Gravitational lensing $H_0$ tension from ultralight axion galactic cores

Kfir Blum\textsuperscript{1} and Luca Teodori\textsuperscript{1}

\textsuperscript{1}Weizmann Institute of Science, Rehovot 7610001, Israel

Gravitational lensing time delays offer an avenue to measure the Hubble parameter $H_0$, with some analyses suggesting a tension with early-type probes of $H_0$. The lensing measurements must mitigate systematic uncertainties due to the mass modelling of lens galaxies. In particular, a core component in the lens density profile would form an approximate local mass sheet degeneracy and could bias $H_0$ in the right direction to solve the lensing tension. We consider ultralight dark matter as a possible mechanism to generate such galactic cores. We show that cores of roughly the required properties could arise naturally if an ultralight axion of mass $m \sim 10^{-25}$ eV makes up a fraction of order ten percent of the total cosmological dark matter density. A relic abundance of this order of magnitude could come from vacuum misalignment. Stellar kinematics measurements of well-resolved massive galaxies (including the Milky Way) may offer a way to test the scenario. Kinematics analyses aiming to test the core hypothesis in massive elliptical lens galaxies should not, in general, adopt the perfect mass sheet limit, as ignoring the finite extent of an actual physical core could lead to significant systematic errors.

I. INTRODUCTION

Measurements of the image and time delay of gravitationally-lensed quasar-host galaxies constrain the expansion rate of the Universe, parameterised via the Hubble constant $H_0$ \cite{Riess1998, Perlmutter1999, Connolly2002}. In a work that summarised the efforts of several groups, the TDCOSMO team\textsuperscript{1} used these data to derive $H_0 = 74.0^{+1.7}_{−1.8}$ km/s/Mpc (TDCOSMO-I \cite{TDCOSMO1}). This result is in tension with measurements based on the cosmic microwave background (CMB) \cite{Planck2013}, which find $H_0 = 67.36 \pm 0.54$ km/s/Mpc, and with large scale structure (LSS) galaxy clustering that is consistent with the CMB \cite{Planck2015, Viel2013, Wechsler2012}. We refer to the apparent discrepancy between the lensing \cite{TDCOSMO1} and the CMB/LSS \cite{Planck2015, Viel2013, Wechsler2012} measurements as the lensing $H_0$ tension.

The lensing $H_0$ measurement of Ref. \cite{TDCOSMO1} is independent of the well-known cepheid-calibrated supernova-la (SNIa) measurements by the SH0ES collaboration, which find $H_0 = 73.2 \pm 1.3$ km/s/Mpc \cite{Riess2011}. The lensing result \cite{TDCOSMO1} is in excellent agreement with the SNIa cepheids result \cite{Riess2011} and both are "late Universe" probes of $H_0$, that is, they involve only low-redshift ($z \sim 1$) dynamics, in contrast to the CMB/LSS measurements which can be considered "early Universe" probes because they hinge crucially on high-redshift ($z \sim 10^3$) dynamics such as the baryonic perturbations sound horizon. Discrepancy between early and late determinations of $H_0$ could indicate a long-awaited breakdown of the $\Lambda$CDM effective description of cosmology \cite{Planck2013, Riess2011}. After all, we understand no more than 5% of the energy budget of the Universe. It is tantalising to think that a clue to the nature of the remaining 95% may come from the $H_0$ tension.

Nevertheless, all of the methods to determine $H_0$ require a careful account of systematic uncertainties. A main concern in the SNIa analyses is the calibration of local distance ladder anchors. The TRGB-calibrated SNIa analysis of Ref. \cite{Freedman2001}, for example, finds a value of $H_0$ that is consistent to $\sim 1\sigma$ with the CMB result, despite a nominal precision that is comparable to the SNIa/cepheids method. (See, however, \cite{Freedman2012}. And of course, there are concerns of systematic issues in the CMB analysis, too \cite{Planck2015}.)

Lensing measurements of $H_0$ are detached from the distance ladder. However, modelling degeneracies couple the inferred value of $H_0$ to the assumed density profile of the lens galaxy \cite{Kenheim2010, TDCOSMO2}. Ref. \cite{Kenheim2010} pointed out that a core component in the lens galaxy density profile could comprise an approximate internal mass sheet degeneracy (MSD), shifting the inferred value of $H_0$ without affecting the image reconstruction and without conflict with estimates of cosmological external convergence. Subsequently, TDCOSMO-IV \cite{TDCOSMO2} added an effective "internal MSD" degree of freedom to their halo model fit; as a result, the error budget on $H_0$ increased to the level expected from stellar kinematics, around 10% \cite{TDCOSMO2, TDCOSMO3}. Interestingly, including galaxies from the Sloan Lens ACS (SLACS) survey \cite{Slacs2009} in the kinematics analysis, and making the additional assumption that SLACS and TDCOSMO galaxies share a self-similar structure, shifted the central value of the lensing $H_0$ to the CMB value while providing some positive evidence for an internal MSD component in the data. The status of the lensing $H_0$ measurements is illustrated in Fig. 1.

In what follows we use the term "core-MSD" instead of "internal MSD", to highlight the fact that a natural interpretation of the added degree of freedom in the halo model corresponds to a physical core feature in the density profile \cite{TDCOSMO2}.

We should emphasise that the hint \cite{TDCOSMO2} for a core-MSD could eventually go away after further scrutiny of uncertainties in conventional halo models \cite{Sawicki2012}. Nevertheless, even setting aside the results of \cite{TDCOSMO2}, it is interesting to examine the possibility of an actual core driving the lensing $H_0$ tension. The question then is, what is the core made of? If the core is not traced by the light profile of the lens, then it is natural to speculate that it could come from dark matter, perhaps providing a clue to
dark matter properties.

We consider the possibility that such cores come from ultralight dark matter (ULDM). ULDM has been studied extensively in recent years, and we do not give a thorough coverage of the literature here; see references to and from [30, 31]. ULDM is known to develop a cored density profile (”soliton”) due to gravitational dynamical relaxation. The phenomenon has been identified in numerical simulations by different groups [32, 33], and is consistent with analytic considerations which show that the soliton is an energy-minimiser at fixed mass, and thus an attractor solution of the equations of motion [34].

Fig. 2 illustrates our idea. It shows the different density components (stellar mass and dark matter) of a would-be lens galaxy. We include an ULDM soliton core with a total mass of $M = 1.4 \times 10^{12} M_\odot$ at a particle mass $m = 2 \times 10^{-25}$ eV. These ULDM parameters ($M$ and $m$) are chosen such that the core extends sufficiently far beyond the projected Einstein radius $R_E$ to keep imaging errors undetectable for typical current lensing reconstruction measurement uncertainties.

From a theoretical perspective, ULDM is a compelling possibility. If the spectrum of particles contains an ultralight boson, like the axions of some string-inspired models [45, 46], then the phenomenon of vacuum misalignment generically predicts that such a boson would behave as dark matter if the particle mass satisfies $m \gtrsim H_0 \approx 10^{-93}$ eV. If the boson $\chi$ is an axion with decay constant $f$, vacuum misalignment predicts [31]

$$\Omega_\chi \approx 0.3 \left(\frac{10^{-21} \text{ eV}}{m}\right)^{1/2} \left(\frac{f}{10^{17} \text{ GeV}}\right)^2,$$  \hspace{1cm} (1)

where $\Omega_\chi / \Omega_{\text{crit}}$ is the ratio of the ULDM relic density to the critical density of the Universe and $\Omega_\chi \approx 0.3$ would saturate the total dark matter contribution $\Omega_{\text{dm}}$ inferred from cosmological data. This puts ULDM with $m \approx 10^{-21}$ eV in the right order of magnitude to make up all of the dark matter if $f$ is around the grand-unification or string scale.

As we shall see, the interesting mass range for our current analysis is actually $m \sim 10^{-25}$ eV, give or take an $\mathcal{O}(1)$ factor. Cosmological and astrophysical observations imply that such ULDM can only comprise a fraction of the total dark matter. We thus define the cosmological ULDM fraction

$$\alpha_\chi = \frac{\Omega_\chi}{\Omega_{\text{dm}}}, \hspace{1cm} (2)$$

we will be led to consider $\alpha_\chi < 1$. In this case, the remaining dark matter must take some other form (e.g., higher-$m$ axions).

Rotation curves of low-surface-brightness galaxies are inconsistent with $\alpha_\chi = 1$ for $m \lesssim 10^{-21}$ eV [47], but these constraints have not been evaluated for $\alpha_\chi < 1$. Recently, Ref. [48] reported constraints that combine galaxy clustering data [49] with Planck15 CMB data [50] (see [51] for an earlier analysis of the CMB data). The constraint on $\alpha_\chi$ depends on the value of $m$; for example, for $m = 10^{-25}$ eV, the $2\sigma$ CL combined limit is $\alpha_\chi \lesssim 0.34$, while for $m = 10^{-26}$ eV the limit tightens to $\alpha_\chi \lesssim 0.035$. Additional constraints come from the Ly-$\alpha$ forest line absorption power spectrum [52], which can be roughly summarised by $\alpha_\chi \lesssim 0.16$ at $2\sigma$ CL for $m < 10^{-22}$ eV. The constraint becomes weaker towards larger $m$ and disappears for $m \gtrsim 10^{-20}$ eV. We note that the Ly-$\alpha$ bound of [52] was not explicitly computed and must be extrapolated to the low values of $m$ where we will use it; keeping that in mind, and noting in addition that systematic uncertainties associated with the heating and ionisation history of the intergalactic medium could affect the Ly-$\alpha$ analyses to some extent, we allow ourselves to explore $\alpha_\chi$ as large as 0.2.

Eq. (1) tells us that ULDM at $m \approx 10^{-25}$ eV could easily make up $\mathcal{O}(10\%)$ of the total dark matter, in the vanilla misalignment scenario with $f \approx 3 \times 10^{17}$ GeV.
The rest of the paper is arranged as follows. In Sec. II we recap the core-MSD set-up of [26], explaining the connection between imaging errors and the possible range of the shift in the inferred value of $H_0$. In Sec. III we show how an ULDM soliton produces a core-MSD profile. Using a simplified prescription to estimate imaging constraints we explore the ULDM parameter space. In Sec. IV we study stellar kinematics. We find that the perfect MSD limit, adopted in the kinematics analysis of TDCOSMO-IV [27], needs to be revised if one wishes to explore a realistic physical core-MSD model.

Our analysis suggests that ULDM could solve the lensing $H_0$ tension, provided it condenses into sufficiently massive solitons in the lens galaxies. In Sec. V we consider the theoretical consistency of this scenario. We show that ULDM solitons of roughly the right mass could produce a core-MSD profile in the lens galaxies. In Sec. VI we consider kinematics and dynamics in well-resolved galaxies, like our own Milky Way. We summarise in Sec. VII.

App. A contains technical details of the distortion of the soliton under a power-law background density profile. App. B contains analyses of mock data, with references to our implementation of the ULDM model in the lensing software package lenstronomy [53]. App. C contains some details of the kinematics analysis.

II. THE CORE-MSD MODEL

Consider a lensing reconstruction model $\kappa_0(\theta)$ for the convergence of the lens. A core-MSD model can be constructed from $\kappa_0(\theta)$ by adding a core component $\kappa_c(\theta)$ while rescaling the original model:

$$\kappa(\theta) = \kappa_c(\theta) + (1 - \kappa_c(\theta_E))\kappa_0(\theta).$$

Here $\theta_E$ is defined by $\kappa_0(\theta_E) = \theta_E$, where $\alpha_0(\theta)$ is the deflection angle due to $\kappa_0(\theta)$. At the same time, the source plane coordinates are rescaled as $\beta = (1 - \kappa_c(\theta_E))\beta_0$. On angular scales $\theta \gg \theta_E$ it is assumed that $\kappa_c(\theta) \to 0$ such that the core-MSD effect commutes with external convergence.

Eq. (3) is an approximate MSD if $\kappa_c(\theta)$ is nearly constant up to $|\theta|$ that is sufficiently larger than $|\theta_E|$. To be quantitative, we can define the correction $\delta \kappa_E$ via:

$$\alpha(\theta_E) = \theta_E (1 + \delta \kappa_E),$$

where $\alpha(\theta)$ is the deflection angle of the full model. $\delta \kappa_E$ quantifies the relative imaging error in the vicinity of $\theta \approx \theta_E$, the angular range where lensing analyses have the most constraining power. For simplicity, in this estimate we assume the system to be spherically symmetric, so that $\alpha(\theta) = 2\theta\int_0^1 dz z \kappa(z \theta)$. Using Eq. (3) we then have

$$\delta \kappa = 2\int_0^1 dz (z \kappa(z \theta_E) - \kappa_c(\theta_E))$$

$$= \frac{\alpha_c(\theta_E)}{\theta_E} - \kappa_c(\theta_E).$$

The first line in Eq. (5) shows that constant $\kappa_c(\theta)$ within $\theta < \theta_E$ produces an MSD, and the second is convenient for quantifying corrections when $\kappa_c(\theta)$ is not exactly constant. While this estimate was given for a spherical lens, it gives a good approximation of the imaging error also for the non-symmetric systems arising in realistic analyses, as we will verify using mock data.

Consider the possibility that a lens galaxy harbours a core component, leading to a true convergence profile resembling Eq. (3) with $\kappa_c(\theta_E) > 0$. In this case, both the null model $\kappa_0(\theta)$ and the core-MSD model $\kappa(\theta)$ would give a good description of the imaging data. However, the true value of $H_0$ would differ from the inferred value in the null model by:

$$\frac{H_0, \text{inferred}}{H_0, \text{true}} = \frac{\delta H_0}{H_0} \approx \kappa_c(\theta_E).$$

Tab. I shows the values of $\delta H_0/H_0$ required to bring the different systems to accord with the CMB result. We see that $\kappa_c(\theta_E) \approx 0.1$, with some variation between systems [54-59], could solve the lensing $H_0$ tension.

III. CORE-MSD WITH AN ULDM SOLITON

An ULDM soliton could produce the $\kappa_c$ term in Eq. (3). We now derive some results that are useful for the lensing analysis; for a detailed discussion and more references concerning ULDM solitons, we refer the reader to [47].

The ULDM soliton field is described by a function $\chi(r)$, where we define the rescaled coordinate $r = m x$. The mass density associated with $\chi$ is given by

$$\rho = \frac{m^2}{4\pi G} \chi^2,$$

where $G$ is Newton's constant. The field $\chi$ and the Newtonian gravitational potential sourced by it, $\Phi$, satisfy the Schrodinger-Poisson equations (SPE) [47]

$$\partial_r^2(r \chi) = 2r (\Phi + \Phi_{\text{ext}} - \tilde{\gamma}) \chi,$$

$$\partial_r^2(r \Phi) = r \chi^2.$$
The convergence, deflection angle, and lensing potential are:

\[ \chi \text{ starts off constant at } r \rightarrow 0 \text{ and decays to zero with no nodes. We solve the } \text{SPE numerically.} \]

The solution is fixed by a single parameter that we can take to be the value of \( \chi \) at \( r = 0 \). We thus define the solution \( \chi(r) \) via

\[ \chi(r = 0) = \lambda^2, \tag{10} \]

with a real parameter \( \lambda \). It is convenient to use the scaling relation \[ \chi(r; \Phi_{\text{ext}}(r)) = \lambda^2 \chi(\lambda r; \lambda^{-2} \Phi_{\text{ext}}(\lambda^{-1} r)), \tag{11} \]

meaning that in numerical investigations, it is always enough to compute \( \chi_1 \). For clarity, in Eq. (11) we explicitly note how the external potential enters the solution.

It is convenient to introduce an approximation with which properties of the soliton can be derived analytically. We choose

\[ \chi_1(r) \approx \frac{1}{1 + a^2 r^2} \beta, \tag{12} \]

where the coefficients \( a \) and \( b \) are fitted numerically to the exact solution. For a self-gravitating soliton (the limit \( \Phi_{\text{ext}} \rightarrow 0 \)), we obtain \( a \approx 0.23 \) and \( b \approx 3.9 \). When \( \Phi_{\text{ext}} \neq 0 \), the coefficients \( a \) and \( b \) depend on \( \lambda \) and \( \Phi_{\text{ext}} \) via the combination \( \lambda^{-2} \Phi_{\text{ext}}(\lambda^{-1} r) \).

With the approximation of Eq. (12), the soliton mass is

\[ M_\lambda = \frac{\lambda}{G m} \int dr r^2 \chi_1^2 \approx \frac{\lambda}{G m} \frac{\sqrt{\pi}}{a^3/4} \Gamma(2b - \frac{3}{2}), \tag{13} \]

The convergence, deflection angle, and lensing potential are:

\[ \kappa_\lambda(\theta) \approx \frac{\lambda^3}{4 \pi G \Sigma_{\text{c}}} \frac{\sqrt{\pi}}{a} \frac{\Gamma(2b - \frac{1}{2})}{\Gamma(2b)} \frac{1}{(1 + \frac{a^2}{\theta^2})^{2b - \frac{1}{2}}}, \tag{14} \]

\[ \alpha_\lambda(\theta) \approx \frac{2 \theta^2}{(4b - 3) \theta} \left( 1 - \left( 1 + \frac{a^2}{\theta^2} \right)^{-2b - \frac{1}{2}} \right), \tag{15} \]

\[ \psi_\lambda(\theta) \approx \frac{\kappa_\lambda(0)}{2} \Gamma_{2} \left[ \{1, 1, 2b - \frac{1}{2} \}, \{2, 2\}; -\frac{\theta^2}{\theta_c^2} \right], \tag{16} \]

where we defined the core angle

\[ \theta_c = \frac{1}{\lambda am D_\ell} \tag{17} \]

The critical density entering the convergence computation is \( \Sigma_{\text{c}}^{-1} = 4 \pi G D_{\ell} D_{ho}/D_{ho} \), where \( D_{l,ho} \) are the angular diameter distances to the lens, source, and between the lens and the source. \( \mathcal{F}_q \left[ \alpha, \beta; z \right] \) is the generalised hypergeometric function. In the MSD limit, \( \theta_c \gg \theta \), one can verify that \( \kappa_\lambda(\theta) \approx \kappa_\lambda(0), \alpha_\lambda(\theta) \approx \kappa_\lambda(0) \theta, \) and \( \psi_\lambda \approx \kappa_\lambda(0) \theta^2/2 \).

Adopting the soliton as our core-MSD component, we set \( \kappa_c(\theta) \equiv \kappa_\lambda(\theta) \) in Eq. (3). From Eqs. (5) and (6) we get

\[ \frac{\delta H_0}{H_0} \approx \frac{\kappa_\lambda(0)}{(1 + \frac{a^2}{\theta^2})^{2b - \frac{1}{2}}} \tag{18} \]

and

\[ \delta_{\text{E}} \approx \frac{(1 + \frac{a^2}{\theta^2})^{2b - \frac{1}{2}} - (2b - \frac{1}{2}) \frac{a^2}{\theta^2} - 1}{(2b - \frac{3}{2}) \frac{a^2}{\theta^2}} \frac{\kappa_\lambda(0)}{(1 + \frac{a^2}{\theta^2})^{2b - \frac{1}{2}}} \frac{\delta H_0}{H_0} \tag{19} \]

For \( \theta_c < \theta \), we have \( \delta_{\text{E}} \approx (b - \frac{1}{4}) \frac{a^2}{\theta^2} \frac{\delta H_0}{H_0} \). This shows how imaging uncertainties, roughly summarised by \( \delta_{\text{E}} \), constrain the shift \( \delta H_0/H_0 \) at given soliton core angle \( \theta_c \).

In App. [A] we calculate how the soliton profile is distorted in the presence of a power-law (PL) background. We find that Eq. (12) is still a good approximation, sufficient for our needs; the effect of the background density is to modify the numerical values of the coefficients \( a \) and \( b \).

Before proceeding to observational constraints, we comment on the parameter space of the model. As noted

\[ \text{TABLE I: Lens systems from [4]. Values for } H_0 \text{ (in km/s/Mpc) are from the PL fit (Fig. 6 in [4]). The reference "true" } H_0 \text{ used to define } \delta H_0/H_0 \text{ is taken from the CMB result } H_0 = 67.36 \pm 0.54 \text{ km/s/Mpc [5]. } \theta_c \text{ is in arcsec. } \sigma^2 \text{ is in km/s. On the last column we show twice the maximum relative error of the velocity anisotropy, useful for comparison with } \delta H_0/H_0 \text{ (see discussion in Sec. IV).} \]

| System | \( \delta H_0/H_0 \) | \( \gamma \) | \( \theta E \) | \( \delta E \) | \( z_1 \) | \( z_n \) | \( \sigma^2 \) | \( 2/\sigma^2 \) |
|--------|------------------|--------|---------|--------|------|--------|---------|--------|
| RXJ1131 | 0.13^{+0.05}_{-0.06} | 1.98 | 1.6 | 0.006 | 0.295 | 0.654 | 320 \pm 20 | 0.125 |
| PG1115 | 0.23^{+0.11}_{-0.10} | 2.18 | 1.1 | 0.02 | 0.311 | 1.722 | 280 \pm 25 | 0.178 |
| HE0435 | 0.06^{+0.07}_{-0.07} | 1.87 | 1.2 | 0.025 | 0.4546 | 1.693 | 220 \pm 15 | 0.136 |
| DESJ0408 | 0.11^{+0.04}_{-0.04} | 2 | 1.9 | 0.01 | 0.597 | 2.37 | 230 \pm 27 | 0.235 |
| WFI2033 | 0.08^{+0.05}_{-0.04} | 1.95 | 0.9 | 0.016 | 0.6575 | 1.662 | 250 \pm 19 | 0.152 |
| J1206 | -0.01^{+0.08}_{-0.07} | 1.95 | 1.2 | 0.025 | 0.745 | 1.789 | 290 \pm 30 | 0.207 |

\[ \text{A rapidly converging expression is } \mathcal{F}_q \left[ \left\{ \frac{1}{2}, 1 \right\}, \left\{ 2, 2 \right\}; z \right] = \frac{1}{(4 - 4z)} \left( \log(1 - z) - \sum_{n=1}^{\infty} \frac{1}{n} \left( \frac{1}{2} - \frac{1}{n} \right) \right). \]
in the beginning of this section, at a fixed value of the ULDM particle mass \( m \), the soliton is a single-parameter function. While the scaling parameter \( \lambda \) from Eq. (10) is convenient for analytical expressions, in making contact with observations we prefer to use the total soliton mass \( M \), substituting \( \lambda \rightarrow \lambda(M, m) \) using Eq. (13). (The detailed matching, but not the basic procedure, is slightly modified with a background potential as explained in App. A.) All other properties of the core (the convergence, for example) then depend only on \( m \) and \( M \). The full parameter space is therefore covered when we analyse our results in terms of \( m \) and \( M \).

A. Constraints on ULDM from TDCOSMO systems

We are ready for a rough assessment of the lensing \( H_0 \) tension in the ULDM model. In Fig. 3 we explore \( \delta H_0/H_0 \) and \( \delta \theta_E \) as function of the ULDM particle mass \( m \) (\( x \)-axis) and soliton mass \( M \) (\( y \)-axis). The different panels correspond to the different systems in Tab. I. The information in the plot is as follows.

We begin with results that include the effect of a background (non-ULDM) external potential, modelled as a pure PL, using the results in App. A. For simplicity, the same PL index \( \gamma = 2 \) is used for all systems, but the value of \( \theta_E \) and \( \delta \theta_E \) for each system is as in Tab. I. In the pale red shaded region, \( \delta \theta_E \) exceeds its corresponding value from Tab. I. This region is disfavoured by the imaging data. Along the blue dashed line, \( \delta H_0/H_0 \) matches the value required to solve the \( H_0 \) tension. The solid blue lines delimit the uncertainty on \( \delta H_0/H_0 \) for each system. (Other curves in Fig. 3 correspond to theoretical constraints and are explained in Sec. IV).

We also show how the imaging constraints change if the external PL density is not included in the soliton computation. The result is shown by the dark red shaded region. The imaging constraint is generally weaker when the PL effect is not included, compared to when it is (i.e., the dark shaded region is contained inside the pale region), because the background potential causes the soliton to contract inwards at fixed \( m \) and \( M \), decreasing \( \theta_E \) and leading to stronger violation of the MSD.

The fact that the PL background analysis provides stronger imaging constraints, compared with the self-gravity case, illustrates the sensitivity of the analysis to the detailed mass profile of the lens. However, the soliton contraction is mostly driven by the cuspy PL mass distribution at small \( r \ll R_E \), where the lensing observables are not well constrained. In fact, the observed stellar surface brightness of the lenses display cores rather than cusps on distances \( r \ll R_E \), where the stellar density dominates over the DM. As a result, physically-motivated composite stellar+DM halo models, adjusted to fit the stellar light profile, predict a contraction effect on the soliton that is less significant than in the PL background. The imaging constraint in these more realistic background models are closer to the self-gravitating soliton result.

In Fig. 2 we show a soliton solution of the lensing \( H_0 \) tension, using a composite stellar+DM model that mimics the properties of the system DESJ0408. The solution has \( m = 2 \times 10^{-25} \text{ eV} \) and \( M \approx 1.4 \times 10^{12} \text{ M}_\odot \), marked in Fig. 3 by a circle (bottom-left panel). This solution has \( \delta H_0/H_0 \approx 0.1 \) and we have verified that it is compatible with the requirement \( |\delta \theta_E| < 0.01 \), valid for this system. The fact that this solution would seem to be excluded in the PL background analysis is due to the exaggerated soliton contraction in the PL case.

We can use the self-gravitating soliton case to understand the imaging constraints parametrically. In this limit \( a \approx 0.23 \) and \( b \approx 3.9 \) can be used in Eqs. (13-19) and the \( H_0 \) shift is

\[
\frac{\delta H_0}{H_0} \approx \frac{2.9 \lambda m}{4 \square G \Sigma_c}.
\]

On the other hand, in the same self-gravity limit we have \( \delta \theta_E \approx 3.6 (\theta_E/\theta_c)^2 \delta H_0/H_0 \). Demanding \( \delta \theta_E \lesssim 0.01 \), as in typical systems, and setting \( \delta H_0/H_0 \approx 0.1 \), we should impose \( \theta_c \gtrsim 6 \theta_E \), or

\[
\lambda m \lesssim \frac{0.7}{D_l \theta_E}. \tag{21}
\]

Combining Eqs. (20) and (21) we obtain,

\[
m \lesssim 10^{-24} \text{ eV} \left( \frac{D_{ls}}{D_l} \right)^\frac{1}{1} \left( \frac{1}{\theta_E} \right)^\frac{2}{1} \left( \frac{1 \text{ Gpc}}{1 \text{ Gpc}} \right)^\frac{0.1}{1} \left( \frac{\delta H_0/H_0}{0.1} \right)^\frac{1}{1}. \tag{22}
\]

Again, the presence of an external potential (PL or composite) contracts the soliton inward to some extent at fixed \( M \) and \( m \), shifting the upper limit of Eq. (22) to somewhat lower \( m \).

Our discussion of the imaging constraints was simplistic, in that we used the rough Einstein angle criterion \( \delta \theta_E \) to constrain the possible shift to \( H_0 \). In comparison, the likelihood function in real lensing analyses contains detailed extended source information as well as multiple modelling parameters, experimental seeing limitations, etc. In App. B we present a numerical study of a mock system, including most of these complications, using lenstronomy. This numerical study serves two purposes. First, we introduce an implementation of the ULDM module in lenstronomy. In future work we plan to use this tool to test the ULDM model including the full lensing likelihood. Second, this exercise allows us to test the accuracy of the simple \( \delta \theta_E \) criterion. We find that the naive \( \delta \theta_E \) criterion is slightly conservative compared with a full analysis: for example, at fixed \( \delta H_0/H_0 \approx 0.1 \), we find that a full numerical analysis yields a constraint on \( \theta_c \) (and therefore, equivalently, on \( m \) at fixed \( M \)) that is about a factor of 2 weaker than the constraint we would obtain using the naive \( \delta \theta_E \) criterion.
FIG. 3: ULDM soliton as a solution of the lensing $H_0$ tension. Panels correspond to different systems in Tab. I. The dark red shaded region shows where $\delta E$ exceeds its limits from Tab. I, computed neglecting the effect of an external mass distribution on the soliton. The pale shaded region shows the same constraint, including the effect of an external PL potential. Along the blue dashed line, $\delta H_0/H_0$ matches the value from Tab. I, with solid lines delimiting the uncertainty; this result also includes the PL background. (In J1206 and HE0435 the central value and/or lower limit of $\delta H_0/H_0$ are compatible with zero.) The relaxation estimate of Eq. (35) (see Sec. V) is shown by green, pink, and purple bands for $\alpha = 0.2, 0.1, 0.05$, respectively, using $\sigma_0 = \sigma^p$. The band width is defined by the uncertainty in $\sigma^p$. We truncate each constant-$\alpha$ band at small $m$ according to the cosmological constraints from Ref. [48]. The saturation estimate $[(K/M)_\chi < 1.5(A/2)]$ is shown by thick dashed grey line. The circle in the panel of DESJ0408 marks the set-up of Fig. 2.

IV. STELLAR KINEMATICS

Stellar kinematics measurements break the MSD, and are the limiting observational factor to a core-MSD shift of $H_0$. The basic observable is the luminosity-weighted velocity dispersion along the line of sight, $\sigma_{\text{los},0}$, given by [59]

$$\sigma^2_{\text{los},0} = \frac{2G}{I(\theta)} \int_1^\infty \frac{dy}{y} K(y, \theta_a) l(yD_\gamma \theta) M(yD_\gamma \theta).$$

(23)

Here, $l(r)$ is the stellar luminosity density, $I(\theta)$ is the surface brightness, $M(r)$ is the total enclosed mass, and the function $K(u, u_a)$ encodes the velocity anisotropy profile with Osipkov-Merritt [60, 61] anisotropy radius $r_a = D_\gamma \theta_a$ [59]. For analytical estimates, we note that the isotropic velocity limit gives $K(u, \infty) = \sqrt{1 - 1/u^2}$.

The core-MSD model enters Eq. (23) via the mass profile, $M(r) = M_c(r) + (1 - \kappa_c)M_0(r)$, where $M_0(r)$ comes from the null model and $M_c(r)$ from the core. The dispersion of the full model is related to that of the null model, $\sigma^2_{\text{los},0}$, via

$$\frac{\sigma^2_{\text{los,0}}}{\sigma^2_{\text{los},0}} = 1 - \kappa_c (1 - \delta_c),$$

(24)

$$\delta_c = \frac{1}{\kappa_c} \frac{\sigma^2_c}{\sigma^2_{\text{los},0}}.$$

(25)

where $\sigma^2_c$ is the velocity dispersion due to the core itself. In general, all of $\sigma^2_{\text{los},0}$, $\sigma^2_c$, $\sigma^2_{\text{los}}$, and $\delta_c$ depend on the measurement point $\theta$. In Eq. (24), the term $\delta_c$ parametrises the deviation from the perfect MSD limit. It becomes small for $\theta_c \gg \theta$, but may be quantitatively relevant once we consider a finite soliton core, and once kinematics data
probing $\theta$ not much smaller than $\theta_e$ is used.

To see this, consider an isothermal PL profile for the null model, for which $M_0(r) = 2\alpha^2 r/G$ where $\alpha^2$ is the physical velocity dispersion. In convenient angular variables we can trade $\sigma^2_e$ for $\theta_e$, noting that $M_0(\theta) = 2\Sigma_e D^2 \theta_e$. We also take the isotropic limit of $K$, and consider the Hernquist profile \(^{[62]}\) for the luminosity density, $l(r) = I_0 r^4 / [2\pi r (r + r_s)^3]$. The parameter $r_s$ is related to the commonly used effective radius $r_e$ via $r_e \approx 1.8 r_s$ \(^{[62]}\). With these simplifications, and using Eq. (12) for the soliton profile (with $b \approx 3.9$), we obtain

$$
\delta_e(\theta) = \pi \theta^2 \overline{\theta} \psi E_1 \left( \frac{\theta_e}{\theta} \right)
$$

$$
\approx 0.31 \left( \frac{10''}{\theta_e} \right) \left( \frac{1}{\overline{\theta}^2} \right)^2 \frac{\theta}{0.5''} f \left( \frac{\theta}{\overline{\theta}}, \frac{\theta}{\overline{\theta}} \right)
$$

where

$$
f(y_c, y_+ = \frac{-2F(1/2) \int_{1}^{\infty} dy \sqrt{y^2 - 1}}{\sqrt{y^2 - 1}} 2F_1 \left( \frac{3}{2}, \frac{5}{2}, \frac{5}{2}, -y^2 \right) \int_{1}^{\infty} dy \sqrt{y^2 - 1} \sqrt{y^2 + y_+^2} \right) \right)
$$

TDCOSMO-IV \(^{[27]}\), in considering the effect of an “internal MSD”, have assumed in practice the perfect MSD limit $\delta_e = 0$ in their kinematics analysis of TDCOSMO and SLACS systems. The approximation was tested using a mock system with $\theta_e = 0.2''$ and several core radii $\theta_c$. However, the parametric breaking of the MSD, captured by Eq. (26), was not explored for different values of $\theta_e$ or the baseline $\theta_E$ and system redshifts (equivalently $\sigma_c$). As Eq. (26) suggests a strong dependence on the kinematics observation point, it is important to check to what extent the MSD limit is expected to hold across different systems.

For TDCOSMO systems \(^{[4]}, [27]\), the kinematics constraints were based on a single effective measurement centred on $\theta = 0$ and averaged over an aperture $A$, weighted by the surface brightness $I(\theta)$. To be more precise, the observationally accessible dispersion $\sigma^P$ is given by \(^{[19, 63]}\)

$$
(\sigma^P)^2 = \frac{\int_A I(\theta)^2 \sigma^2_{\text{los}}(\theta) \sigma^P(\theta) \, d\theta}{\int_A I(\theta)^2 \sigma^2(\theta) \, d\theta},
$$

where $P$ is the seeing. It is natural to define

$$
(\sigma^P)^2 = 1 - \kappa_c (1 - \Delta_c),
$$

$$
\Delta_c = \frac{\int_A d\theta I(\theta) \sigma^2_{\text{los},0}(\theta) \delta_e(\theta) \, d\theta}{\int_A d\theta I(\theta) \sigma^2_{\text{los},0}(\theta)}.
$$

From this expression and the previously quoted results, the correction term $\Delta_c$ can be evaluated numerically. It depends on $\theta_e$, $\theta_c$ (equivalently $\theta_e$), the aperture $A$, and the seeing $P$. The main point to explore is how $\Delta_c$ reacts to different values of $\theta_e$ and $\theta_c$.

In Fig. 4 we plot $\Delta_c$ vs. $\theta_e$ for different values of $\theta_c$. The null model is defined with $\theta_E = 1''$. The aperture is defined to be a circular region of radius 1", and the null model has $\theta_E = 1''$.

![FIG. 4: The finite-core correction $\Delta_c$ modifying the MSD limit in the kinematics analysis (see Eqs. (29-30) and text). Here we neglect the seeing, the aperture is defined to be a circular region of radius 1", and the null model has $\theta_E = 1''$.](image)

From the calculation in Fig. 4 does not include the effect of a finite PSF, velocity anisotropy, lens ellipticity, etc. In App. C we repeat a similar calculation using a full mock system that includes all of these effects. The result of a full computation is compatible with that in Fig. 4 numerically to 50% or so.

If a real physical core component is behind the lensing $H_0$ tension, then the kinematics constraints must be considered with care, because the MSD limit could introduce large systematic errors. In general, the breaking of the MSD manifests in a smaller deviation of $\sigma^P$ from the null model: instead of $\delta\sigma^P/\sigma^P = -0.5\kappa_c$ we have $\delta\sigma^P/\sigma^P = -0.5\kappa_c (1 - \Delta_c)$, with $\Delta_c > 0$. This calls into question the kinematics analysis of some TDCOSMO systems and certainly of resolved SLACS systems in \(^{[27]}\).

Finally, while we think that the kinematics data needs to be reconsidered, this is unlikely to change the conclusion.
that a core-MSD solution for the lensing $H_0$ tension is consistent with the data. Even if we conservatively take the MSD limit, Tab. 1 shows that the TDCOSMO systems driving the tension satisfy $\delta \sigma^N/\sigma^I \gtrsim \delta H_0/H_0$ for all but PG1115, and there the inequality holds to 0.5σ or so.

V. THEORETICAL PERSPECTIVE

To explain the lensing $H_0$ tension, the ULDM soliton mass in the lens galaxy must be large enough. How much ULDM is needed, and how does this requirement compare to the soliton predicted by numerical and analytic considerations?

Numerical simulations have shown that the soliton grows by accreting ULDM from the surrounding halo via gravitational dynamical relaxation, with a characteristic timescale

$$\tau_g \approx \frac{\sqrt{2} b}{12 \pi^3 G^2 \rho \Lambda}. \quad (31)$$

Here, $\rho$, is the density of ULDM, $\sigma$ is the velocity dispersion, $\Lambda$ is the Coulomb logarithm, and the numerical factor $b \approx 0.7$ was calibrated in numerical simulations (for recent analyses, see also [40–42]). Below we will set $\sqrt{2} b \approx 1$.

A first estimate of the maximal mass of an ULDM soliton that could form in a galaxy can be obtained by calculating the ULDM mass contained inside the galactocentric radius $R_g$, within which $\tau_g(R_g) < t_{\text{gal}}$, where $t_{\text{gal}}$ is the age of the galaxy. Near this radial boundary we expect that $\rho_\chi \approx \chi \rho = \chi a R_0/\Lambda$, where $\rho$ is the total DM density (ULDM+non-ULDM) and $\rho_0$ is the background density in non-ULDM DM ($\alpha_\chi$ is the cosmological ULDM fraction defined in Eq. (2)). We can thus estimate $R_g$ from solving

$$t_{\text{gal}} \approx \frac{1}{12 \pi^3 G^2 \alpha^2 \rho \Lambda} \cdot \quad (32)$$

A rough upper bound on the mass of a soliton is then

$$M < \alpha_\chi M_{\text{halo}}(R_g). \quad (33)$$

For an isothermal power-law halo with $\rho \propto R^{-2}$ and constant $\sigma = \sigma_0$, we have $\sigma_0^2 = c_0 G M(R_0)/R_0 \approx c_0 4 \pi G \rho(R_0) R_0^2$ where we expect $c_0 \approx 1/2$. With this we have

$$R_g^4 \approx \frac{12 \pi^3 G^2 \alpha^2 \rho^2(R_0) R_0^4}{m^3 \sigma_0^2} t_{\text{gal}} \quad (34)$$

On the other hand, $M(R_g) \approx \sigma_0^2 R_g/(c_0 G)$, so using Eq. (33) the soliton upper bound reads

$$M < \left( \frac{\sigma_0^2}{m} \right)^{3/4} \left( \frac{1}{G} \right)^{3/2} \left( \frac{3 \pi}{4} M_{\text{gal}} \right)^{1/4}. \quad (35)$$

In Fig. 3 we show how the estimate of Eq. (35) compare with the imaging and $H_0$ constraints. The upper bound is shown by the green, pink, and purple bands, corresponding to $\alpha_\chi = 0.2, 0.1$, and 0.05, respectively. The upper and lower limits of each of the bands are obtained by setting $\sigma_0 = \sigma^I$ in Eq. (35) and using the upper and lower uncertainty estimates for $\sigma^I$ from Tab. 1. The age of each lens galaxy $\tau_{\text{gal}}$ in Eq. (35) is taken as the FRW time between $z = 20$ and the lens redshift $z_1$.

We truncate each constant-$\alpha_\chi$ band at small $m$ according to the cosmological constraints from Ref. [48]. We also adhere, roughly, to the limit of [52] by restricting to $\alpha_\chi \lesssim 0.2$. Inspecting the result, it is clear that the cosmological constraints on $\alpha_\chi$ play an important role in the scenario. While the imaging constraints eliminate $m > 10^{-24}$ eV or so, the combination of the dynamical relaxation consideration with the cosmological bounds [48] [52] disfavours $m \lesssim 10^{-25}$. This defines the interesting parameter space of the model to a rather narrow window.

Apart from the dynamical relaxation upper bound, another consideration comes from the saturation of the growth of the soliton: while Eq. (35) estimates the maximal amount of ULDM mass that is available for condensation into a soliton, it is possible that only a fraction of this total available mass would actually condense. The soliton growth slows from $M \propto (t/\tau_g)_1^{1/2}$ to $M \propto (t/\tau_g)_1^{1/8}$ when the specific kinetic energy of the soliton (kinetic energy per unit mass, $(K/M)$) becomes comparable to the specific kinetic energy in the surrounding halo. Both the $M \propto 1/8$ growth phase and its saturation into $M \propto 1/8$ were observed in numerical simulations [40–42], and are consistent with the soliton–host halo relation originally discovered in [33, 34], and then shown to be equivalent to $(K/M)$ equilibration in [47] [65]. The reason for this saturation is that once the $(K/M)$ threshold is crossed, the velocity dispersion at the outskirts of the soliton, and thus the dynamical timescale $\tau_g$, becomes dominated by the gravitational potential of the soliton itself. This causes $\tau_g$ to depend on $M$ with larger $M$ corresponding to larger $\tau_g$, leading to self-regulation of the growth rate.

With the parameterization of Eq. (12) we can compute the soliton specific kinetic energy,

$$\left( \frac{K}{M} \right)_\lambda = \chi^2 \int \frac{d r^2 (\partial_\chi^2)^2}{2 \int d r^2 \chi^2} \approx \chi^2 \frac{3 \sigma_0^2 \Lambda^2}{\Gamma(2 b + 1)} \frac{\Gamma(2 b - \frac{1}{2})}{\Gamma(2 b + 2) \Gamma(2 b - \frac{3}{2})}. \quad (36)$$

In the limit of low-mass soliton, where the background gravitational potential completely dominates the structure of $\chi$, $(K/M)_\lambda$ is independent of the parameter $\lambda$ because

---

5 See [64], Ch.4.3. We keep track of the constant $c_0$ here because in a realistic scenario it could vary by $O(1)$, contributing to the uncertainty in the relaxation estimate.
the ULDM profile simply reflects the wave function of an ULDM particle bound in the external potential. Indeed, using the PL external potential in this limit gives \((K/M)_\lambda \approx A/2\), consistent with the virial theorem\(^6\). We can estimate the self-regulation threshold by letting the soliton mass grow until \((K/M)_\lambda\) starts to exceed the background-dominated result.

In Fig. 5 we illustrate the growth saturation limit, computed for the system RXJ1131 from Tab. I. On the x-axis we plot \((K/M)_\lambda\) normalised to its asymptotic small-\(M\) value. On the y-axis we plot the product \(M_\lambda m_{25}\), where \(m_{25}\) corresponds to the ULDM particle mass via \(m_{25} = m/(10^{-25} \text{ eV})\). As noted above, at small \(M_\lambda\) the value of \((K/M)_\lambda\) becomes independent of \(M_\lambda\) (or equivalently, of \(A\)). As \(M_\lambda\) increases, the soliton self-gravity begins to dominate \((K/M)_\lambda\). In Fig. 5 we mark by a red dot the value of \(M_\lambda\) at which \((K/M)_\lambda\) exceeds the small-\(M\) result by 50%. From Eq. (13), we know that in the self-gravitation limit the parameter \(\lambda\) fixes the combination \(M_\lambda m\); thus, the saturation limit also fixes the combination \(M_\lambda m\). This is the reason why we use the product \(M_\lambda m_{25}\) for the y-axis in Fig. 5.

With some arbitrariness, we will estimate the growth saturation limit (roughly) by imposing, for each halo, \((K/M)_\lambda < 1.5(A/2)\), similar to the illustration in Fig. 5. The result of this calculation is shown by the thick dashed grey lines in Fig. 5.

For all of the systems of Tab. I, the growth saturation limit is weaker than or comparable to the dynamical relaxation time-scale constraint obtained with ULDM fraction \(\alpha_5 = 0.2\). This suggests that for the range of \(m\) plotted in Fig. 3 ULDM solitons are still growing in the lens galaxies, and the limiting factor for the soliton mass may be the total ULDM mass available within the dynamically relaxed region of the halo.

---

\(^6\) To be precise, the large-\(A/\lambda^2\) limit of Eq. (36) gives \((K/M)_\lambda \rightarrow 0.454 A\). The small mismatch from 1/2 can be expected given that Eq. (12) is merely an approximation for the soliton.

VI. ADDITIONAL DISCUSSION

A. Looking for a large-core soliton in near-by galaxies?

Stellar kinematics in well-resolved galaxies — including, e.g., the Milky Way (MW) itself — may provide additional constraints on ULDM. To our knowledge, the parametric region we consider here with \(m \sim 10^{-25} \text{ eV}\) and ULDM fraction \(\alpha_5 \sim 0.1\) has not been systematically studied yet.

In a MW-like galaxy, the radius of the core would fall in the dozens of kpc range (comparable to the core radius for the massive elliptical lens galaxies in the cosmography analysis). Inwards of the core radius, ULDM would make a small perturbation to the total mass budget of the galaxy, and its presence may be difficult to detect. Near the core radius, however, ULDM may become observationally relevant. Fig. 6 illustrates how a soliton satisfying the soliton–halo relation \([33, 34]\) at \(m = 10^{-24.5} \text{ eV}\) looks like in comparison to the observed kinematic mass budget of the MW. Clearly, a dedicated analysis of relevant data, notably from the GAIA mission \([66–68]\), could probe the scenario.

B. Fluctuations and dynamical heating

Ref. \([69]\) estimated the dynamical heating due to ULDM fluctuations on MW disk stars (see also \([70, 71]\)). For the case \(\alpha_5 = 1\), where all of DM is ULDM, they quote a bound \(m \gtrsim 10^{-22} \text{ eV}\) by considering the vertical velocity dispersion of thick disk stars. Formally, in an infinite system, the rate of dynamical heating scales as \(m^{-3} \alpha_5^2\), so a model with \(\alpha_5 \sim 0.1\) and \(m \sim 10^{-25} \text{ eV}\) could naively be thought to violate the bound. However, the MW is finite and in our model ULDM in the central few dozens of kpc (or even \(\sim 100\) kpc) is actually expected to be condensed in the coherent soliton (see Fig. 6). In this region the stochastic heating analysis...
of [69,71] is not valid. Instead of stochastic fluctuations, dynamical heating may still be transmitted to some extent to stars via core quasimodal mode fluctuations [32,33]. This analysis, for stellar orbits at the outskirts of the galactic disk, is yet to be done. (A related study [72] considered soliton fluctuations heating a star cluster in a dwarf galaxy. These are very different regions in ULDM parameter space and system size.)

VII. SUMMARY

The possibility of a real tension between early- and late-type determinations of $H_0$ is exciting, and could signal a breakdown of $\Lambda$CDM [12]. After all, the $\Lambda$CDM model is merely an effective theory. Gravitational lensing analyses, notably led by the TDCOSMO team, provide an important way to measure the local $H_0$. Accepting certain minimal assumptions about lens galaxy structure, the lensing analyses seem to reinforce the tension [4].

We follow up on the suggestion of Ref. [26], that a core component in the density profile of lens galaxies would behave as an approximate internal mass sheet degeneracy (MSD) and could bring the lensing $H_0$ measurement down to the CMB value, solving the lensing part of the $H_0$ tension. A preliminary test of this proposition on the data was reported in TDCOSMO-IV [27], finding a possible positive hint in the data. However, while [27] took an important step towards mitigating possible systematics related to the core-MSD proposal, they did not address the question of the physical origin of a core component.

We explored ultralight dark matter (ULDM) as a concrete, well-motivated model that could naturally produce the required cores. If ULDM exists, then it is known to produce cores ("solitons") in the centre of galaxies, due to gravitational dynamical relaxation. We studied the lensing imprint of these cores and demonstrated that they could indeed address the lensing $H_0$ tension, if the ULDM particle mass is in the ballpark of $m \sim 10^{-25} \text{ eV}$. Cosmological constraints [51, 52] imply that such light ULDM can only comprise $\lesssim 20\%$ of the total dark matter. This puts pressure on our scenario, because it limits the rate at which dynamical relaxation can operate and form the solitons. However, for ULDM abundance near this limit, the predicted cores are very close to the level required for $H_0$: clarifying this issue further would require numerical simulations that account for the background halo potential (tools of this type are already operational [33–42], but have so far been used to explore different parametric regions of ULDM).

From a theoretical perspective, the required ULDM abundance could be realised via simple vacuum misalignment for an axion-like particle with a decay constant around the grand unification or string scale.

Our study shows that strong galaxy lensing, combined with other cosmological probes like the CMB, could be sensitive to the presence of a subdominant component of dark matter in the form of ultralight fields or axions. It would be exciting if the lensing $H_0$ tension is the first hint for such fields, which could be the harbingers of otherwise inaccessible aspects of the UV theory. A promising path to test this idea is by dedicated kinematics studies, considering both massive elliptical galaxies of the type dominating the lensing analyses as well as near-by systems, including our own Milky Way.

Acknowledgments

We are grateful to N. Bar, E. Castorina, M. Simonović, and S. Suyu for comments on the manuscript, and to S. Birrer for useful comments and discussions and for guidance in using lenstronomy. This work made use of the following public software packages: lenstronomy [53, 73], emcee [74], corner [75], astropy [76, 77], and FASTELL [78]. KB was supported by grant 1784/20 from the Israel Science Foundation, and is incumbent of the Dewey David Stone and Harry Levine career development chair. LT thanks R. Porto for hospitality at DESY Hamburg. The work was supported by the International Helmholtz-Weizmann Research School for Multimessenger Astronomy, largely funded through the Initiative and Networking Fund of the Helmholtz Association.

"No hypergeometric functions were hurt in the preparation of this work.

[1] S. Refsdal, “On the possibility of determining Hubble’s parameter and the masses of galaxies from the gravitational lens effect,” *MNRAS* 128 (Jan, 1964) 307
[2] C. S. Kochanek, “What do gravitational lens time delays measure?,” *The Astrophysical Journal* 578 no. 1, (Oct, 2002) 25–32, https://doi.org/10.1086 0.1086
[3] C. S. Kochanek, “Part 2: Strong gravitational lensing,” in *Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro*, G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, and J. Wambsganss, eds., pp. 91–268, Jan, 2006.
[4] M. Millon et al., “TDCOSMO. I. An exploration of systematic uncertainties in the inference of $H_0$ from time-delay cosmography,” *Astron. Astrophys. 639* (2020) A101, arXiv:1912.08027 [astro-ph.CO]
[5] Planck Collaboration, N. Aghanim et al., “Planck 2018 results. I. Overview and the cosmological legacy of Planck,” *Astron. Astrophys. 641* (2020) A1, arXiv:1807.06205 [astro-ph.CO]
[6] DES Collaboration, T. M. C. Abbott et al., “Dark Energy Survey Year 1 Results: A Precise H0 Estimate from DES Y1, BAO, and D/H Data,” *Mon. Not. Roy. Astron. Soc. 480* no. 3, (2018) 3879–3888, arXiv:1711.00403 [astro-ph.CO]
Appendix A: Power-law background fitting formula

Here we consider how an external mass distribution affects the soliton profile. The background is taken to be a pure power-law (PL). Lensing analyses have often adopted this approximation, which leads to results for $H_0$ that are consistent with more realistic composite DM+stars halo models [4]. In a realistic analysis, the halo is axisymmetric to accommodate quad geodesics, and we include axi-symmetry when we analyse mock data in App. [B]. For simplicity, however, in modelling the impact of the external potential on the structure of the soliton we assume spherical symmetry. This approximation is justified by the disk galaxy study of Ref. [55], which showed that the soliton remains nearly spherical even with significant a-sphericity of the background.

The spherical PL density profile can be parameterised by

$$\rho_0(x) = \frac{\Sigma_c}{4\pi G D_\theta} \left( \frac{x}{D_\theta} \right)^{-\gamma}.$$  \hfill (A1)

This profile has two parameters: the PL slope $\gamma$ and the normalization, fixed here by $D_\theta$. (For a lensing model containing the PL $\rho_0$ alone, the parameter $D_\theta$ would match the observable Einstein angle $\theta_E$. This is no longer true once we consider composite models as in Eq. (3).) The values of $D_\theta$ and $\Sigma_c$ are fixed by the system redshift and cosmology. To simplify matters further we set $\gamma = 2$, close to the slopes inferred for the galaxies in Tab. [1]. The external potential entering Eq. (8) is then given by:

$$\Phi_{\text{ext}}(r) = A \ln r + C,$$  \hfill (A2)

$$A = 2\Sigma_c D_\theta.$$  \hfill (A3)

Note that the factor $\Sigma_c D_\theta = D_\theta / (4\pi G D_\theta)$ is independent of $H_0$. To gain some physical intuition, note that if we define $M_{PL}(1/m)$ as the mass included in the PL profile up to a distance equal to the ULDM Compton radius $1/m$, then $A = GM_{PL}(1/m)m$. Conveniently, for $\gamma = 2$, $M_{PL}(1/m)m$ is independent of $m$. The constant C in Eq. (A2) is unimportant.

Because $\Phi_{\text{ext}}$ breaks the scale invariance of the self-gravitating soliton, the coefficients $a$ and $b$ in the approximation of Eq. (12) now depend on the combination $A/\lambda^2$. We find that Eq. (12) still provides a good fit for any value of $A/\lambda^2$, with the fitting formula:

$$a(z) = 0.23 \sqrt{1 + 7.5 z \tanh(1.5z^{0.24})},$$  \hfill (A4)

$$b(z) = 1.69 + \frac{2.33}{(1 + 2.2z)^{1.24}},$$  \hfill (A5)

where $z = A/\lambda^2$.

Appendix B: Power-law background: mock analysis

Here we use the gravitational lens model software package lenstronomy [53] to study the core-MSD soliton model in mock data analysis. Our main purpose is to check how well the simple $\delta_E$ imaging error criterion described in Sec. [11] (see Eqs. (15) and (19)) captures the observational constraints on the model. In addition, the implementation of the soliton core module in lenstronomy would be useful to test the model directly against data in forthcoming work.

The mock data is as follows. The truth model has the convergence of Eq. (3), with $\kappa_0$ given by an elliptic PL profile (so as to produce a quad image) and $\kappa_c = \kappa_\lambda = 1$ of an ULDM soliton with $m = 10^{-25} \, eV$ and $M = 1.4 \times 10^{12} \, M_\odot$. The parameters are chosen to produce an effective $\kappa_\lambda(\theta_E) \approx 0.1$ and $\theta_E \approx 10''$. The truth value of $H_0$ is set to $H_0 = 67.4 \, km/s/Mpc$, mimicking the CMB result [5]. In figure 7 we show the mock alongside a reconstructed image, done by running the MCMC using the core-MSD model with a Gaussian prior on $H_0$ set at its truth value.
To demonstrate the outcome of using an inference model which does not include a core component (the case of, e.g. [4, 54–58]), we run the MCMC using a pure (elliptic) PL. Fig. 8 shows the posterior triangle plot obtained for this model. As expected, the MCMC converges to $H_0 \approx 75$ km/s/Mpc, in a good fit without detectable imaging residuals. A lensing analysis that does not utilise the core-MSD model would converge to this biased result.

In the top panel of Fig. 9 we re-run the MCMC, this time using the core-MSD model in the inference. (For convenience in the implementation, we use $1/\theta_e$ and $\kappa_\Lambda(\theta_E)$, rather than $m$ and $M$, as the sampling parameters in the fit.) The MSD leads to a significant broadening of the $H_0$ posterior, corresponding to the $\kappa_\Lambda(\theta_E)$-H$_0$ degeneracy. A low value of $H_0 \approx 60$ km/s/Mpc, accompanied by an $M \approx 10^{12}$ M$_\odot$ soliton at $m \lesssim 10^{-25}$ eV, produces a comparably good fit as the original $H_0 \approx 75$ km/s/Mpc model with a vanishing soliton (Fig. 8).

In the bottom panel of Fig. 9 we repeat the exercise, this time adding an external CMB prior on $H_0$ coincident with the truth value of the mock. The posterior now converges to an upper limit of $1/\theta_e \approx 0.13/(1^\prime)$ at 95% CL. This, together with the most probable value for $\kappa_\Lambda(\theta_E)$, correspond to $M \approx 9 \times 10^{10}$ M$_\odot$ and $m \approx 2 \times 10^{-25}$ eV.

To study how well Eq. (19) approximates realistic imaging constraints on the soliton, in Fig. 10 we show $\delta_E$ as a function of $1/\theta_e$, computed using Eq. (4) for a specific value of $\kappa_\Lambda(\theta_E)$. In this calculation, $\alpha(\theta_E)$ entering Eq. (4) is the deviation angle of the full core-MSD model, computed at a fixed angle corresponding to the peak posterior value of $\theta_E$ found in the pure PL MCMC run of Fig. 8. In green, we show the value given by Eq. (19). We see that Eq. (19) leads to a bound on $\theta_e$ which is a factor of 2 or so stronger (that is, more conservative) than the MCMC bound.

**Appendix C: MSD-breaking kinematics correction**

In Fig. 11, we show the MSD-breaking kinematics correction $\Delta c$, computed semi-analytically (see Sec. IV).
for model parameters mimicking the nine SLACS systems of [27] with resolved kinematics data (see Fig.15 in [27]).

The values of $\theta_c$ and $\theta_E$ for these system are taken from Tab. E1 in [27]; from left to right, the systems in the plot are: SDSSJ1627-0053, SDSSJ2303+1422, SDSSJ1250+0523, SDSSJ1204+0358, SDSSJ0037-0942, SDSSJ0912+0029, SDSSJ2321-0939, SDSSJ0216-0813, SDSSJ1451-0239. Circle, triangle, and diamond markers correspond to the angular bins $(0''',1''), (1'',2''), (2'',3''$, respectively. In the left panel we use the core toy model of [27]. In the right, we repeat the exercise for the physical ULDM soliton model. Both models are defined with $\theta_c = 10\theta_E$.

In computing the effect for the core model of [27], we use the fact that the density profile in this model matches (the square of) Eq. (12). It follows that Eqs. (26-27) are still valid for this model. The only adjustment needed is to set $b = 3/4$ for the toy model (compared to $b \approx 3.9$ for a self-gravitating soliton). The parameter $\theta_c$ has the same role in both cases. Fig. 11 shows that for small apertures, the toy $\Delta_c$ is roughly half that of a soliton defined at the same $\theta_c$.

We can calculate $\Delta_c$, numerically, including effects like velocity anisotropy, lens ellipticity, PSF, and realistic aperture definitions that were lacking above and in Sec. IV. Fig. 12 shows a full numerical computation of $\Delta_c$, calculated directly from the definition Eq. (29). The mock is defined with $\theta_E = 1.2''$, compared to $\theta_E = 1''$ in Fig. 4. This means that if the PSF, aperture, anisotropy, and axi-symmetry effects were not important, we would expect $\Delta_c$ computed from the mock in Fig. 12 to be smaller by a factor $\approx 0.83$ compared to Fig. 4. In practice, with all of the above effects included, $\Delta_c$ in Fig. 12 is slightly larger. The parametric dependence on $\theta_c$ and the rough size of the effect are well reproduced. Lastly, we verified that the full numerical procedure coincides very accurately (to $O(1\%)$) with the analytical calculation when lens ellipticity and velocity anisotropy are set to zero.

FIG. 9: Lensing analysis for the same mock data as in Fig. 8, this time adding a core component to the fit. The blue lines correspond to the true values used for the mock. Top: Model inference with flat prior on $H_0$. We remark that the mock is consistent with a no-core solution; hence, the median value shown for $1/\theta_c$ is an artefact of the finite range of the prior. The MSD is manifest by the broadening of the $H_0$ posterior distribution. Code: © Bottom: Same as in the top panel, this time adding a CMB prior on $H_0$. As expected, the no-core solution is now disfavoured. Code: ©

FIG. 10: Triangle plot of $\delta_E$ and $1/\theta_c$, calculated for a fixed $\kappa_\lambda(\theta_E)$. In green, the $\delta_E$ coming from Eq. (19). Notice that the maximum $\delta_E$ allowed for our mock is $|\delta_E| \lesssim 0.0008$ (horizontal red lines). Combining this with Eq. (19) would yield the naive bound $\theta_c \gtrsim 25''$ (blue vertical line, marking the intersection of the green and red lines). However, the region explored by the MCMC suggests that the more realistic bound is somewhat weaker, $\theta_c \gtrsim 10''$. Code: ©
FIG. 11: MSD-breaking kinematics correction $\Delta_c$, computed semi-analytically (see Sec. [IV]) for model parameters mimicking the nine SLACS systems with resolved kinematics (see Fig.15 in [27]). Circle, triangle, and diamond markers correspond to the angular bins $(0'', 1'')$, $(1'', 2'')$, $(2'', 3'')$, respectively. **Left:** for the toy model of [27]. **Right:** for the ULDM soliton. Both models are defined with $\theta_c = 10\theta_e$.

FIG. 12: MSD-breaking kinematics correction $\Delta_c$, computed numerically for the mock system including velocity anisotropy, lens ellipticity, PSF effects, and realistic aperture definitions. Compare this result to the semi-analytic result in Fig. [IV].