Hoffmann Samantha L., Scolnic Dan, Casertano Stefano, Filippenko Alexei V., Tucker Brad E., Reid Mark J., Jones David O., Silverman Jeffrey M., Chornock Ryan, Challis Peter, Yuan Wenlong, Brown Peter J., Foley Ryan J. A 2.4

Riess Adam G, Yuan Wenlong, Macri Lucas M, Scolnic Dan, Brout Dillon, Casertano Stefano, Jones David O, Murakami Yukei, Anand Gagandeep S, Breuval Louise, others . A comprehensive measurement of the local value of the Hubble constant with 1 km s- 1 Mpc- 1 uncertainty from the Hubble Space Telescope and the SH0ES team // The Astrophysical journal letters. 2022. 934, 1. L7.

Riffeser Arno. Gravitational microlensing toward the Andromeda Galaxy: A search for dark matter in M31. I 2006.

Rizzo Francesca. A strong gravitational lensing view on the dynamical properties of high-redshift star-forming galaxies. 2020.

Roberts-Borsani Guido, Treu Tommaso, Chen Wenlei, Morishita Takahiro, Vanzella Eros, Zitrin Adi, Bergamini Pietro, Castellano Marco, Fontana Adriano, Glazebrook Karl, others . The nature of an ultra-faint galaxy in the cosmic dark ages seen with JWST // Nature. 2023. 618, 7965. 480–483.

Rojas Karina, Motta Verónica, Mediavilla E, Jimenez-Vicente J, Falco E, Fian C. Microlensing Analysis for the gravitational lens systems SDSS0924+ 0219, Q1355-2257, and SDSS1029+ 2623 // The Astrophysical Journal. 2020. 890, 1. 3.

Rubin Vera C, Ford Jr W Kent, Thonnard Norbert. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605/R= 4kpc/to UGC 2885/R= 122 kpc // Astrophysical Journal, Part 1, vol. 238, June 1, 1980, p. 471-487. 1980. 238. 471–487.

Rusu Cristian E, Wong Kenneth C, Bonvin Vivien, Sluse Dominique, Suyu Sherry H, Fassnacht Christopher D, Chan James HH, Hilbert Stefan, Auger Matthew W, Sonnenfeld Alessandro, others . H0LiCOW XII. Lens mass model of WFI2033- 4723 and blind measurement of its time-delay distance and H 0 // Monthly Notices of the Royal Astronomical Society. 2020. 498, 1. 1440–1468.

STSI Space Telescope Science Institute. HST Data Search. 07/10/2023. Accessed: 9/11/2023.

STScI . Photometric Keywords in SCI Extensions of ACS Images. 05/30/2024. Accessed:26/08/2024.

Saglia RP, Maraston C, Greggio L, Bender R, Ziegler B. The evolution of the color gradients of early-type cluster galaxies // arXiv preprint astro-ph/0007038. 2000.

Saha Prasenjit, Sluse Dominique, Wagner Jenny, Williams Liliya LR. Essentials of strong gravitational lensing // Space Science Reviews. 2024. 220, 1. 12.

Sánchez Ariel G. Arguments against using h- 1 Mpc units in observational cosmology // Physical Review D. 2020. 102, 12. 123511.

Schmidt RW, Wambsganss J. Quasar microlensing // General Relativity and Gravitation. 2010. 42. 2127–2150.

Schmidt T, Treu T, Birrer S, Shajib AJ, Lemon C, Millon M, Sluse Dominique, Agnello A, Anguita T, Auger-Williams MW, others . STRIDES: Automated uniform models for 30 quadruply imaged quasars // Monthly Notices of the Royal Astronomical Society. 2022. 518, 1. 1260–1300.

Schneider Peter, Ehlers Jürgen, Falco Emilio. Gravitational Lenses. 1992.

Schneider Peter, Kochanek Christopher S, Wambsganss Joachim, Wambsganss J. Gravitational Microlensing // Gravitational Lensing: Strong, Weak and Micro. 2006. 453–540.

Schneider Peter, Sluse Dominique. Mass-sheet degeneracy, power-law models and external convergence: Impact on the determination of the Hubble constant from gravitational lensing // Astronomy & Astrophysics. 2013. 559. A37.

Schneider Peter, Sluse Dominique. Source-position transformation: an approximate invariance in strong gravitational lensing // Astronomy & Astrophysics. 2014. 564. A103.

Scolnic D, Riess AG, Wu J, Li S, Anand GS, Beaton R, Casertano S, Anderson RI, Dhawan S, Ke X. CATS: The Hubble Constant from Standardized TRGB and Type Ia Supernova Measurements // The Astrophysical Journal Letters. 2023. 954, 1. L31.

Astrophysics lab: Strong gravitational lensing. // . 1998.

Sérsic JL. Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy // Boletin de la Asociacion Argentina de Astronomia La Plata Argentina. 1963. 6. 41–43.

Shajib Anowar J, Birrer S, Treu T, Agnello Adriano, Buckley-Geer EJ, Chan JHH, Christensen L, Lemon C, Lin H, Millon M, others . STRIDES: a 3.9 per cent measurement of the Hubble constant from the strong lens system DES J0408- 5354 // Monthly Notices of the Royal Astronomical Society. 2020. 494, 4. 6072–6102.

Shajib Anowar J, Birrer Simon, Treu T, Auger MW, Agnello A, Anguita T, Buckley-Geer EJ, Chan JHH, Collett TE, Courbin F, others . Is every strong lens model unhappy in its own way? Uniform modelling of a sample of 13 quadruply+ imaged quasars // Monthly Notices of the Royal Astronomical Society. 2019. 483, 4. 5649–5671.

Shajib Anowar J, Mozumdar Pritom, Chen Geoff C-F, Treu Tommaso, Cappellari Michele, Knabel Shawn, Suyu Sherry H, Bennert Vardha N, Frieman Joshua A, Sluse Dominique, others . TDCOSMO-XII. Improved Hubble constant measurement from lensing time delays using spatially resolved stellar kinematics of the lens galaxy // Astronomy & Astrophysics. 2023. 673. A9.

Shapiro Irwin I. Fourth test of general relativity // Physical Review Letters. 1964. 13, 26. 789.

Soldner Johann von. Über die Ablenkung eines Lichtstrahls von seiner geradlinigen Bewegung durch die Attraktion eines Weltkörpers, am welchem er nahe vorbeigeht // Berliner Astronomisches Jahrbuch. 1804.

Springel Volker, White Simon DM, Jenkins Adrian, Frenk Carlos S, Yoshida Naoki, Gao Liang, Navarro Julio, Thacker Robert, Croton Darren, Helly John, others . Simulations of the formation, evolution and clustering of galaxies and quasars // nature. 2005. 435, 7042. 629–636.

Suyu S. H., Treu T., Hilbert S., Sonnenfeld A., Auger M. W., Blandford R. D., Collett T., Courbin F., Fassnacht C. D., Koopmans L. V. E., Marshall P. J., Meylan G., Spiniello C., Tewes M. COSMOLOGY FROM GRAVITATIONAL LENS TIME DELAYS AND PLANCK DATA // The Astrophysical Journal Letters. jun 2014. 788, 2. L35.

Suyu SH, Auger MW, Hilbert S, Marshall PJ, Tewes M, Treu T, Fassnacht CD, Koopmans LVE, Sluse Dominique, Blandford RD, others . Two accurate time-delay distances from strong lensing: Implications for cosmology // The Astrophysical Journal. 2013. 766, 2. 70.

Suyu SH, Huber S, Cañameras R, Kromer M, Schuldt S, Taubenberger S, Yıldırım A, Bonvin V, Chan JHH, Courbin F, others . Holismokes-i. highly optimised lensing investigations of supernovae, microlensing objects, and kinematics of ellipticals and spirals // Astronomy & Astrophysics. 2020. 644. A162.

Suyu SH, Marshall PJ, Auger MW, Hilbert S, Blandford RD, Koopmans LVE, Fassnacht CD, Treu T. Dissecting the gravitational lens B1608+ 656. II. Precision measurements of the Hubble constant, spatial curvature, and the dark energy equation of state // The Astrophysical Journal. 2010. 711, 1. 201.

Tewes M, Courbin F, Meylan G. Cosmograil: the cosmological monitoring of gravitational lenses-XI. techniques for time delay measurement in presence of microlensing // Astronomy & Astrophysics. 2013. 553. A120.

Thorne Kip S, Will Clifford M. Theoretical frameworks for testing relativistic gravity. I. Foundations // Astrophysical Journal, vol. 163, p. 595. 1971. 163. 595.

Tie S. S., Kochanek C. S. Microlensing makes lensed quasar time delays significantly time variable // Monthly Notices of the Royal Astronomical Society. 09 2017. 473, 1. 80–90.

Times The New York. Lights All Askew in the Heavens // The New York Times. November 1919. Available at: `https://timesmachine.nytimes.com/timesmachine/1919/11/10/issue.html`.

Treu T, Agnello A, Baumer M A, Birrer S, Buckley-Geer E J, Courbin F, Kim Y J, Lin H, Marshall P J, Nord B, Schechter P L, Sivakumar P R, Abramson L E, Anguita T, Apostolovski Y, Auger M W, Chan J H H, Chen G C F, Collett T E, Fassnacht C D, Hsueh J-W, Lemon C, McMahon R G, Motta V, Ostrovski F, Rojas K, Rusu C E, Williams P, Frieman J, Meylan G, Suyu S H, Abbott T M C, Abdalla F B, Allam S, Annis J, Avila S, Banerji M, Brooks D, Rosell A Carnero, Kind M Carrasco, Carretero J, Castander F J, D'Andrea C B, daCosta L N, DeVicente J, Doel P, Eifler T F, Flaugher B, Fosalba P, García-Bellido J, Goldstein D A, Gruen D, Gruendl R A, Gutierrez G, Hartley W G, Hollowood D, Honscheid K, James D J, Kuehn K, Kuropatkin N, Lima M, Maia M A G, Martini P, Menanteau F, Miquel R, Plazas A A, Romer A K, Sanchez E, Scarpine V, Schindler R, Schubnell M, Sevilla-Noarbe I, Smith M, Smith R C, Soares-Santos M, Sobreira F, Suchyta E, Swanson M E C, Tarle G, Thomas D, Tucker D L, Walker A R. The STRong lensing Insights into the Dark Energy Survey (STRIDES) 2016 follow-up campaign – I. Overview and classification of candidates selected by two techniques // Monthly Notices of the Royal Astronomical Society. 08 2018. 481, 1. 1041–1054.

Tully R Brent, Fisher J Richard. A new method of determining distances to galaxies // Astronomy and Astrophysics, vol. 54, no. 3, Feb. 1977, p. 661-673. 1977. 54. 661–673.

Vegetti S., Koopmans L. V. E. Bayesian strong gravitational-lens modelling on adaptive grids: objective detection of mass substructure in Galaxies // Monthly Notices of the Royal Astronomical Society. 01 2009. 392, 3. 945–963.

Velten Hermano ES, Vom Marttens RF, Zimdahl Winifried. Aspects of the cosmological "coincidence problem" // The European Physical Journal C. 2014. 74. 1–8.

Verde Licia, Treu Tommaso, Riess Adam G. Tensions between the early and late Universe // Nature Astronomy. IX 2019. 3. 891–895.

Wall J. V., Jenkins C. R. Practical Statistics for Astronomers. 2003. (Cambridge Observing Handbooks for Research Astronomers).

Weinberg Steven. The cosmological constant problem // Reviews of modern physics. 1989. 61, 1. 1.

Wells Patrick, Fassnacht Christopher D, Rusu CE. TDCOSMO-XIV. Practical techniques for estimating external convergence of strong gravitational lens systems and applications to the SDSS J0924+ 0219 system // Astronomy & Astrophysics. 2023. 676. A95.

Willmer Christopher NA. The absolute magnitude of the sun in several filters // The Astrophysical Journal Supplement Series. 2018. 236, 2. 47.

Wills Beverley J, Brotherton MS, Fang D, Steidel Charles C, Sargent Wallace LW. Statistics of QSO broad emission-line profiles. I. The C IV lambda 1549 line and the lambda 1400 feature // Astrophysical Journal v. 415, p. 563. 1993. 415. 563.

Wong Kenneth C, Suyu Sherry H, Auger Matthew W, Bonvin Vivien, Courbin Frederic, Fassnacht Christopher D, Halkola Aleksi, Rusu Cristian E, Sluse Dominique, Sonnenfeld Alessandro, others . H0LiCOW–IV. Lens mass model of HE 0435- 1223 and blind measurement of its time-delay distance for cosmology // Monthly Notices of the Royal Astronomical Society. 2017. 465, 4. 4895–4913.

Wong Kenneth C, Suyu Sherry H, Chen Geoff CF, Rusu Cristian E, Millon Martin, Sluse Dominique, Bonvin Vivien, Fassnacht Christopher D, Taubenberger Stefan, Auger Matthew W, others . H0LiCOW–XIII. A 2.4 per cent measurement of H 0 from lensed quasars: 5.3 σ tension between early-and late-Universe probes // Monthly Notices of the Royal Astronomical Society. 2020. 498, 1. 1420–1439.

Yang Tao, Birrer Simon, Hu Bin. The first simultaneous measurement of Hubble constant and post-Newtonian parameter from Time-Delay Strong Lensing // Monthly Notices of the Royal Astronomical Society: Letters. 2020. 497, 1. L56–L61.

York Donald G, Adelman J, Anderson Jr John E, Anderson Scott F, Annis James, Bahcall Neta A, Bakken JA, Barkhouser Robert, Bastian Steven, Berman Eileen, others . The sloan digital sky survey: Technical summary // The Astronomical Journal. 2000. 120, 3. 1579.

Zarrouk Pauline, Burtin Etienne, Gil-Marín Héctor, Ross Ashley J, Tojeiro Rita, Pâris Isabelle, Dawson Kyle S, Myers Adam D, Percival Will J, Chuang Chia-Hsun, others . The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2 // Monthly Notices of the Royal Astronomical Society. 2018. 477, 2. 1639–1663.

Zhang CY. Robust Estimation and Image Combining // Astronomical Data Analysis Software and Systems IV. 77. 1995. 514.

Zwicky Fritz. Die rotverschiebung von extragalaktischen nebeln // Helvetica Physica Acta, Vol. 6, p. 110-127. 1933. 6. 110–127.

Zwicky Fritz. Nebulae as gravitational lenses // Physical Review. 1937. 51, 4. 290.

Acknowledgments

My first gratitude goes to my advisor, Dr. Stella Seitz, who, since she first met me at the beginning of my Master's program, always inspired me to do science and had more faith in my capabilities than what I myself had. Her determined and unrelenting curiosity for all fields of astrophysics and beyond has always pushed me to improve and work harder. I also want to thank Dr. Arno Riffeser, who was part of this project since day one, always ready to listen to my problems in the analysis and discuss them with an open mind. Not only his technical abilities, but even more his kindness and support have had a great impact on my PhD.

Furthermore, I want to express my gratitude to Prof. Dr. Ralf Bender for offering me the great opportunity to work as a PhD in the OPINAS group at MPE and for the numerous opportunities that I had during these years to expand my scientific and professional horizons.

I then want to thank the many people who made the project possible, most notably for the excellent data obtained from the 2.1-meter Wendelstein Telescope. Firstly, the builders of the telescope, Dr. Claus Gössl, Dr. Ulrich Hopp and Prof. Dr. Ralf Bender. Secondly, the observers, who worked sleepless nights for me to obtain wonderful lighcurves which were cardinal for this analysis: Christoph Ries, Michael Schmidt, Dr. Claus Gössl, Dr. Arno Riffeser, Raphael Zöller, Dr. Matthias Kluge.

I furthermore want to thank Leon Ecker for the numerous fruitful discussions we had in these years, and I wish him the best of luck in the continuation of his career.

To my two office mates, Anik Halder and Laurence Gong, I want to give a warm thank you for all the help and encouragement you have offered me during these years, both scientific and not. In particular, I want to thank Anik, as he was the one who "recruited" me to join the group, and he was always ready to lend a hand in the moment of need. I wish them both a successful career and a fulfilling life.

I then want to expand my thank you and best wishes to the USM ExGal group, which always had a welcoming environment and was always a space of interesting and enriching discussions.

Furthermore, I wish to give a special thank you to Prof. Dr. Sherry Suyu, who offered Leon and me the opportunity to access her group's weekly strong lensing seminar and was always very supportive and helpful.

On a more personal note, I want to thank my family and in particular my sister, Elena Queirolo, and my mother.

To Elena: thank you for always being on my side, no matter the physical distance between us, and for helping me see beyond the obstacles of life, showing me the bigger picture.
To my mother: thank you for supporting me and at the same time allowing me to have my space while maintaining a clear and reliable presence in my life. I then want to thank all my friends, both close and far; some of you directly motivated me during my PhD, and for that, I am grateful; all of you made my life enjoyable during it, and for that, I am even more grateful.
A special thank you goes to Martina; although our paths have now diverged, your support over the past three years has been unwavering and invaluable. I am deeply grateful for your help and encouragement, and I wish you all the best for the future.
I hope I will have the opportunity to repay all of you one day.