Free-form Lens Model and Mass Estimation of the High-redshift Galaxy Cluster
ACT-CL J0102-4915, “El Gordo”

Jose M. Diego1, S. M. Molnar2, C. Cerny3, T. Broadhurst4,5, R. Windhorst6, A. Zitrin7, R. Bouwens8, D. Coe9, and K. Sharon3

1 Institut de Física de Cantabria (CSIC-UC). Avda. Los Castros s/n. E-39005 Santander, Spain; jdiego@ifca.unican.es
2 Institute of Astronomy and Astrophysics, Academia Sinica, P.O. Box 23-141, Taipei 10617, Taiwan
3 Department of Astronomy, University of Michigan, 1085 South University Avenue, Ann Arbor, MI 48109, USA
4 Department of Theoretical Physics, University of the Basque Country, Bilbao E-48080, Spain
5 Ikerbasque, Basque Foundation for Science, Alameda Urquijo, 36-5 Plaza Bizkaia E-48011, Bilbao, Spain
6 School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA
7 Department of Physics, Ben-Gurion University, Be’er Sheva 84105, Israel
8 Leiden Observatory, Leiden University, NL-2300 RA Leiden, The Netherlands
9 Space Telescope Science Institute, Baltimore, MD, USA
10 Centre for Astronomy and Particle Theory, School of Physics & Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK

Received 2019 December 13; revised 2020 October 5; accepted 2020 October 6; published 2020 November 25

Abstract
We examine the massive colliding cluster El Gordo, one of the most massive clusters at high redshift. We use a free-form lensing reconstruction method that avoids making assumptions about the mass distribution. We use data from the RELICS program and identify new multiply lensed system candidates. The new set of constraints and free-form method provide a new independent mass estimate of this intriguing colliding cluster. Our results are found to be consistent with earlier parametric models, indirectly confirming the assumptions made in earlier work. By fitting a double gNFW profile to the lens model and extrapolating to the virial radius, we infer a total mass for the cluster of $M_{200c} = (1.08 \pm 0.12) \times 10^{15} M_{\odot}$. We estimate the uncertainty in the mass due to errors in the photometric redshifts and discuss the uncertainty in the inferred virial mass due to the extrapolation from the lens model. We also find in our lens map a mass overdensity corresponding to the large cometary tail of hot gas, reinforcing its interpretation as a large tidal feature predicted by hydrodynamical simulations that mimic El Gordo. Finally, we discuss the observed relation between the plasma and the mass map, finding that the peak in the projected mass map may be associated with a large concentration of colder gas exhibiting possible star formation. El Gordo is one of the first clusters that will be observed with JWST, which is expected to unveil new high-redshift lensed galaxies around this interesting cluster and provide a more accurate estimation of its mass.

Unified Astronomy Thesaurus concepts: Gravitational lensing (670); Dark matter (353); High-redshift galaxy clusters (2007); Rich galaxy clusters (2005)

1. Introduction
The galaxy cluster ACT-CL J0102-4915, also known as El Gordo, is a relatively high-redshift $z = 0.870$ with a rich, bimodal galaxy distribution (Williamson et al. 2011; Menanteau et al. 2012). Its mass has been estimated in earlier work using different techniques, including combined dynamical, X-ray, and Sunyaev–Zel’дович (SZ) data (Menanteau et al. 2012); strong lensing data (Zitrin et al. 2013; Cerny et al. 2018); and weak lensing data (Jee et al. 2014). El Gordo has also been the subject of several dynamical studies, including numerical N-body simulations (Donnert 2014) and hydrodynamical simulations (Molnar & Broadhurst 2015; Zhang et al. 2015, 2018). These studies have highlighted the impressively large-scale cometary structure clearly visible in X-ray images, which appears to imply that El Gordo is being observed right after a collision of two subgroups (Molnar & Broadhurst 2015; Zhang et al. 2015), similar to the iconic Bullet cluster. This interpretation is supported by the presence of two radio relics ahead and behind the X-ray cometary structure (Lindner et al. 2014; Molnar & Broadhurst 2015). Based on the X-ray and radio morphology, as well as on a preliminary lens model for the mass distribution, Ng et al. (2015) argued that El Gordo is in a return phase after first core passage. This means that the cluster is being observed after the phase of apocenter, and that the two groups are moving toward each other rather than away from each other. Part of this conclusion is based on a lens model that relies on weak lensing and assigns more mass to the NW clump than the SE clump. This interpretation is, however, challenged by lens models based on strong lensing data that place most of the mass in the SE group (Zitrin et al. 2013; Cerny et al. 2018).

The El Gordo cluster is an extreme cluster at several levels. It is the most massive cluster at $z \approx 0.9$ with an estimated mass ranging from $M_{200c} = 1.8 \times 10^{15}$ to $2 \times 10^{15} M_{\odot}$. Here $M_{200c}$ is the mass within the sphere of radius $r_{200}$. This radius is defined as the radius where the mass density enclosed in the sphere with the same radius and centered in the object is 200 times the critical density of the universe at the cluster redshift. Some authors estimate the overdensity in relation to the average density of the universe, $\rho_\text{m} = \rho_c \Omega_m$. In this case, the mass is denoted as $M_{200\rho_m}$. Using SZ data, Williamson et al. (2011) estimated a mass of $M_{200\rho_m} = 1.89 \pm 0.45 \times 10^{15} M_{\odot}$. Menanteau et al. (2012) obtained a mass estimate of $M_{200\rho_m} = 2.16 \pm 0.32 \times 10^{15} M_{\odot}$, which relied on different scaling relations. Based on an extrapolation of the strong lensing mass model, Zitrin et al. (2013) estimated a total mass of $M_{200\rho_m} \approx 2.3 \times 10^{15} M_{\odot}$. Using weak lensing obtained with the Hubble Space Telescope (HST), Jee et al. (2014) estimated $M_{200\rho_m} = 3.13 \pm 0.56 \times 10^{15} M_{\odot}$.

At the high end of the mass range for El Gordo, these masses are in tension with the standard ΛCDM model (see, for instance,
Jee et al. (2014), which predicts that the maximum virialized mass at this redshift should be less than $M_{200m} \approx 1.7 \times 10^{15} M_\odot$ (Harrison & Coles 2012). A similar conclusion is reached from large N-body simulations. Using the very large 630 Gpc$^3$ N-body simulation Jubilee (based on a standard ΛCDM model), Watson et al. (2014) found that the most massive cluster in the simulation and at $z = 0.9$ is $M_{200m} \approx 1.5 \times 10^{15} M_\odot$ (see their Figure 5). Note that in Watson et al. (2014), the masses are defined as $M_{1780k}$ rather than $M_{200m}$ or $M_{500c}$. For a Navarro–Frenk–White (NFW) profile, $M_{1780k} \approx 1.2 M_{200m}$ and is about 4% times higher than $M_{200m}$ (Waizmann et al. 2012). Even though El Gordo is still in the process of merging, its final virial mass is expected to be well approximated by the sum of the masses of the two subclusters. Given the fact that El Gordo was found in a relatively small area of the sky (the footprint of the original ACT survey covers less than 2% of the area of the sky), it raises the question of its significance as possible evidence for tension with the ΛCDM model. How likely is it to find such an extreme object in an area smaller than 2% of the sky? The tension could be reduced if previous mass estimates are found to be too high. Additional mass estimates, based on alternative methods, can explore the uncertainty in the mass of the cluster. In this paper, we present such an alternative estimate based on the first free-form lensing modeling of this cluster. Free-form models are affected by different systematics than parametric models. Hence, it is important to estimate the mass of extreme clusters (like El Gordo) with as many independent estimators as possible. We note that, after correcting for redshift and Eddington bias, Waizmann et al. (2012) found El Gordo not to be in tension with ΛCDM. However, in their analysis, the survey area is assumed to be 3.7 times higher than the area where El Gordo was originally found.

As of 2014, El Gordo was also the highest-redshift cluster known to host radio relics (Lindner et al. 2014). The X-ray emission exhibits an interesting offset between the peak of the X-ray emission and the position of the Brightest Cluster Galaxy (BCG). Contrary to what happens in the Bullet cluster, the X-ray peak seems to be ahead of the BCG. However, in the interpretation of Ng et al. (2015), the BCG would be moving toward the second group, so the X-ray peak would be trailing the BCG. The returning phase interpretation of El Gordo is challenged based on results from dedicated N-body/hydrodynamical simulations reproducing most of the observations of El Gordo (Molnar & Broadhurst 2015; Zhang et al. 2015). Molnar & Broadhurst (2018) demonstrated that the speed of the outgoing shocks can be very large (4000–5000 km s$^{-1}$) in a massive merging cluster like El Gordo, thus leaving the system before the first turnaround.

El Gordo is also unique in the sense that it is a powerful lens at relatively high redshift. One of the features that makes El Gordo an attractive target for lensing studies is the fact that for sources at high redshift, critical curves form at relatively large distances from the member galaxies. This is particularly true in the gap between the two clusters, where the critical curves are relatively undisturbed by nearby member galaxies. Having undisturbed critical curves is relevant to observe caustic crossing events of distant stars (Diego et al. 2018; Kelly et al. 2018), since in this case, the maximum magnification can be larger than in situations where critical curves are affected by microlenses in member galaxies or from the intracluster medium. Caustic crossing events have been proposed as a technique useful to study Population III stars and stellar-mass black hole accretion disks in Windhorst et al. (2018) with the James Webb Space Telescope (JWST). Because El Gordo is the highest-redshift known cluster with potentially such significant transverse motion—based on the X-ray morphology and the two lensing mass centers discussed in this paper—it is an ideal target for JWST follow-up to search for caustic transits at $z > 1$ and possibly for first-light caustic transits at $z > 7$. For this reason, El Gordo is a JWST GTO target that will be observed in Cycle 1 (JWST program 1176; PI: Windhorst). It is our sincere hope that JWST guest observers will propouse to observe El Gordo in many successive epochs, among other clusters to find caustic transits at $z > 1$.

In this paper, we derive the mass distribution and study this interesting cluster using the latest public data from the Reionization Lensing Cluster Survey (RELICS) program and newly identified strong lensing system candidates. We use our free-form lensing reconstruction code WSLAP+ (Diego et al. 2005, 2007; Sendra et al. 2014; Diego et al. 2016), which does not rely on assumptions about the distribution of dark matter. Our results offer an important cross-check with previous results, since any disagreement between our free-form method and results obtained by previous parametric methods could signal potential systematic problems in one (or both) types of modeling. We pay special attention to the integrated mass as a function of radius and the effect that extrapolations of the derived mass profile up to the virial radius have on the inferred total mass of the cluster.

This paper is organized as follows. In Section 2, we describe the data and simulations used in this work. In Section 3, we briefly describe the algorithm used to perform the lensing reconstruction. Results are presented in Section 4 and discussed in Section 5. We summarize and conclude in Section 6. We adopt a standard flat cosmological model with $\Omega_m = 0.3$ and $h = 0.7$. At the redshift of the lens and for this cosmological model, 1″ corresponds to 7.8 kpc.

2. Observational and Simulated Data

In this section, we briefly describe the data used in this work, as well as previous N-body simulations of the El Gordo cluster, that will be used to compare with our results.

2.1. Optical Data

We use public Hubble imaging data from programs GO 12755 (PI: J. Hughes), GO 12477 (PI: F. High), and GO 14096 (PI: D. Coe). These Advanced Camera for Surveys (ACS) and WFC3/IR observations include data in 10 filters spanning wavelengths of ~0.4–1.7 μm. RELICS (Coe et al. 2019) delivered reduced images combining data from all of these HST programs, including their own (14096). RELICS also delivered photometric redshift catalogs of objects detected in these images. We retrieved these data products from the Mikulski Archive for Space Telescope (MAST). From the reduced images, we produce color images by combining the optical and IR bands. A parametric lens model derived using the new RELICS (and previous HST) data is presented in Cerny et al. (2018).

2.2. X-Ray Data

To explore the dynamical state of El Gordo, we also produce an X-ray image using public Chandra data. In particular, we use data from the ACIS instrument acquired in 2011–2012 and with ObsIDs 12258, 14022, and 14023 (PI: J. Hughes) totaling ~250 ks. The X-ray data are smoothed using the code ASMOOTH (Ebeling et al. 2006). Figure 1 shows a composite...
image made after combining different exposures from the Hubble Frontier Fields (HFF), and overlaying the contours of the smoothed X-ray data. The distribution of X-rays shows a cometary structure with a bright region near the BCG and a double tail aligned in the NW direction, like in the Bullet cluster that also shows a cometary structure. The peak of the X-ray emission is offset with respect to the BCG by \( \approx 70 \) kpc.

2.3. Simulated Data

In order to study the impact of extrapolating a strong lensing-derived profile (i.e., a profile that extends a limited range in radii) up to the virial radius, we use dedicated N-body/hydrodynamical simulations that mimic El Gordo (for details of the simulations, see Molnar & Broadhurst 2015). Our simulations were constrained by multifrequency data: X-ray, radio (SZ observations), and optical (for gravitational lensing and dynamics). Our best simulated cluster (in terms of reproducing the observed morphology in El Gordo) assumed initial total masses of \( 1.4 \times 10^{15} \) and \( 7.5 \times 10^{14} \ M_\odot \) for the main and infalling cluster, respectively; an impact parameter of 300 kpc; and a relative initial infall velocity of 2250 km s\(^{-1}\) when separated by the sum of the two virial radii. This model explains most of the observed features of El Gordo: the distinctive cometary feature with a twin-tailed wake (see Figure 2 in Molnar & Broadhurst 2015) observed in the X-ray morphology, the locations of the two peaks of the dark matter components, and the position of the SZ peak. In this paper, we use the total mass distribution from our best model to derive the surface mass distribution and compare that to the surface mass distribution we derived from gravitational lensing. The initial conditions of the simulations are comparable to earlier attempts to simulate this cluster in Donnert (2014), except for the impact parameter that is larger in Molnar & Broadhurst (2015; \( \approx 300 \) kpc versus \( \approx 20 \) kpc. Other parameters, like core radius and concentration parameter, are also slightly different. Using a larger impact parameter, Molnar & Broadhurst (2015) were able to reproduce the observed twin-tailed cometary structure that is missing in Donnert (2014).

2.4. Multiply Lensed Galaxies

In this section, we discuss the lensed systems used to constrain the mass model. The identification of multiply lensed galaxy systems in El Gordo is particularly challenging because no multiply lensed galaxies with spectroscopic redshifts have been confirmed so far, despite the recent attempts made with LDSS3 in Magellan by Cerny et al. (2018). However, with the increased depth of the new RELICS data, we can improve upon previous identifications. As a default, we adopt the original
naming scheme of Zitrin et al. (2013) for the lensed system candidates, who identified the first families of strongly lensed galaxies, and performed the first strong lensing analysis of this cluster based on the three HST bands available at the time (the compilation of systems is detailed in Appendix A). Notably, two systems (1 and 2 in Figure 1) exhibit well-resolved morphological features that, together with robust photometric redshifts, allow one to unambiguously confirm these two families of images. Systems 3, 4, and 5 also contain morphological information and reliable photometric redshifts, which makes the identification of these systems equally robust. Systems 1–4 are similar to the systems defined in Zitrin et al. (2013) and Cerny et al. (2018). We note that systems 10 and 20 in Cerny et al. (2018) are part of our systems 1 and 2, where we identify different knots in the systems that are used as additional constraints. Our new system 5 was also independently identified by Cerny et al. (2018) as system 13 in their work. In the original scheme of Zitrin et al. (2013), system 4 was composed of three tangential counterimages and two possible radial images. A preliminary model clearly disfavors radial image 4.2 and candidate image 4.3 (in Zitrin et al. 2013) as part of system 4. We note that Cerny et al. (2018) also rejected these two counterimages as part of system 4 (as do some updated models by Zitrin et al. 2020, private communication). Instead, we suggest that the two alleged arclets, 4.2 and 4.3, in Zitrin et al. (2013) are likely features in the galaxy cluster associated with the cooling of the plasma (see Section 5.2). Using the robust systems 1–5, we derive a first model based on these reliable systems. This model is later used to unveil new system families (listed in Table 1). We refer to this first model for the mass distribution as the driver model.

Table 1 lists all of the arclets, including some candidates listed here for completeness but not used in the lens reconstruction. Also for completeness, we include system 8 as originally defined in Zitrin et al. (2013). The driver model disfavors this system, so we do not include it in our lens reconstruction. Also, Cerny et al. (2018) discarded this system. The systems in Table 1 are divided into three categories: A, B, and C. Arclets in category A are robustly confirmed based on their color, morphology, and photometric redshift. As mentioned above, we use these arclets to derive the driver model. Systems in category B are highly compatible with the driver model. In addition, color and, when available, morphology and photometric redshift are also consistent among the different members of the same family of images. Note that in some cases, the images are unresolved, so they lack detailed morphological information. Also, gaps in the CCDs or masked regions in some bands may result in color artifacts. The systems labeled A and B are used to derive an alternative model that we name the full model. Arclets in category C are still consistent with the driver model, but lack of morphological information, a mismatch in the alignment of the predicted image (compared with the observed one), or tension between the predicted and observed magnification ratios reduces the reliability of the identification. Arclets marked with label C are not used in the mass reconstruction but are still included in Table 1. Future data will confirm or reject these system candidates. The systems in Table 1 that are new identifications are marked in bold. Systems that were fully included in previous work are indicated in the Comments column. Our new system 6 has an estimated redshift (from the lens model) of $z \approx 4.3$, which is consistent with the photometric redshift. For system 7, we identify a new candidate for 7c that differs from the candidate in Zitrin et al. (2013). System 10 is a new system with a photometric redshift of $5.1$ (for 10a). The driver model is fully consistent with this system and redshift. System 11 is a new redefinition of system 5 in Zitrin et al. (2013). The driver model suggests that the big arclet forming part of system 5 in Zitrin et al. (2013) consists of two images merging at the critical curve. The corresponding third image is identified with the tail of a bright galaxy that may itself be lensed by another member galaxy. Based on the driver model, the redshift of this galaxy should be $z \approx 3.1$, while the photometric redshift for the arc is $z \approx 2.2$. When this system is included in the lens reconstruction (i.e., in the full model), we adopt the photometric redshift for this system. System 12 is a redefinition of system 14 in Cerny et al. (2018), based on the driver model and color+morphological information. Our 12a matches 14a in Cerny et al. (2018), but we identify two different counterimages. The driver model predicts a redshift of $z \approx 3$, consistent with the photometric redshift of $z = 3.4$ of 12a. System 13 has no photometric redshift. The driver model predicts a redshift of $z = 3$ for this system. System 14 has a photometric redshift of $z = 2.7$ (14a), but we adopt the redshift predicted by the driver model, i.e., $z = 4$, for this system. System 15 corresponds to system 8 in Cerny et al. (2018), which was also independently identified. Both the photometric redshift ($z = 2.7$) and the redshift predicted by the driver model ($z = 2.65$) agree reasonably well. System 17 is a newly identified system with two identifiable knots and a redshift consistent with the nearby system 3, so systems 3 and 17 may be forming a close pair of galaxies at $z \approx 5$. Systems 18 and 19 are all new candidates, but a lack of morphological features does not allow us to confirm their association based on the morphology of the predicted images.

Finally, we do not consider system 5 in Cerny et al. (2018). Although the driver model is consistent with the positions of system 5 in Cerny et al. (2018), a third image is clearly predicted but not observed, casting doubt on the feasibility of this system. However, we should note that it is also possible that the driver model fails at correctly predicting the position of the third counterimage, this image could be hidden underneath one of the bright member galaxies, or a small unobserved substructure close to the predicted position of this third image demagnifies this image.

We emphasize that, although no spectroscopic redshifts are yet available for this cluster, the photometric redshifts of the systems used to derive the driver model show a high degree of consistency and relatively small scatter. The driver model can then geometrically predict the redshifts of the remaining system candidates (see, for instance, Chan et al. 2020, for an application of this method). Future spectroscopic observations will confirm some of these systems and/or discard others but will also allow testing of the blind predictions made by the driver model, hence validating the process of estimating redshifts in a geometric way. It is also important to note that since we use photometric redshifts, errors in the resulting lens model are expected to be larger than when spectroscopic redshifts are available. As shown by Johnson & Sharon (2016), in situations where no spectroscopic redshifts are available, errors up to 10% are expected. Interestingly, the recent result of Caputi et al. (2020; which appeared during the revision of this work) presents the first spectroscopic measurements of lensed

---

11 Both systems 3 and 17 were recently spectroscopically confirmed in Caputi et al. (2020), confirming their physical association.
systems in this cluster. In particular, systems 3 and 17 are confirmed by MUSE observations and with spectroscopic redshifts consistent with the photometric estimates.

3. Formalism

The mass reconstruction is based on our method WSLAP+. The reader can find the details of the method in our previous papers (Diego et al. 2005, 2007; Sendra et al. 2014; Diego et al. 2016). Here we give a brief summary of the most essential elements.

The lens equation is defined as follows:

$$\beta = \theta - \alpha(\theta, \Sigma),$$  \hspace{1cm} (1)

where $\theta$ is the observed position of the source, $\alpha$ is the deflection angle, $\Sigma(\theta)$ is the surface mass density of the cluster at the position $\theta$, and $\beta$ is the position of the background source. Both the strong and weak lensing observables can be expressed in terms of derivatives of the lensing potential,$^{12}$

$$\psi(\theta) = \frac{4GDlDs}{c^2Di} \int d^2\theta' \Sigma(\theta') \ln(|\theta - \theta'|),$$  \hspace{1cm} (2)

where $D_l$, $D_s$, and $D_i$ are the angular diameter distances to the lens, to the source, and from the lens to the source, respectively. The unknowns of the lensing problem are, in general, the surface mass density and the positions of the background sources in the source plane. The surface mass density is described by the combination of two components: (i) a soft (or diffuse) component (usually parameterized as a superposition of Gaussians) and (ii) a compact component that accounts for the mass associated with the individual halos (galaxies) in the cluster.

For the diffuse component, different bases can be used, but we find that Gaussian functions provide a good compromise between the desired compactness and smoothness of the basis function. A Gaussian basis offers several advantages, including a fast analytical computation of the integrated mass for a given radius, a smooth and nearly constant amplitude between overlapping Gaussians (with equal amplitudes) located at the right distances, and an orthogonality between relatively distant Gaussians that help reduce unwanted correlations. For the compact component, we directly adopt the light distribution in the IR band (F160W) around the brightest member galaxies in the cluster. For each galaxy, we assign a mass proportional to its surface brightness. This mass is later readjusted as part of the optimization process.

As shown by Diego et al. (2005, 2007), the strong and weak lensing problem can be expressed as a system of linear equations that can be represented in a compact form,

$$\Theta = \Gamma X,$$  \hspace{1cm} (3)

where the measured strong lensing observables (and weak, if available) are contained in the array $\Theta$ of dimension $N_0 = 2N_d$, the unknown surface mass density and source positions are in the array $X$ of dimension $N_X = N_i + N_g + 2N_s$, and the matrix $\Gamma$ is known (for a given grid configuration and fiducial galaxy deflection field) and has the dimensions $N_0 \times N_X$.

Here $N_i$ is the number of strong lensing observables (each one contributing with two constraints, $x$ and $y$), and $N_g$ is the number of grid points (or cells) that we use to divide the field of view. Each grid point contains a Gaussian function. The widths of the Gaussians are chosen in such a way that two neighboring grid points with the same amplitude produce a horizontal plateau in between the two overlapping Gaussians. In this work, we consider only regular grid configurations. Irregular grids are useful when there is a clear peak in the mass distribution, for instance, when the cluster has a well-defined center or a single BCG. Here $N_i$ (in Equation (4)) is the number of deflection fields (from cluster members) that we consider. It can be seen as the number of mass layers, each one containing one or several galaxies at the distance of the cluster. In this work, we set $N_i$ equal to 1, i.e., all galaxies are forced to have the same mass-to-light ratio.

Finally, $N_l$ in Equation (4) is the number of background sources (each contributes with two unknowns, $\beta_x$ and $\beta_y$), which in our particular case ranges from $N_l = 5$ when only the subset of reliable systems is used (driver model in Section 2) to $N_l = 16$ when all systems labeled A or B in Table 1 are used in the reconstruction (full model). The solution, $X$, of the system of Equation (3) is found after minimizing a quadratic function of $X$ (derived from the system of Equation (3) as described in Diego et al. 2005). The minimization of the quadratic function is done with the constraint that the solution, $X$, must be positive. The vector $X$ contains the three types of unknowns in the problem, i) the grid masses, ii) the renormalization factors for the galaxy deflection field, and iii) the background source positions. All these quantities are always positive (the zero of the source positions is set to the bottom left corner of the field of view). Imposing $X > 0$ helps constrain the space of meaningful solutions and regularize the solution, as it avoids unwanted large negative and positive contiguous fluctuations. The quadratic algorithm convergence is fast (a few minutes on a standard laptop), allowing for multiple solutions to be explored in a relatively short time. Different solutions can be obtained after modifying the starting point in the optimization and/or the redshifts of the systems without spectroscopic redshift. A detailed discussion of the quadratic algorithm can be found in Diego et al. (2005). For a discussion of its convergence and performance (based on simulated data), see Sendra et al. (2014). In order to study the dependency of the lens model with the uncertainty in the photometric redshifts and the initial condition, we also derive 100 lens models with the settings of the driver model (i.e., using A systems only). For each of the models, we vary the initial condition in the optimization process and the redshift of all five background sources. For the redshift, we adopt a Gaussian error of 0.5 for all five systems, which is consistent with the dispersion in photometric redshifts for each system as shown in Table 1.

4. Results

Thanks to the new RELICS data, we can revise the multiple images identified in this cluster and assign them a rank ranging from A (most reliable) to C (least reliable). Based on the set of images ranked A (see Table 1), we derive the driver model, which is later used to uncover new multiple images or reveal issues with previous identifications. Even though the driver model is based on a relatively small subset of only five families of images, the relatively uniform spatial distribution of these five families allows us to derive a relatively reliable lens model.
The driver model disfavors the radial counterimage candidates 4.2 and 4.3 in Zitrin et al. (2013; these images were also discarded by Cerny et al. 2018); instead, we suggest that these may be signatures of cooling flows or jets near the BCG (see Section 5.2 for discussion). System 8 in Zitrin et al. (2013) shows a relatively good consistency with the driver model in terms of predicted versus observed positions, but the morphology of the observed images does not match well with the predicted morphology, so we do not use this system in any of our lensing reconstructions as well (this system is still included in Table 1 for completeness). Some of the counterimages postulated in earlier work as candidates (for instance, 7c and 9c) are in general consistent in terms of position, but their morphology is not well reproduced by the lens model. We also unveil new image candidates, some of them independently identified in Cerny et al. (2018).

In addition to these, we identify additional new families as described in Appendix A. System 15 in Cerny et al. (2018) is consistent with the driver model, but a third image is clearly predicted by the driver model and not observed. Consequently, we do not use this system in our reconstruction, although we should note that the predicted position for the third image is only a few arcseconds from the BCG. Hence, it is possible that the driver model is not accurate enough around the BCG and that the third image lurks behind the bright BCG, with a smaller magnification than the one predicted by the driver model. A smaller magnification is possible if the BCG has a larger mass-to-light ratio in the central region, for instance, through a central spike in the mass distribution, or if a supermassive black hole is at or near the center.

Based on the driver model, we expand the number of reliable systems and estimate their redshifts based on the available photometric redshift information and/or the redshift predicted by the driver model. Using the expanded set of systems (ranked A and B in Table 1), we derive a new model, which we refer to as the full model. The mass distributions of the two models are compared in Figure 2. For these plots, we have subtracted the contribution from the member galaxies to better show the diffuse component. The two models look similar to first order, but some differences can be appreciated, especially around the BCG, where the full model places the peak of the diffuse component at several arcseconds from the BCG. In particular, the peak of the soft component correlates very well with the peak in the X-ray emission. Similar correlations between the diffuse component obtained by our method and the observed X-ray emission were found in earlier work, where we discussed the possibility that the lensing data are also sensitive to the mass of the hot gas. This possibility is discussed in more detail in Section 5.2. Note that in massive clusters, the hot plasma is expected to contribute to \( \approx 10\% \) of the total mass, on average.

In terms of integrated mass, both models also look similar, but with the full model having slightly more mass, especially in the SE group, where the X-ray emission is also more intense. A quantitative comparison of the integrated mass as a function of radius for each subgroup and the two models is presented in Figure 3. Here we include the dispersion of the 100 driver models produced after varying the initial condition and photometric redshifts (colored bands). The mass increase in the full model around the SE group is due mostly to the smaller photometric redshift of system 11 used to derive the full model compared with the larger redshift predicted by the driver model. The driver model does a good job of predicting the arc position and morphology for a source redshift of \( \approx 3 \). In the full model, the system is assumed to be at the photometric redshift of \( z = 2.2 \) instead, which results in a mass increase in the SE.
Figure 3. Integrated mass as a function of aperture radius. The purple lines correspond to the driver model, and the red lines correspond to the model derived with all arcs. For comparison, we also show the corresponding integrated mass for the mass model from Cerny et al. (2018) as black lines. In all cases, the solid lines are for profiles centered in the NW group, while the dashed lines are for profiles centered in the SE group. The colored regions show the 1σ interval derived from the 100 realizations, where both the initial conditions and photo-z are varied.

In this section, we compare our models with previous results derived from the same RELICS data and presented in Cerny et al. (2018). We should note that the constraints used in this work and Cerny et al. (2018) are not exactly the same, so some of the differences can be attributed to this fact. Our lensed candidates were derived independently from Cerny et al. (2018); in some cases, our system candidates coincide, but not in all cases. System 3, with photometric redshift \( z = 7.42 \) in Cerny et al. (2018), is claimed to be a newly identified system. However, the positions of 3.1, 3.2, and 3.3 are very similar (within a fraction of an arcsecond) to the positions of system 3 in Zitrin et al. (2013), but in this case, they have a different photometric redshift of 4.16. For the position and redshift of system 3, we adopt the values of Zitrin et al. (2013) based on the color of the images and positions relative to the critical curve of preliminary models. We have not used system 8 from Cerny et al. (2018). System 8 appears as a likely multiple lensed arc with two pairs of images very close to each other and possibly merging around a critical curve. A third predicted image cannot be found near the predicted position, suggesting that this is not a real system or that the assumed redshift is significantly different from the real one. Alternatively, the predicted image may be buried behind a member galaxy. System 10 in Cerny et al. (2018) is included as an extra knot in our system 1. Similarly, we have included the positions of system 20 of Cerny et al. (2018) as additional knots of system 2. For the driver model, all systems are also included in Cerny et al. (2018; although see the difference in redshift for system 3). For the alternative full model, we include all systems used in Cerny et al. (2018), except system 8, as mentioned above. In the full model, we also include systems not included in Cerny et al. (2018). These are the new system candidates 6, 10, 13, 14, 16, and 17 and the redefined system candidates 11, 12, and 14 listed in the table in Appendix A. System 14 in Cerny et al. (2018) was not included in their model. Here we use a redefined version of this system as our new system 12. As for the new redefined system 11, Figure 5 shows the predicted images for 11a and 11b based on image 11c for the driver and full models.

The critical curves of our two models and the model in Cerny et al. (2018) are compared in Figure 6. The position of the critical curves is consistent between both models, although our model predicts slightly wider critical curves, suggesting a rounder distribution for the projected surface mass density in our model. In contrast, Cerny et al. (2018) predicted a more elongated distribution of matter, with the mass being more concentrated around the line intersecting the two clumps. The models show a better agreement (in terms of positions of the critical curve) around the position of the constraints. The figure shows the estimated observed position of the critical curve based on symmetry arguments for the giant arc of system 2 (at \( z \approx 3.3 \)). All models agree relatively well with this.
position by placing the critical curve (at the redshift of system 2) very close to or intersecting the estimated position of the critical curve. In the SE part of the lens, the differences between models are larger, reflecting the relatively smaller density of the constraints in this part of the lens (see Figure 1), but possibly also the fact that parametric methods assume explicit mass profiles that can extend the mass distribution beyond the range of distances covered by the lensing constraints. A more quantitative comparison of the magnification between the different models can be made by comparing the curves, \( A(>\mu) \), of the area above a given magnification.

These curves are computed by integrating the differential area curves, i.e., \( A(>\mu) = \int_{\mu}^{\mu_{\text{max}}} \frac{dA}{d\mu} d\mu \), where \( \mu_{\text{max}} \) is the maximum magnification considered (220 in this case) and \( dA/d\mu \) is the area in the lens plane with magnification \( \mu \) and in the interval \( d\mu \). An additional division by the factor \( \mu \) is needed to compute the corresponding area in the source plane. The curves \( A(>\mu) \) follow the usual \( A_{\mu} \mu^{-2} \) above magnification \( \mu \approx 10 \). The values of the normalization for the different models and at \( z_s = 3.3 \) are (in arcmin\(^2\)) \( A_0 = 4.5 \) (Cerny18 model), 10 (driver model), and 8.5 (full model). Here \( A(>\mu) \) can be interpreted as the probability of a galaxy being lensed by a factor larger than \( \mu \). At high magnifications, the driver and full models predict about twice the probability compared with the model in Cerny et al. (2018). This difference is mostly due to the shallower profiles in the driver and full models around the position of the critical curves. The values of \( A_0 \) put El Gordo at a level comparable to other powerful lenses, like the Hubble Frontier Fields clusters, in terms of lensing efficiency (see, for instance, Vega-Ferrero et al. 2019). This means that future observations, like the planned ones with JWST, promise to reveal many additional high-redshift lensed galaxies.

Due to the relatively large separation between the two groups, the critical curve on the east-central side of the cluster is relatively unperturbed by cluster members (this can be appreciated in Figure 1, where the critical curves are very smooth in this part of the cluster). At even higher redshift, the critical curves move outward, where the distortion by cluster members is expected to be even smaller. This has important implications for the probability of observing caustic crossing events of distant stars, for instance, Population III stars at \( z > 7 \), as suggested by Windhorst et al. (2018). Pristine critical curves (that is, critical curves that are not perturbed by microlensing from stars or remnants in the cluster members or intracluster medium) can produce lensed images in their vicinity with magnification factors of order \( 10^6 \) when the
background source has the size of a Population III star. In contrast, critical curves that are close to the cluster center (for instance, for background objects at relatively low redshifts of \(z \approx 2\) or less) are normally perturbed by such microlenses, resulting in maximum magnification factors of order \(10^4\) for background sources with sizes comparable to giant stars (see, for instance, Diego 2019, for details). The relatively unperturbed sections of the critical curves in El Gordo are hence good target regions for JWST to search for highly magnified distant stars.

In terms of total mass, the agreement between the models is made more evident when looking at the integrated mass as a function of aperture radius. In order to better account for the asymmetric nature of the cluster, we set the center of the aperture at the position of the two main galaxies (or BCGs). For each center, we compute the projected mass within a given aperture as a function of the aperture radius. The resulting profiles are shown in Figure 3. All models agree well, especially between \(\approx 100\) and \(300\) kpc, which is the range where lensing constraints are more abundant. At small radii (\(r < 100\) kpc), the model in Cerny et al. (2018) predicts slightly more mass than our free-form models, especially in the SE clump. At radii larger than \(\approx 400\) kpc, our free-form models fall below the prediction of the model of Cerny et al. (2018). This is an expected behavior, since the free-form models have systematically low masses in areas that extend beyond the realm of the lensing constraints. This is simply a memory effect of the algorithm that does not constrain distant regions in the field of view, leaving their masses close to their initial value before the minimization (these masses are originally assigned small random values).

5. Discussion

The results from the previous section suggest that the mass in El Gordo is relatively well constrained in the inner 500 kpc region. Within this range, Cerny et al. (2018) found that the masses within the 500 kpc radius for each clump have a mass ratio of 1.19 (for SE/NW). Compared with our results, we find that within the same radius, the SE/NW ratio is 0.98 for the driver model and 1.11 for the model with all systems. At 100 kpc, this ratio grows to 1.17 and 1.18 for the driver and full models, respectively. This can be compared with the dynamical masses inferred in Menanteau et al. (2012), where for the SE/NW ratio (within the virial radius), they found a value of 0.6 \(\pm 0.4\) and hence consistent with a ratio of \(\approx 1\) at 1\(\sigma\) with their measurement. The weak lensing analysis in Jee et al. (2014) finds a more discrepant ratio of the SE/NW groups (in the virial masses) of 0.56 \(\pm 0.17\) (statistical), in contrast with our results. This discrepancy may be due to systematic effects in either analysis, but it is also possible that the NW group becomes more massive than the SE group beyond the 500 kpc radius studied in the strong lensing analysis.

One of the more puzzling aspects of the El Gordo cluster is the position of the X-ray emission in relation to the peak in the mass distribution. Botteon et al. (2016) studied this cluster with X-rays and inferred a very high velocity for the shock (with a Mach number of 3 or above), which is spatially coincident with one of the radio relics. Ng et al. (2015) combined different observations from the El Gordo cluster to constrain the dynamical state of the cluster. Based on the separation of the two subgroups, the morphology of the radio relics, and their polarization angle, they concluded that the cluster is most likely in a return phase. This naturally explains the relative position of the X-ray peak and the main BCG, which seems to be lagging behind the X-ray peak. Hallman & Markevitch (2004) studied the low-redshift cluster A168, which resembles El Gordo. Like in El Gordo, in A168, the peak of the X-ray emission seems to have moved ahead of the dominant galaxies in the cluster. They concluded that the “subcluster gas slingshots past the dark matter center, becomes unbound from the subcluster and expands adiabatically.” This type of adiabatic expansion has been observed in N-body hydrodynamical simulations (Mathis et al. 2005), and it can result in substantial cooling of the gas as it leaves the potential well. For a cluster in a return phase after a head-on collision, the gas leaves the potential well twice: the first time as it drags behind the peak of the mass distribution due to ram pressure, and the second time when the peak of the mass distribution falls back toward the center of mass sprinting through the gas. Similar slingshot mechanisms have been recalled to interpret other clusters (Merten et al. 2011; Medezinski et al. 2016).

The results from our driver model seem to agree better with the Ng et al. (2015) interpretation (returning phase), since the X-ray peak is ahead of the mass peak. On the contrary, the full model seems to agree better with the interpretation of Molnar & Broadhurst (2015) and Zhang et al. (2015), since the mass peak coincides with the X-ray peak and not the BCG. This would suggest that the BCG was perturbed out from the potential well of the infalling cluster, which may be explained by the merging. Based on results from full N-body hydro simulations (Molnar & Broadhurst 2015; Zhang et al. 2015), the infall velocity for El Gordo is not as large as for the Bullet cluster. If the velocity is not large enough, the ram pressure does not produce noticeable offsets between the mass and gas centers of the infalling cluster. Future spectroscopic confirmation of lensed systems, plus the addition of new systems from deeper JWST observations, will allow one to resolve this question by unambiguously determining the peak position of the mass in the SE group. A confirmation of the candidate systems (and their redshift) used in the full model would favor the scenario proposed by Molnar & Broadhurst (2015) and Zhang et al. (2015).

5.1. Total Mass

In this section, we estimate the total mass of the lens model within \(r_{200c}\) by simultaneously fitting two spherically symmetric gNFW (generic Navarro Frenk and White) profiles to the lens model. We place one gNFW at the center of each clump and simultaneously fit the two profiles of the lens model. In this case, as opposed to what was shown in Section 4 and Figures 3 and 4, we do not mask the other subgroup when computing the profiles of each subgroup, since the simultaneous fit already considers the superposition of the two profiles in the intermediate region (and the other regions in the field of view). The gNFW is a generalization of the NFW profile and is given by

\[
\rho(r) = \frac{\rho_s}{\left(\frac{r}{r_s}\right)^2 \left[\left(1 + \frac{r}{r_s}\right)^2 - \gamma\right]^{(\beta - 3)/\alpha}},
\]

where \(\rho_s\) is the normalization and \(\gamma, \alpha, \beta\) are the inner, middle, and outer slopes of the profile.

In a standard NFW profile, \(\gamma = 1, \alpha = 2, \beta = 3\). Since the lens model can only constrain the inner \(\approx 300\) kpc, the
slopes $\alpha$ and $\beta$ are unconstrained. Hence, we fix only for the normalization, $\rho_0$; inner slope, $\gamma$; and scale radius, $r_s$. We fix $\alpha$ to the value for the standard NFW (i.e., $\alpha = 2$). For $\beta$, we consider two values, since the value of $\beta$ largely determines the integrated mass up to the virial radius. In particular, we consider the two cases of $\beta = 3$ (NFW-like case) and 2 (isothermal-like case) to explore the uncertainty due to the extrapolation of the double gNFW profile up to the virial radius. While the case of $\beta = 3$ is consistently reproduced by $N$-body simulations, shallower values are expected at large radii when one considers contributions from the two-halo term. We find that shallow values for the central slope of $\gamma = 0.2$, together with a scale radius in the range 285–380 kpc, are needed in order to reproduce the mass distribution of the lens model. The small-scale radii correspond to the isothermal-like profiles, $\beta = 2$, while larger-scale radii are needed when adopting the NFW-like profiles, $\beta = 3$.

By adding the two gNFW profiles and extrapolating to larger radii, it is, in principle, possible to estimate the mass at the virial radius. Note that the mass within the virial radius estimated this way is simply an approximation, and it is expected to be uncertain given the unknown behavior of the unmodelled mass beyond the region containing the lensing constraints. To account for the uncertainty due to the unknown profile in the outer regions and errors in the photometric redshifts, we fit the two models bracketing the $1\sigma$ regions from the 100 models described at the end of Section 3. The model in the upper $1\sigma$ limit is assigned the shallower $\beta = 2$ slope, which will correspond to the upper limit of the extrapolated mass. The model in the lower $1\sigma$ limit is assigned the steeper $\beta = 3$ slope, and it will correspond to the lower limit of the extrapolated mass.

Finally, the two gNFW models are added up together to account for the total mass. The resulting interval is shown as an orange shaded region in Figure 7. The upper limit corresponds to the upper $1\sigma$ model (among the 100 lens models) extrapolated with $\beta = 2$, and the lower limit corresponds to the lower $1\sigma$ model (among the 100 lens models) extrapolated with $\beta = 3$. The black solid line in Figure 7 shows the resulting extrapolation for the driver model. For comparison, the black dashed line shows the double gNFW fit to the simulated cluster. Note how the driver model closely follows the simulation below $\approx 300$ kpc. Above $\approx 300$ kpc, the departure is mostly due to the memory effect at large radii, that is, the fact that the lens model is unconstrained beyond some radius and the solution recovers the initial condition of the optimization (small random values). The same departure is observed when we select the center of the profile as the point in between the two subgroups and we skip the fit to the gNFW profiles. In this case (blue solid line in Figure 7), the departure is observed at $\approx 400$ kpc instead. When the center of the profile is in between the two subgroups, the integrated mass of the simulated cluster (blue dashed line) converges to the mass of the double gNFW profile at a radius of $\approx 1$ Mpc. The light blue line marks the mass enclosed in a sphere with 200 times the critical density. The intersection with this curve marks the virial radius and mass.

Based on the uncertainty of the extrapolation and the dispersion due to photometric redshifts and the randomness of the initial condition in the optimization, we can estimate the mass at the virial radius from the intersection with the light blue line, $M_{200c} = 1.08^{+0.05}_{-0.12} \times 10^{15} M_\odot$. The extrapolation of the double gNFW model seems to imply that the lens model has a smaller virial mass than the simulated cluster. This tentative conclusion will need the JWST images and spectra to be confirmed by extending the range in radius of the lensing constraints. The question of the total mass of El Gordo will only be settled with new data. The JWST will provide new strong lensing constraints (for instance, high-$z$ galaxies extending farther away from the cluster central region), as well as weak lensing data, which is scarce in current data given the high redshift of the cluster.

5.2. The Contribution from the Gas and Filamentary Structure around the SE BCG

The position of the peak in the mass distribution of the full model is coincident with the peak in the X-ray emission. This coincidence is intriguing, as it raises the possibility that the X-ray-emitting gas is contributing substantially to the projected mass in this region of the lens plane. This is not all that surprising, since clusters are expected to contain a baryon fraction representative of the cosmological value, and most of the baryons in massive clusters are in the form of a hot plasma. One expects the hot plasma to account for $\approx 10\%$–$15\%$ of the total mass in the entire cluster. Due to the more efficient cooling of baryons, this fraction increases toward the center of clusters, so around the BCG, one could expect an even larger contribution to the total mass from the hot plasma. No significant radio emission is observed near this position, except for a compact radio source named U7 in Figure 16 of Lindner et al. (2014) and at 75 kpc SE from the BCG. This radio source is spatially coincident with the peak of the X-ray emission (see Figure 8). Based on X-ray data, Menanteau et al. (2012) constrained the electron density to values between 0.023 and 0.045 cm$^{-3}$ within a region of diameter $\approx 170$ kpc (or $\approx 22''$). Based on this electron density, the gas surface mass density (projected on 170 kpc along the line of sight) is then $\Sigma_{gas} \approx 100$–$200 M_\odot$ pc$^{-2}$, which should be compared with the critical surface mass density of $\Sigma_{crit} = 2800 M_\odot$ pc$^{-2}$ for a source at $z_s = 3$ (note that the critical surface mass density is projected along a much longer line of sight). The contribution from the gas
to the convergence, $\kappa$ (where $\kappa$ is defined as the ratio between the surface mass density and the critical surface mass density at the given lens and source redshifts), projected along this relatively small interval of 170 kpc is then 0.035–0.07. At the peak of the X-ray emission and projecting over larger distances, the gas can easily contribute over 0.1 to $\kappa$. This may be sufficient to explain the correlation between the total mass peak and the X-ray emission shown in Figure 2.

Interestingly, as discussed earlier, the peak of the X-ray emission coincides with blue features observed in the UV-optical bands (A and B in Figure 8). Based on the driver model, if these two features are multiply lensed objects, they need to be at a redshift above $z = 1.8$. However, at this redshift, the predicted images would not form radially oriented arcs in our models, but rather tangential arcs. Radially oriented arcs at this position in the lens plane appear for redshifts $z \sim 2$ or above, but the lens model cannot reproduce arcs with a morphology similar to the observed ones. For redshifts between $z = 4.5$ and 5.5, counterimages for the arcs are expected at positions compatible with the three arclets of system 4. This correspondence explains the original classification of these features as part of system 4, but the morphology of the predicted images differs significantly from the observation, making this possibility unlikely. The two features are at distances of $\approx 35$ and $\approx 50$ kpc, respectively, from the center of the BCG. The tight correlation shown in Figure 8 between these two features and the offset peak of the X-ray emission suggests that these two features may be the optical counterpart of a cooling flow (alternatively, they could be associated with a jet emitting in UV-optical and X-rays). A well-studied case that could serve as a similar example is the nearby cluster A2597 ($z = 0.0821$), where an arc-like feature also correlates very well with the peak in the X-ray emission. Tremblay et al. (2012) studied this cluster combining data from X-ray, UV/optical, near-IR, and radio observations. They found evidence for a cooling flow in the X-ray band and filamentary features in the far-UV and optical bands that resemble the blue features shown in Figure 8, which could be associated with precipitation of the gas (Voit et al. 2015). If cold and dense gas is responsible for this feature, this could explain the apparent correlation between the X-ray and mass discussed in the previous subsection.

### 6. Conclusions

We derive a new lens model and mass estimate for the El Gordo cluster using data from the RELICS program. This cluster is of special interest, since earlier estimates suggest that it is the most massive cluster at redshift $z \sim 0.9$; hence, it can be used to constrain the cosmological model, for instance, by using extreme value statistics (Harrison & Coles 2012). Understanding the uncertainties in the mass estimate emerging from the uncertainties in the lens model is of pivotal importance. In this work, we present the first free-form model of this cluster that does not rely on assumptions about the mass distribution that, if they are erroneous, could bias the mass estimation. Free-form models offer a complementary cross-check with parametric models and are of particular interest in highly irregular clusters, like El Gordo, due to the lack of assumptions about the underlying distribution of dark matter.

We first derive a robust model (nicknamed the driver model) based on a reliable subsample of lensed galaxies. Using the driver model, we unveil new strongly lensed system candidates and infer their redshifts. With the full set of lens systems, we derive an alternative model (or full model) for the mass distribution. Both models are similar to each other, but small differences can be identified, especially in the SE sector of the cluster. We explicitly compare our models with the one derived by Cerny et al. (2018) using the same RELICS data but a different sample of lensed galaxies (although with substantial overlap between our sample and theirs). We find that our lens model predicts wider critical curves, but the integrated mass as a function of aperture is consistent with the model of Cerny et al. (2018), thus suggesting that the mass estimate obtained with earlier parametric methods may not be affected by erroneous choices in the parameterization. Our new model also predicts nearly twice the lensing efficiency above a given magnification factor (at large magnifications).

We explore uncertainties in the lens model due to errors in the assumed redshifts. Adopting an uncertainty of 0.5 in the redshift of all systems, we produce a set of 100 models where we vary the redshift with this uncertainty. We find that the mass within 300 kpc of each subgroup varies by relatively small amounts due to this uncertainty in redshift. By fitting our full lens model to an NFW profile and extrapolating up to $R_{200c}$, we find a mass $M_{200c} = (1.08 \pm 0.05) \times 10^{15} M_\odot$, where the uncertainty comes mostly from the poorly constrained scale radius (and the external slope $\beta$). Due to the highly nonsymmetric nature of the cluster, the extrapolation to the virial radius is not as reliable as it would be in more symmetric clusters. Hence, this mass estimate should be taken with caution. Improved identification and redshift determinations of the lensed galaxies, together with the addition of dense weak lensing measurements, will be needed to improve the mass estimate up to the virial radius. Our mass estimate is smaller than previous estimates. If confirmed, this would relax the tension with standard $\Lambda$CDM models that predict that clusters with masses above $M_{200c} = 1.7 \times 10^{15} M_\odot$ at this redshift should be very rare (this cluster was serendipitously found in a relatively small 2% portion of the sky). We test the accuracy of the profile extrapolation using an N-body simulation that matches most of the observed features in El Gordo and conclude that our mass estimate may still be underestimated the real mass due to systematics in the extrapolation of the profile. A combination of future strong and weak lensing data (for instance, with JWST, which will be able to detect the high-redshift...
background galaxies needed for the weak lensing analysis of this high-
$z$ cluster) should allow for better constraint of the total mass of this cluster.

We find evidence for the lens model being sensitive to the gas mass. In particular, we find that the peak of the smooth component of the mass distribution in the full lens model agrees well with the peak of the X-ray emission (which is offset with respect to the nearby BCG). We discuss the possibility that two features at the locations of these peaks, which are observed in the optical-UV bands and interpreted in the past as background lensed galaxies, are instead the optical counterpart of a cooling flow or a precipitation mechanism from the hot plasma.

New lens models will be valuable when El Gordo is observed as part of Cycle 1 of JWST (as part of the JWST Medium Deep Fields program; PI: R. Windhorst). New arc systems, including several at high redshift, are expected to be discovered with the new JWST data. Our lens model can be used to identify new strongly lensed system candidates, as well as to estimate their redshifts.

J.M.D. acknowledges the support of projects AYA2015-64508-P (MINECO/FEDER, UE), funded by the Ministerio de Economía y Competitividad, and PGC2018-101814-B-100 (MCIU/AEI/MINECO/FEDER, UE), Ministerio de Ciencia, Investigación y Universidades. This project was funded by the Agencia Estatal de Investigación, Unidad de Excelencia María de Maeztu, ref. MDM-2017-0765. This work was funded by NASA JWST Interdisciplinary Scientist grants NAG5-12460, NNX14AN10G, and 80NSSC18K0200 to R.A.W. from GSFC. We would like to thank Harald Ebeling for making the code ASMOOTH (Ebeling et al. 2006) available. J.M.D. acknowledges the hospitality of the Physics Department at the University of Pennsylvania for hosting him during the preparation of this work. This work is based on observations made with the NASA/ESA Hubble Space Telescope, operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-2655. Some of the data for this study were retrieved from the Mikulski Archive for Space Telescope (MAST). The authors would like to thank the RELICS team for making the reduced data set available to the community. The scientific results reported in this paper are based in part on data obtained from the Chandra Data Archive.

### Appendix A

Compilation of Arc Positions

This appendix presents the sample of secure and likely lensed multiple images detected behind El Gordo using the

| Knot ID | R.A. | Decl. | $\zeta_{\text{phot}}$ | $\zeta_{\text{model}}$ | $\zeta_{\text{used}}$ | Rank | Comments |
|---------|------|-------|----------------------|----------------------|----------------------|------|----------|
| 1.1.1   | 1 02 53.275 | $-$49 15 16.49 | 3 | 3 | A | Z13(1) |
| 1.2.1   | 1 02 52.819 | $-$49 15 18.29 | 2.93 | A |
| 1.3.1   | 1 02 55.411 | $-$49 14 59.90 | A |
| 1.1.2   | 1 02 53.340 | $-$49 15 16.36 | A |
| 1.2.2   | 1 02 52.763 | $-$49 15 18.70 | 2.8 | A |
| 1.3.2   | 1 02 55.391 | $-$49 15 00.33 | 2.91 | A |
| 1.1.3   | 1 02 53.480 | $-$49 15 16.01 | A |
| 1.2.3   | 1 02 52.600 | $-$49 15 19.68 | A |
| 1.3.3   | 1 02 55.320 | $-$49 15 01.18 | 3.26 | A |
| 2.1.1   | 1 02 55.828 | $-$49 15 52.37 | 3.21 | 3.3 | A | Z13(2) |
| 2.2.1   | 1 02 56.749 | $-$49 15 46.01 | 3.39 | A |
| 2.3.1   | 1 02 54.429 | $-$49 16 04.63 | 3.3 | A |
| 2.1.2   | 1 02 55.671 | $-$49 15 53.54 | A |
| 2.2.2   | 1 02 56.885 | $-$49 15 45.17 | 3.27 | A |
| 2.3.2   | 1 02 54.456 | $-$49 16 04.00 | 2.9 | A |
| 2.1.3   | 1 02 55.983 | $-$49 15 51.24 | A |
| 2.2.3   | 1 02 56.573 | $-$49 15 47.06 | A |
| 2.3.3   | 1 02 54.383 | $-$49 16 04.61 | A |
| 3.1.1   | 1 02 56.257 | $-$49 15 07.03 | 4.4 | 4.4 | A | Z13(3) |
| 3.2.1   | 1 02 54.751 | $-$49 15 19.54 | A |
| 3.3.1   | 1 02 51.536 | $-$49 15 38.47 | 4.54 | A |
| 4.1.1   | 1 02 59.986 | $-$49 15 49.54 | 3.98 | 3.2 | 4 | A | Z13(4) |
| 4.2.1   | 1 02 55.362 | $-$49 16 26.09 | 4.0 | A |
| 4.3.1   | 1 02 56.599 | $-$49 16 08.45 | A |
| 5.1.1   | 1 02 54.539 | $-$49 14 58.60 | 2.4 | 2.8 | 2.8 | A | C18(13) |
| 5.2.1   | 1 02 53.230 | $-$49 15 07.11 | A |
| 5.3.1   | 1 02 51.803 | $-$49 15 17.05 | 2.2, 2.5 | A |

13 ivo://ADS/Sa.CXO#obs/12258
14 ivo://ADS/Sa.CXO#obs/14022
15 ivo://ADS/Sa.CXO#obs/14023
updated imaging from the RELICS program. Table 1 lists the complete sample of images and their redshifts, assigning IDs to each of them.

The first column shows the system ID following the original notation of Zitrin et al. (2013) (ID1.D2.ID3 = System.Image.Knot) and ranks (A, B, and C). Systems 1–4 were initially presented by Zitrin et al. (2013). The IDs marked in bold are new systems presented in this work. Photometric redshifts are given in column $z_{\text{phot}}$. The systems having spectroscopic redshifts are marked in bold. Redshifts predicted by the lens model are given in column $z_{\text{model}}$. The redshift used to reconstruct the lens is given in column $z_{\text{used}}$.

The column labeled Rank shows the quality of the system. Systems marked with rank A are very reliable and used to derive the driver model. Systems marked with B are used to derive (together with systems having rank A) the full model.
Systems marked with C are less reliable but still highly consistent with the driver lens model. In the last column, 1, 2, and 3 refer to previous work, where these systems are defined; Z13 stands for Zitrin et al. (2013), while C18 stands for Cerny et al. (2018). The number in parentheses next to the reference indicates the system ID in the corresponding publication. During the refereeing process of this manuscript, the first spectroscopic measurements for these systems were presented in Caputi et al. (2020), which confirmed systems 3 and 17, both with redshift $z_{\text{spec}} = 4.32$ and consistent with our lens model.

**Appendix B**

**Predicted Arcs**

In this section, we show the position and morphology of the images predicted by the driver model. In Figure 9, we show the entire field of view and define eight regions that contain the multiply lensed images. These regions are later shown separately in Figure 10, where we represent both the data (in panels labeled A, B, C, D, E, F, G, and H) and the prediction made by the driver model (panels labeled with the corresponding primed letters). The orientation, scale, and reference position are the same for each pair of panels. For each multiply lensed system, one of the counterimages is chosen (dashed circles), delensed into the source plane, and relensed into the image plane (solid circles). We remind the reader that the driver model is derived from just five multiple lensed systems, shown as yellow circles and marked with yellow labels in Figure 10. Hence, deviations between the data and the prediction made by the model are expected to be larger for the systems (like systems 10, 11, and 13) that were not used to derive the driver lens model (i.e., white circles in Figure 10).

Note that when systems are not used as constraints in the model, deviations in position similar to those shown in

![Figure 9](image-url)
Figure 10. Zoomed regions shown in Figure 9. Panels with primed letters denote the images predicted by the driver model. Dashed circles show the images that are used to predict the lensed images in the panels with primed letters. Solid circles mark the images being predicted by the driver model (and shown in the panels with primed letters). The systems used to derive the driver model are marked in yellow. The scale and reference points are the same for each pair of panels. Blue circles mark possible alternative candidates. The red circle in panel H' marks the position of a counterimage that is predicted but not observed.

Figure 10 are expected, especially in regions of the lens plane where constraints are more scarce. In some cases, we mark possible alternative candidates with blue circles. A red circle is used in panel H' to mark the position of an arclet predicted by the model but not observed, which would correspond to system 4. However, this arclet falls within the BCG and overlaps with the blue UV regions discussed in Section 5.2, so it is possible that it is merged with the UV regions.

ORCID iDs
Jose M. Diego © https://orcid.org/0000-0001-9065-3926
S. M. Molnar © https://orcid.org/0000-0002-4227-956X
References

Botteon, A., Gastaldello, F., Brunetti, G., & Kale, R. 2016, MNRAS, 463, 1534
Caputi, K. I., Caminha, G. B., Fujimoto, S., et al. 2020, arXiv:2009.04838
Cerny, C., Sharon, K., Andrade-Santos, F., et al. 2018, ApJ, 859, 159
Chan, B. M. Y., Broadhurst, T., Lim, J., et al. 2020, ApJ, 888, 35
Coe, D., Salmon, B., Bradac, M., et al. 2019, ApJ, 884, 85
Diego, J. M. 2019, A&A, 625, A84
Diego, J. M., Broadhurst, T., Chen, C., et al. 2016, MNRAS, 456, 356
Diego, J. M., Kaiser, N., Broadhurst, T., et al. 2018, ApJ, 857, 25
Diego, J. M., Protopapas, P., Sandvik, H. B., & Tegmark, M. 2005, MNRAS, 360, 477
Diego, J. M., Tegmark, M., Protopapas, P., & Sandvik, H. B. 2007, MNRAS, 375, 958
Donnert, J. M. F. 2014, MNRAS, 438, 1971
Ebeling, H., White, D. A., & Rangarajan, F. V. N. 2006, MNRAS, 368, 65
Hallman, E. J., & Markevitch, M. 2004, ApJL, 610, L81
Harrison, I., & Coles, P. 2012, MNRAS, 421, L19
Jee, M. J., Hughes, J. P., Menanteau, F., et al. 2014, ApJ, 785, 20
Johnson, T. L., & Sharon, K. 2016, ApJ, 832, 82
Kelly, P. L., Diego, J. M., Rodney, S., et al. 2018, NatAs, 2, 334
Lindner, R. R., Baker, A. J., Hughes, J. P., et al. 2014, ApJ, 786, 49
Mathis, H., Lavaux, G., Diego, J. M., & Silk, J. 2005, MNRAS, 357, 801
Medezinski, E., Umetsu, K., Okabe, N., et al. 2016, ApJ, 817, 24
Menanteau, F., Hughes, J. P., Sifón, C., et al. 2012, ApJ, 748, 7
Mertan, J., Coe, D., Dupke, R., et al. 2011, MNRAS, 417, 333
Molnar, S. M., & Broadhurst, T. 2015, ApJ, 800, 37
Molnar, S. M., & Broadhurst, T. 2018, ApJ, 862, 112
Ng, K. Y., Dawson, W. A., Wittman, D., et al. 2015, MNRAS, 453, 1531
Sebesta, K., Williams, L. L. R., Liesenborgs, J., Medezinski, E., & Okabe, N. 2019, MNRAS, 488, 3251
Sendra, I., Diego, J. M., Broadhurst, T., & Lazkoz, R. 2014, MNRAS, 437, 2642
Tremblay, G. R., O’Dea, C. P., Baum, S. A., et al. 2012, MNRAS, 424, 1042
Vega-Ferrero, J., Diego, J. M., & Bernstein, G. M. 2019, MNRAS, 486, 5414
Voit, G. M., Donahue, M., Bryan, G. L., & McDonald, M. 2015, Natur, 519, 203
Waizmann, J. C., Ettori, S., & Moscardini, L. 2012, MNRAS, 420, 1754
Watson, W. A., Illiev, I. T., Diego, J. M., et al. 2014, MNRAS, 437, 3776
Williamson, R., Benson, B. A., High, F. W., et al. 2011, ApJ, 738, 139
Windhorst, R. A., Timmes, F. X., Wytse, J. S. B., et al. 2018, ApJS, 234, 41
Zhang, C., Yu, Q., & Lu, Y. 2015, ApJ, 813, 129
Zhang, C., Yu, Q., & Lu, Y. 2018, ApJ, 855, 36
Zitrin, A., Menanteau, F., Hughes, J. P., et al. 2013, ApJL, 770, L15