

A massive blow for Λ CDM – the high redshift, mass, and collision velocity of the interacting galaxy cluster El Gordo contradicts concordance cosmology

Elena Asencio,¹★ Indranil Banik² and Pavel Kroupa^{1,2}

¹*Helmholtz-Institut für Strahlen und Kernphysik (HISKP), University of Bonn, Nussallee 14–16, D-53115 Bonn, Germany*

²*Astronomical Institute, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic*

Accepted 2020 October 28. Received 2020 September 23; in original form 2020 June 4

ABSTRACT

El Gordo (ACT-CL J0102-4915) is an extremely massive galaxy cluster ($M_{200} \approx 3 \times 10^{15} M_{\odot}$) at redshift $z = 0.87$ composed of two subclusters with a mass ratio of 3.6 merging at speed $V_{\text{infall}} \approx 2500 \text{ km s}^{-1}$. Such a fast collision between individually rare massive clusters is unexpected in Lambda cold dark matter (Λ CDM) cosmology at such high z . However, this is required for non-cosmological hydrodynamical simulations of the merger to match its observed properties. Here, we determine the probability of finding a similar object in a Λ CDM context using the Jubilee simulation box with a side length of $6 h^{-1} \text{ Gpc}$. We search for galaxy cluster pairs that have turned around from the cosmic expansion with properties similar to El Gordo in terms of total mass, mass ratio, redshift, and collision velocity relative to virial velocity. We fit the distribution of pair total mass quite accurately, with the fits used in two methods to infer the probability of observing El Gordo in the surveyed region. The more conservative (and detailed) method involves considering the expected distribution of pairwise mass and redshift for analogue pairs with similar dimensionless parameters to El Gordo in the past light-cone of a $z = 0$ observer. Detecting one pair with its mass and redshift rules out Λ CDM cosmology at 6.16σ . We also use the results of Kraljic and Sarkar to show that the Bullet Cluster is in 2.78σ tension once the sky coverage of its discovery survey is accounted for. Using a χ^2 approach, the combined tension can be estimated as 6.43σ . Both collisions arise naturally in a Milgromian dynamics (MOND) cosmology with light sterile neutrinos.

Key words: gravitation – methods: statistical – galaxies: clusters: individual: El Gordo – galaxies: clusters: individual: Bullet Cluster – dark matter – large-scale structure of Universe.

1 INTRODUCTION

According to the hierarchical structure formation paradigm, smaller structures formed first and then merged into larger structures, eventually forming galaxy clusters (Davis et al. 1985). These are the largest gravitationally bound structures in the Universe. The efficiency and time-scale proposed for their formation vary depending on the cosmological model adopted. The most generally accepted model nowadays is Lambda cold dark matter (Λ CDM; Ostriker & Steinhardt 1995). Λ CDM is built on the assumption of a homogeneous and isotropically expanding Universe on very large scales, with all its matter–energy content appearing at very early times in a hot big bang. Two other elements had to be added in order to explain certain astronomical phenomena: The unexpected faintness of distant Type Ia supernovae at fixed redshift z suggested that the Universe was undergoing an accelerated expansion, which implied that Einstein’s field equations should have a non-zero cosmological constant Λ (Riess et al. 1998); and the observed flat rotation curves of galaxies (e.g. Babcock 1939; Rubin & Ford 1970; Rogstad & Shostak 1972; Bosma 1978) indicated that the Newtonian gravity of the

baryons alone was not enough to hold them together, which led to the postulation of undetected mass that would surround galaxies in the form of a halo (Ostriker & Peebles 1973). Gravitational microlensing results show that the Galactic halo cannot consist of compact objects with planetary or stellar mass (e.g. Alcock et al. 2000; Tisserand et al. 2007), so it is instead thought to consist of non-baryonic particles that do not exist in the well-tested standard model of particle physics. Null detection of non-gravitational signals from the postulated CDM particles places stringent constraints on their allowed properties (e.g. Abazajian et al. 2020; Hoof, Geringer-Sameth & Trotta 2020). None the less, Λ CDM has proven successful in explaining, e.g. the cosmic microwave background (CMB; Planck Collaboration XXVII 2014).

However, its predicted efficiency for the formation of galaxy clusters has been challenged by observations of several massive galaxy clusters at high z (e.g. Mullis et al. 2005; Ebeling et al. 2007; Lamer et al. 2008; Jee et al. 2009; Menanteau et al. 2010; Foley et al. 2011; Stanford et al. 2012; Stalder et al. 2013; Buddendiek et al. 2015; Tozzi et al. 2015; Wang et al. 2016; Miller et al. 2018; Wen & Han 2018). The most massive cluster at $z > 1$ is SPT-CL J2106-5844, with $z = 1.132^{+0.002}_{-0.003}$ and mass $M_{200} = (1.27 \pm 0.21) \times 10^{15} M_{\odot}$ (Foley et al. 2011). This was confirmed by the more recent study of Kim et al. (2019), which gives a slightly lower mass estimate of $M_{200} = (1.04^{+0.33}_{-0.30}) \times 10^{15} M_{\odot}$. Their fig. 6 shows two weak lensing

* E-mail: s6elena@uni-bonn.de (EA); ibanik@astro.uni-bonn.de (IB)

peaks, indicating that the system is likely a merger product. Further studies are required to pin down the pre-merger configuration. The Bullet Cluster (1E 0657-56) at redshift $z = 0.296$ and El Gordo (ACT-CL J0102-4915) at $z = 0.870$ are the galaxy cluster encounters most well known for having a very high relative velocity (Markevitch et al. 2004) and for being the most massive distant galaxy cluster (Menanteau et al. 2012), respectively. These observations suggest that the formation of large structures took place earlier than expected in Λ CDM.

On even larger scales of tens of Mpc, galaxy clusters are often part of a supercluster. The existence of very massive superclusters at high redshift also seems to pose problems for the Λ CDM model (Cucciati et al. 2018; Hayashi et al. 2019). The former work identified a $4.8 \times 10^{15} M_{\odot}$ supercluster at $z = 2.45$, which is consistent with the predicted mass function at $z = 1$. Though a relevant simulation snapshot was unavailable for a more meaningful comparison at $z = 2.5$, the rapidly rising number density of such massive superclusters around the epoch $z = 1$ (fig. 2 of Lim & Lee 2014) suggests that there would be significant tension with the observations of Cucciati et al. (2018).

If structure formation is enhanced compared to Λ CDM expectations, then not only overdensities but also underdensities should be more pronounced than expected. There is actually strong evidence for a large local underdensity extending out to ≈ 300 Mpc (Keenan, Barger & Cowie 2013). This is in 6.04σ tension with Λ CDM expectations (section 2 of Haslbauer, Banik & Kroupa 2020). Those authors showed that such a large and deep void would cause the local expansion rate of the Universe to exceed the average by ≈ 11 per cent (see their equation 5). This would nicely resolve the Hubble tension, i.e. the difficulty in reconciling the observed CMB anisotropies (Aiola et al. 2020; Planck Collaboration VI 2020) with multiple independent determinations of the local expansion rate (Riess 2020, and references therein).

The problem of early structure formation also extends down to galaxy scales (e.g. Kang & Im 2015; Girelli, Bolzonella & Cimatti 2019; Forrest et al. 2020) and their central super massive black holes (SMBHs). In this respect, the recent discovery of the luminous quasar J1007+2115 at $z = 7.515$ is particularly noteworthy – it contains an SMBH of mass $(1.5 \pm 0.2) \times 10^9 M_{\odot}$. This discovery challenges models of SMBH growth, since these would require a seed black hole of $\geq 10^4 M_{\odot}$ at $z = 30$ to explain the observations (Yang et al. 2020).

Many of the publications addressing this rapid structure formation tension on the cluster scale focus on the Bullet Cluster (Tucker, Tananbaum & Remillard 1995) as one of the main objects with which to test the likelihood of finding such massive high- z clusters in a Λ CDM context (Hayashi & White 2006; Angus & McGaugh 2008; Lee & Komatsu 2010; Thompson & Nagamine 2012; Katz et al. 2013; Lage & Farrar 2014; Watson et al. 2014a; Kraljic & Sarkar 2015). This is because the Bullet Cluster consists of two colliding galaxy clusters with a collisional trajectory roughly normal to the line of sight, making it easier to study the interaction. It also has a high mass for a cluster at $z = 0.296$ – the main cluster has $M_{200} \approx 1.9 \times 10^{15} M_{\odot}$, while the subcluster mass is $M_{200} \approx 2.6 \times 10^{14} M_{\odot}$ (Lage & Farrar 2014). Most importantly, its shock and subcluster velocity were initially calculated to be $v \approx 4500 \text{ km s}^{-1}$ (Markevitch et al. 2004), making this cluster quite problematic for Λ CDM (e.g. Thompson & Nagamine 2012). More recent hydrodynamical simulations show that a lower infall velocity of $v \approx 3000 \text{ km s}^{-1}$ with an impact parameter of $b = 0.256 \text{ Mpc}$ is also compatible with observations (Lage & Farrar 2014). This alleviated the tension significantly, since the Bullet Cluster infall velocity is now

below its escape velocity (see their table 1). Nevertheless, according to Kraljic & Sarkar (2015), this cluster is still considered a rare object for a Λ CDM cosmology: Only 0.1 similar systems can be expected within $z < 0.3$ if we require that the collision has already occurred, as noticed by their referee. The Bullet Cluster was discovered in a survey covering only 5.4 per cent of the whole sky, making it a 2.78σ outlier for Λ CDM (Section 3.4). Finding yet more objects like this could imply a problem for the model.

The second aforementioned cluster of interest, El Gordo, presents significantly more striking features than the Bullet Cluster. El Gordo was observed for the first time by Menanteau et al. (2010) in the Atacama Cosmology Telescope’s 2010 survey of galaxy clusters, which were selected via the Sunyaev–Zel’dovich (SZ) effect (Sunyaev & Zeldovich 1970). El Gordo consists of two merging subclusters at $z = 0.87$ with cores separated by a projected distance of $\approx 700 \text{ kpc}$ (fig. 5 of Jee et al. 2014), a high infall velocity $V_{\text{inf}} \approx 2500 \text{ km s}^{-1}$, and a total mass $M_{200} \approx (2-3) \times 10^{15} M_{\odot}$. These values are estimated using both observations (Menanteau et al. 2012; Zitrin et al. 2013; Jee et al. 2014) and idealized hydrodynamical simulations of the interaction (Molnar & Broadhurst 2015; Zhang, Yu & Lu 2015). This makes El Gordo the most massive galaxy cluster at $z > 0.6$. Due to the highly energetic interaction of its two subclusters, it is also the hottest, most X-ray luminous, and brightest SZ effect galaxy cluster at this redshift (Menanteau et al. 2012). Its bright X-ray emission has a single peak and a characteristic elongated shape with two faint tails that extend beyond this peak. The X-ray emission can be detected even in the outer regions of the merging cluster. Other distinctive features of El Gordo are the large offsets between the X-ray and SZ centroids and between the SZ centroid and the mass centre of the main cluster. These offsets are expected features in high-velocity merging galaxy clusters (Molnar, Hearn & Stadel 2012), supporting the published high infall velocity estimates for El Gordo.

Even though the mass and redshift might appear as the most obvious oddities of El Gordo, we show in Section 2.3.3 that its infall velocity significantly exceeds the escape velocity according to the Zhang et al. (2015) results. This is likely not the case in the Bullet Cluster (table 1 of Lage & Farrar 2014). For two objects to be infalling faster than their escape velocity, there must have been a third object that pushed them towards each other – any peculiar velocities at high z are rapidly redshifted away by Hubble drag. If finding two objects this massive sufficiently close to each other is already rare, finding three objects like this is extremely unlikely. Moreover, the mass ratio between the two components of El Gordo is almost certainly in the range 2–5 (Zhang et al. 2015), with their best guess being 3.6 while other studies suggest ≈ 2 (Jee et al. 2014; Molnar & Broadhurst 2015). This makes El Gordo a major merger. Such events are expected to be much less common than minor mergers like the Bullet Cluster, where the mass ratio is ≈ 7 (Lage & Farrar 2014). To summarize, El Gordo is a much more exceptional object than the Bullet Cluster in terms of the total mass, mass ratio, redshift, and infall velocity.

Deeper surveys are needed to find higher redshift objects, which usually requires a trade-off against the surveyed area. El Gordo was found in a very small sky region of 455 deg^2 (1.1 per cent of the whole sky; Menanteau et al. 2010). However, the survey that obtained the first constraints on El Gordo’s properties was slightly larger: It covered 755 deg^2 (1.8 per cent of the sky; Menanteau et al. 2012). In our analysis, we adopt the latter value to be more conservative.

Despite its extreme nature, there are not so many works analysing the chance of finding El Gordo in a Λ CDM universe. This is partly because it was discovered quite recently, at least in comparison to the Bullet Cluster. Due to its large mass, a very large simulation box is

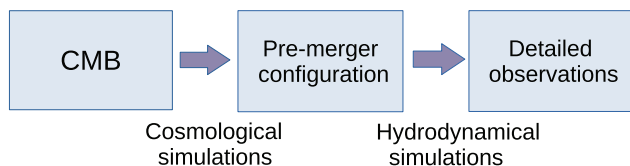


Figure 1. Diagram summarizing the two-step logic of this work: The baryon acoustic oscillations in the CMB constrain the cosmological parameters in a Λ CDM universe, fixing the abundance and evolution of its matter content. This information is used in large N -body cosmological simulations that allow us to estimate the distribution of structures at any epoch, in particular the cluster pairwise mass function. If Λ CDM is correct, the distribution at $z = 1$ contains a pair in a pre-merger configuration that, when advanced with a hydrodynamical simulation, reproduces the detailed properties of El Gordo. Such a detailed analysis is not possible without hydrodynamics, while including it precludes the use of a very large simulation volume and the associated statistical power. Previous studies like Zhang et al. (2015) obtained a good fit to El Gordo using hydrodynamical simulations, but did not quantify whether the required pre-merger configuration is plausible in the surveyed region if the observed anisotropies in the CMB are evolved to $z = 1$ in a Λ CDM context.

needed to study El Gordo in a statistically meaningful way. The few attempts to do so (Menanteau et al. 2012; Katz et al. 2013; Jee et al. 2014; Sahlén, Zubeldía & Silk 2016) do not provide a very detailed analysis. Except for Katz et al. (2013), these works consider El Gordo to be a single massive object instead of two massive objects at a very close distance infalling at high speed. The latter configuration is significantly less likely for clusters that are individually rather rare. Using the exclusion curve method of Mortonson, Hu & Huterer (2011), Menanteau et al. (2012) and Jee et al. (2014) found that El Gordo is a rare object but it is not in tension with Λ CDM. Jee et al. (2014) also added that, since El Gordo is only marginally allowed in Λ CDM, a more accurate measurement of its properties could give rise to significant tension. Sahlén et al. (2016) estimated the likelihood of finding an El Gordo analogue according to the cluster and void mass functions predicted by Λ CDM. Using a Markov chain Monte Carlo method, they also concluded that El Gordo is not in tension with Λ CDM. Katz et al. (2013) commented on the low likelihood of this object arising in a Λ CDM universe according to the mass–redshift relation of Mortonson et al. (2011). Importantly, Katz et al. (2013) mentioned that the probability is even lower if one takes into account the high infall velocity of El Gordo, which can make it a more serious problem for Λ CDM. Indeed, in their cubic simulation box with sides of $1 h^{-1}$ co-moving Gpc (cGpc), they found no pairs with the required mass, redshift, and collision velocity.

Finding analogues to massive objects like El Gordo in a cosmological simulation requires a very large box size to reliably estimate the occurrence rate. This precludes the use of hydrodynamical simulations. However, with only dark matter, it is not possible to try and match, e.g. the observed X-ray morphology of El Gordo. Instead, we rely on idealized hydrodynamical simulations to tell us the pre-merger configuration. We then seek pairs in the cosmological simulation with similar total mass, redshift, and collision velocity to the El Gordo progenitors. These properties should be little affected by small-scale baryonic processes, and so should be well reproduced in a dark matter-only simulation.

In this paper, we conduct a rigorous analysis to find the probability that a pair of progenitor galaxy clusters at $z \approx 1$ could have given rise to an object like El Gordo within the surveyed region if Λ CDM is the correct cosmological model. The overall logic of the project is summarized in Fig. 1. For the second step, we make use of the results

obtained by Zhang et al. (2015) in their hydrodynamical simulations of El Gordo. The main purpose of our work is to check how often the initial conditions of their plausible models arise in a very large Λ CDM cosmological simulation with a side length of $6 h^{-1}$ cGpc developed by the Juropan Hubble Volume Simulation Project (Watson et al. 2013).¹ We refer to this as the Jubilee simulation.

The structure of this paper is as follows: We begin by describing how we find analogues to El Gordo (Section 2). The pre-merger configuration obtained from hydrodynamical simulations (Section 2.1) is compared with the properties of galaxy cluster pairs in the Jubilee simulation (Section 2.2) using appropriate selection criteria (Section 2.3). We then describe our statistical analysis and its results in Section 3. We perform the analysis in two different ways: the power-law method (Section 3.2) and light-cone tomography (Section 3.3). We combine our results with the Kraljic & Sarkar (2015) analysis of the Bullet Cluster (Section 3.4). We then discuss our results in Section 4, both in a standard context (Section 4.2) and with a non-standard extended gravity law (Section 4.3). Finally, we summarize our most relevant conclusions in Section 5.

2 FINDING ANALOGUES TO EL GORDO

2.1 Hydrodynamical simulations of the merger

Donnert (2014) conducted the first attempt to reproduce the main characteristics of El Gordo with hydrodynamical simulations in a Λ CDM cosmology. Their simulation used a Hernquist dark matter density profile (Hernquist 1990), with the gas following the β -model of Cavaliere & Fusco-Femiano (1978) with $\beta = 2/3$. They inferred the following model parameters for the El Gordo cluster using the Menanteau et al. (2012) observations: main cluster mass $M_{200,1} = 1.9 \times 10^{15} M_{\odot}$, secondary cluster mass $M_{200,2} = 8.1 \times 10^{14} M_{\odot}$, main cluster radius $R_{200,1} = 2.55$ Mpc, secondary cluster radius $R_{200,2} = 1.925$ Mpc, relative infall velocity $V_{\text{inf}} = 2600$ km s⁻¹, gas fraction $f_b(< r_{200}) = 0.17$ for both clusters, and a small impact parameter of $b = 20$ kpc. Their model was implemented with the magnetohydrodynamics-smoothed particle hydrodynamics (MHD-SPH) code GADGET-3 (Springel 2005; Dolag & Stasyszyn 2009). The model was able to reproduce the total X-ray luminosity, the core distance of the clusters, and the observed offset between the X-ray and SZ centroids. However, it was not able to properly reproduce the X-ray morphology: Only one tail comes out of the peak emission instead of the observed two. Donnert (2014) attributed this to the lack of substructure in the simulation, and to the possibility that the parent cluster of El Gordo was a highly disturbed system even before it reached the observed configuration.

The hydrodynamical simulations of Molnar & Broadhurst (2015) used a Navarro–Frenk–White (NFW; Navarro, Frenk & White 1997) dark matter density profile and a truncated non-isothermal β -model with $\beta = 1$ for the gas. They also constrained their model parameters with the Menanteau et al. (2012) observations, but used slightly different values to Donnert (2014): $M_{200,1} = 1.13 \times 10^{15} M_{\odot}$, $M_{200,2} = 1.02 \times 10^{15} M_{\odot}$, $R_{200,1} = 2.304$ Mpc, $R_{200,2} = 1.944$ Mpc, $V_{\text{inf}} = 2250$ km s⁻¹, $f_b(< r_{200}) = 0.14$ for both clusters, and $b = 300$ kpc. They used the adaptive mesh refinement (AMR) code FLASH (Fryxell et al. 2000). Their use of an AMR code instead of an SPH code (as used in Donnert 2014) was motivated by the fact that an SPH code is prone to suppress turbulent mixing of the gas, which can play an important role in El Gordo (Bauer & Springel 2012). With this

¹<https://jubilee.ft.uam.es/>

model, Molnar & Broadhurst (2015) managed to reproduce its two-tailed X-ray morphology, but they obtained a low X-ray luminosity with respect to the observations. The projected distance between the clusters and the offset between their SZ and X-ray centroids were also larger than those observed.

The most recent simulation of El Gordo was conducted by Zhang et al. (2015), who explored the parameter space in greater detail. They chose the NFW profile for the dark matter and the Burkert (1995) profile for the gas. Their simulations used both the SPH code and an AMR code as each offers different advantages. In particular, the GADGET-2 code (SPH based) was used to perform a large number of merger simulations, exploiting the higher computational speed and effective resolution of SPH codes. Once they identified which simulations could be potential candidates to reproduce the El Gordo merger, they used the FLASH code (AMR based) to resimulate those mergers with a better handling of the shocks, eddies, and fluid instabilities that arise in the merging process. They ran 123 simulations for different parameters to check which gave a better fit to the El Gordo observations of Menanteau et al. (2012). Their simulations can be divided into the following two classes of model:

(i) Model A: extremely energetic head-on collisions. This category includes the Donnert (2014) and Molnar & Broadhurst (2015) simulations. The parameters derived by Zhang et al. (2015) that best match the observed features of El Gordo are: $M_{200,1} = 1.3 \times 10^{15} M_{\odot}$, $M_{200,2} = 6.5 \times 10^{14} M_{\odot}$, $R_{200,1} = 1.66$ Mpc, $R_{200,2} = 1.32$ Mpc, $V_{\text{inf}} = 3000 \text{ km s}^{-1}$, $f_b(< r_{200}) = 0.1$ for both clusters, and $b = 300$ kpc. This set of parameters – referred to as ‘fiducial model A’ by the authors – can generate an X-ray surface brightness distribution similar to the observations. It can also reproduce the observed mass density distribution. However, the twin-tailed X-ray morphology is smaller and more asymmetric than that observed, and only appears when the clusters are at a projected distance of 600 kpc, smaller than the weak lensing analysis of Jee et al. (2014) that shows the cores are separated by ≈ 700 kpc. The offset between the SZ and X-ray centroids is also smaller than that observed, while there is a lack of extended X-ray emission in the outer region of the merger. From this and the other Model A simulations of Zhang et al. (2015), they reached the following conclusions: (a) Collisions with smaller impact parameter are more violent; (b) having a smaller radius for the secondary progenitor cluster may lead to an increase of the X-ray emission in the cluster centre, but only one tail is formed; and (c) having a more unequal mass ratio in this scenario would make the collision less violent, preventing the destruction of the primary cluster’s gas core and thus leading to the formation of two peaks in the X-ray emission, contradicting observations. These conclusions are in agreement with the results that Donnert (2014) and Molnar & Broadhurst (2015) obtained from their simulations.

(ii) Model B: These simulations include off-centre collisions of two massive clusters with $b \geq 500$ kpc, making the collisions less violent than case A mergers. The parameters that provide the best fit for the observed characteristics of El Gordo are: $M_{200,1} = 2.5 \times 10^{15} M_{\odot}$, $M_{200,2} = 0.7 \times 10^{15} M_{\odot}$, $R_{200,1} = 2.06$ Mpc, $R_{200,2} = 1.35$ Mpc, $V_{\text{inf}} = 2500 \text{ km s}^{-1}$, gas fractions $(f_{b1}, f_{b2}) = (0.05, 0.1)$ for the main and secondary clusters, respectively, and $b = 800$ kpc. This is referred to as ‘fiducial model B’. The simulation of a merger with these parameters reproduces the temperature and X-ray luminosity of El Gordo, as well as a two-tailed X-ray morphology with a closer resemblance to observations than Model A. The projected distance between the two clusters when this structure is observed is 780 kpc, higher than that estimated observationally but still closer to the observational value than the model A result. The

offset between the SZ and X-ray centroids is also slightly smaller than that observed. An aspect of El Gordo that the model could not reproduce is the X-ray emission in the outer region of the merging cluster. In an attempt to correct for this, Zhang et al. (2015) changed the gas fraction of the main and secondary clusters to 0.11 and 0.12, respectively, while leaving unchanged the other parameters of fiducial model B. This ‘Extended Model B’ reproduces the X-ray extension while also getting an offset between the SZ and X-ray centroids closer to the observed offset. However, this changes the distance between the cluster centres to 890 kpc, significantly larger than that observed. After carrying out several Model B simulations, the authors concluded that: (a) the relative velocity needed to reproduce the two-tailed X-ray morphology is lower in Model B ($V_{\text{inf}} \approx 1500\text{--}2500 \text{ km s}^{-1}$ rather than $\approx 3000 \text{ km s}^{-1}$); (b) to have a single X-ray peak in a Model B scenario, the gas fraction of the main cluster must be lower than that of the secondary, but if the gas fraction of the secondary is too large, this would lead to the formation of an unrealistic bright gas core in the centre; (c) mergers with a mass ratio $\lesssim 2$ or $\gtrsim 5$ lead to the formation of a more asymmetric X-ray morphology; (d) a smaller core radius of the secondary cluster leads to brighter X-ray emission in the core and a larger gradient in the X-ray emission; and (e) a relatively large ($\approx 2.5 \times 10^{15} M_{\odot}$) main cluster mass is needed to generate the total X-ray luminosity of El Gordo, which is well constrained observationally.

Out of all these models, it is clear that the observed properties of El Gordo are best reproduced in the Zhang et al. (2015) fiducial Model B. The values chosen in this model reproduce El Gordo’s characteristics fairly well, and are also in agreement with weak lensing observations – the total mass $M_{200,T} = 3.2 \times 10^{15} M_{\odot}$ is very consistent with the weak lensing analysis of Jee et al. (2014). The mass ratio between the clusters is slightly higher (more unequal) in fiducial Model B than that in the weak lensing observations, though still within uncertainties. We expect that detailed modelling of the collision provides a far better guide to the mass ratio. Therefore, our main analysis uses the parameters of Model B in Zhang et al. (2015) as our reference for the pre-merger configuration of El Gordo. It is also their recommended model, as evidenced by them conducting an Extended Model B to further optimize the fit to observations.² In Section 4.2, we discuss their Model A and rerun some of our analyses for this case.

2.2 The Jubilee simulation

The Jubilee project consists mainly of two large N -body simulations: Small Jubilee, a $3.072 h^{-1} \text{ cGpc}$ cubic box with 3072^3 particles; and Big Jubilee, a $6 h^{-1} \text{ cGpc}$ cubic box with 6000^3 particles. Both use the N -body code CUBEP³M (Harnois-Déraps et al. 2013) and were run on the Juropa Supercomputer at the Jülich Supercomputer Centre.³ These simulations assume a Λ CDM cosmology with cosmological parameters $\Omega_{\text{m},0} = 0.27$, $\Omega_{\Lambda,0} = 0.73$, $h = 0.7$, $\sigma_8 = 0.8$, $n_s = 0.96$, and $\Omega_{\text{b},0} = 0.044$. The particle mass is $7.49 \times 10^{10} h^{-1} M_{\odot}$ (section 2 of Watson et al. 2014b). The Jubilee project currently only provides access to results of the Big Jubilee simulation, which have been post-processed in different catalogues with the algorithms known as Amiga Halo Finder (AHF), Friends of Friends (FoF), and Spherical Overdensity (SO), with the latter still under construction

²Since our analysis considers only the total mass of each cluster, it is not sensitive to their gas fractions. Thus, it is not relevant to discuss whether Extended Model B would be preferred over the fiducial Model B.

³<https://www.fz-juelich.de/portal/EN/Home/>

(for citations and comparisons between these algorithms, see Knebe et al. 2011 and Watson et al. 2014a). For this project, we choose the AHF catalogue since it provides a more detailed output with more available parameters.

AHF is an AMR code that identifies and hierarchically classifies structures inside a N -particle simulation box (Gill 2004; Knollmann & Knebe 2009). Its volume is covered with a regular grid of user-defined size. In each cell, the particle densities are calculated using a triangular shaped cloud weighting scheme (Hockney & Eastwood 1988). If the particle density exceeds a given threshold, the cell will be refined and covered by a finer grid with half the linear cell size. This process is repeated until a grid is reached that needs no further cell refinement. Following this procedure yields a grid hierarchy starting from the finest grid to the coarsest in such a way that isolated regions from the finer grids identified as possible haloes are linked to their corresponding volumes in the coarser grid.⁴ From this, a tree of nested grids is constructed. In case two patches that are isolated on one level link into the same patch on the next coarser grid, the two branches of the grid tree join. Once this is done for each nest of grids, the grid forest is constructed and the classification of substructure can be made – starting from the coarsest level downward to the finer levels, once the finer level splits up into two or more isolated patches, the patch containing the most particles is chosen as the main branch, while the others are classified as substructures.

The main properties of each halo are its catalogue radius r_{Jub} and catalogue mass M_{Jub} , with ‘Jub’ subscripts denoting values derived from the Jubilee catalogue. This determines the average density $\bar{\rho} (< r_{\text{Jub}})$ within the radius r_{Jub} . We use this to obtain the overdensity

$$\Delta_{\text{Jub}} \equiv \frac{\bar{\rho} (< r_{\text{Jub}})}{\rho_c}, \quad \text{where} \quad (1)$$

$$\rho_c \equiv \frac{3H^2}{8\pi G} \quad (2)$$

is the critical density of the universe at the snapshot redshift, when the Hubble constant is H . In the Jubilee catalogues, $\Delta_{\text{Jub}} = 178$. For a few cases, we checked that the values of Δ_{Jub} listed in the catalogue correspond to our calculations using equation (1). In the catalogue files, distances are in cMpc/h , masses in M_\odot/h , and peculiar velocities in km s^{-1} , where h is the present Hubble constant H_0 in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. We assume $h = 0.7$ throughout this article for consistency with the Jubilee simulations. We convert co-moving quantities to physical for our analysis, requiring us to add the Hubble flow velocity to the peculiar velocities listed in the catalogue.

The overdensity Δ_{Jub} used in the Jubilee catalogues is not the standard value of 200. For comparison with constraints derived from hydrodynamical simulations (Section 2.1), we convert the Jubilee quantities using the procedure described below. We define the scaled density within the radius r_{Jub} as $\tilde{\rho} \equiv \Delta_{\text{Jub}}/200$, where we use \tilde{q} to denote q/q_{200} for any quantity q with value q_{200} when defined within the radius r for which $\bar{\rho} (< r) = 200 \rho_c$. We guess the scaled catalogue mass \tilde{m} and thereby determine

$$\tilde{r} = \sqrt[3]{\frac{\tilde{m}}{\tilde{\rho}}}. \quad (3)$$

To make further progress, we assume that the cluster follows the NFW profile (Navarro et al. 1997) with concentration parameter c .⁵

⁴The volume covered by a fine grid is a subset of the volume covered by the coarser grids.

⁵ $r_{200} \equiv cr_s$, where r_s is the radial scale of the profile.

We use equation (4) of Duffy et al. (2008) to get that

$$c = 6.71a^{0.44} \left(\frac{M_{200}}{2 \times 10^{12} h^{-1} M_\odot} \right)^{-0.091}, \quad (4)$$

where $a \equiv 1/(1+z)$ is the cosmic scale factor. For a NFW profile, we must have that

$$\tilde{m} \left[\ln(1+c) - \frac{c}{1+c} \right] = \ln(1+c\tilde{r}) - \frac{c\tilde{r}}{1+c\tilde{r}}. \quad (5)$$

To ensure that this equation is satisfied, we set the difference between its left and right sides to 0 by varying our guess for \tilde{m} using the Newton–Raphson algorithm. In this way, we obtain m_{200} and r_{200} for use in subsequent analyses. We call these quantities the virial mass and radius, respectively.

2.3 Selection criteria

Our main goal is to find galaxy cluster pairs on an orbit similar to that of the El Gordo progenitors, whose properties we obtain using the hydrodynamical simulations discussed in Section 2.1. To avoid contamination from galaxies, we restrict our sample to only those haloes with $M_{200} \geq 3.5 \times 10^{13} M_\odot$. This leaves us with 4129 462 haloes, which we assume are all in the galaxy cluster regime.

Comparing all of these haloes with each other would be extremely computationally expensive. To improve the efficiency of our pair-finding algorithm, we only consider the region within some distance r_{max} of a halo with virial mass M , where

$$r_{\text{max}} = r_0 \sqrt[3]{\frac{M}{M_0}}, \quad (6)$$

with $r_0 = 40 \text{ Mpc}$ and $M_0 = 10^{13} M_\odot$. We can estimate the orbital period of any pair by applying Kepler’s laws. Since $r_{\text{max}} \propto M^{1/3}$, the period is independent of M , allowing us to calculate it for a cluster with mass M_0 . In this case, the minimum possible period P_0 arises if the total mass takes the highest possible value of $4.6 M_0$ (equation 9), the apocentre is r_0 , and the pericentre is 0, making the semimajor axis $r_0/2$. For this configuration,

$$P_0 = 2\pi \sqrt{\frac{\left(\frac{r_0}{2}\right)^3}{4.6 G M_0}} = 1220 \text{ Gyr}. \quad (7)$$

Thus, it is clear that we do not miss any pairs that could have turned around from their initial expansion and subsequently reached pericentre within the lifetime of the Universe. None the less, our restriction on the maximum separation of each pair greatly reduces the computational cost of our algorithm as the simulation box size is $(6 h^{-1} \text{ cGpc})^3$.

For a pair to be analogous to El Gordo, it should also have turned around from the cosmic expansion (Section 2.3.1), and have an appropriate total mass and mass ratio (Section 2.3.2), infall velocity (Section 2.3.3), redshift (Section 2.3.4), and impact parameter (Section 2.3.5). We discuss these criteria next, but usually omit the last condition to be conservative.

2.3.1 Requiring turnaround

Among the pairs with separation $< r_{\text{max}}$, we consider only those that have turned around from the cosmic expansion to avoid, e.g. two clusters separated by 30 Mpc with velocities close to the Hubble flow. In other words, the pairwise relative separation \mathbf{r}_{rel} and relative velocity \mathbf{v}_{rel} must satisfy

$$\mathbf{r}_{\text{rel}} \cdot \mathbf{v}_{\text{rel}} < 0. \quad (8)$$

This restricts our sample to cluster pairs that may subsequently undergo a close interaction reminiscent of El Gordo. Since most of the time in an orbit is spent close to apocentre, we should easily be able to find El Gordo analogues in this manner if they are common in the Jubilee simulation. A very small number of analogues could be missed because, e.g. at the time of the snapshot, an interacting pair is caught shortly after pericentre. However, given the significant amount of dynamical friction expected in an interaction like El Gordo, we expect this to be a fairly short-lived phase compared to the pre-merger phase. Even if a cluster pair is after pericentre at the snapshot time, it could still satisfy equation (8) provided the pair has turned around for a second time. This may be appropriate to the case of El Gordo (Ng et al. 2015). Our analysis also misses El Gordo progenitors that will turn around after the snapshot time. However, the time required to turn around and reach pericentre means that we would need to find such pairs at a much higher redshift than the $z = 0.87$ of El Gordo. It would also be very difficult to know for sure whether a pair will indeed turn around subsequent to the snapshot. In what follows, we will assume that only a small fraction of possible El Gordo progenitors are missed by imposing the turnaround condition.

2.3.2 Mass

The fact that El Gordo is an interaction between similar mass clusters is expected to significantly lower the predicted number of analogous pairs in the survey region. This is because minor mergers are much more common than major mergers. According to Model B of Zhang et al. (2015), the mass ratio of the El Gordo progenitors was 3.6. Therefore, we require each galaxy cluster pair to satisfy

$$\frac{m_{\max}}{m_{\min}} \leq 3.6, \quad (9)$$

where m_{\min} and m_{\max} are the virial masses of the member with the lower and higher mass, respectively.

The total mass M of El Gordo is one of the most important parameters. Model B of Zhang et al. (2015) implies that $M = 3.2 \times 10^{15} M_{\odot}$, so we require

$$\tilde{M} \equiv \log_{10} \left(\frac{m_{\min} + m_{\max}}{M_{\odot}} \right) > \tilde{M}_{\text{EG}} = 15.50. \quad (10)$$

This condition leaves us with no analogous systems in the entire Jubilee volume, so our statistical analysis (Section 3) is based on how the number of analogues changes as the limit on \tilde{M} is increased from a much lower value. The results are extrapolated up to the El Gordo value of 15.50, allowing us to obtain reliable statistics even though $\ll 1$ analogous pair exists in the Jubilee volume. The accuracy of this approach is discussed further in Section 4.1.

2.3.3 Infall velocity

To enable a comparison between galaxy cluster pairs caught at different phases of their orbit in a Jubilee snapshot, we use energy conservation under a point mass Newtonian potential to determine the relative velocity $v(2R_{\text{T}})$ when the clusters are separated by $2R_{\text{T}}$, where R_{T} is the sum of their virial radii. We then define a scaled velocity \tilde{v} to facilitate a fair comparison between cluster pairs with a very different total mass.

$$\tilde{v} \equiv v(2R_{\text{T}}) \div \sqrt{\frac{GM}{R_{\text{T}}}}, \quad (11)$$

where M is the sum of the cluster virial masses. The idea is to scale the relative velocity to the escape velocity v_{esc} . We set $\tilde{v} = 0$ for a pair with insufficient energy to reach a separation of $2R_{\text{T}}$. To match the observed properties of El Gordo, we require that $\tilde{v} > \tilde{v}_{\text{EG}} = 1.24$ for our nominal Model B. If instead we use Model A, \tilde{v}_{EG} becomes 1.77. Although a detailed analysis of the Bullet Cluster is beyond the scope of this work, for comparison we use table 2 of Lage & Farrar (2014) to estimate that its \tilde{v}_{BC} is only 0.80.

2.3.4 Redshift

The observed merging configuration of El Gordo is viewed at $z = 0.870$. However, we expect our procedure to find possible progenitor pairs before they interact. This is because:

- (i) it is difficult for the position-based AHF (Knollmann & Knebe 2009) to distinguish closely interacting pairs, and
- (ii) most of the time in an orbit is spent near apocentre.

Thus, we can only explain the observed configuration of El Gordo if we find progenitors at a slightly higher redshift. For this, the temporally closest snapshot available in the Jubilee simulation is at $z = 1$. The time difference between $z = 1$ and 0.870 is $\Delta t = 559$ Myr.

To check if this time lag suits the aforementioned purpose, we compare it with the pre-merger timeframes obtained with hydrodynamical simulations of El Gordo. Zhang et al. (2015) estimate that its observed state corresponds to 140 Myr after pericentre. The total time lag should be much larger as clusters spend most of their time going from apocentre to pericentre.

To get an idea of how long the clusters spend before pericentre, we assume that they are in free fall at their escape velocity ($\tilde{v} = 1$) and that their initial separation $r = R_{\text{T}}$, since it would be difficult for AHF to distinguish clusters with overlapping virial volumes. Assuming that the clusters are point masses of M_{200} , the free-fall time t_{ff} is

$$t_{\text{ff}} = \frac{r^{3/2} \sqrt{2}}{3 \sqrt{GM_{200,T}}} = 748 \text{ Myr}. \quad (12)$$

Since \tilde{v} is actually 1.24, a better estimate would be if we reduce t_{ff} by this factor, yielding 603 Myr. Adding this to the post-pericentre time lag of 140 Myr given in table 2 of Zhang et al. (2015), we obtain a total time lag of $\Delta t = 743$ Myr. Thus, our estimate of 559 Myr is conservative even if all our candidate pairs in the Jubilee simulation have a separation of R_{T} . However, the actual separations are typically much larger (Fig. 2). This is partly because clusters spend relatively little time near pericentre, but also due to the difficulties of AHF in resolving clusters that are merging. Other halo finders such as ROCKSTAR (Behroozi, Wechsler & Wu 2013) use velocity data to alleviate this problem. This is not the case with AHF, so the pairs we identify must be more widely separated. According to equation (12), this implies a much longer Δt .

Other studies also obtain a significantly larger time lag than that estimated above. Donnert (2014) used an initial time ($t = 0$) when the cluster separation is slightly above the sum of their virial radii (see their fig. 4). They found that the observed configuration of El Gordo takes place at $t = 1750$ Myr, shortly after a first passage through pericentre. Ng et al. (2015) propose two different scenarios that could describe today's observations of El Gordo: The clusters could either be right after their first passage through pericentre (as suggested by Donnert 2014), or they could be returning to pericentre for a second encounter. The first case corresponds to a time since pericentre of $t_{\text{out}} = 460$ Myr, while the second scenario – favoured in the author's conclusions – corresponds to a time since pericentre

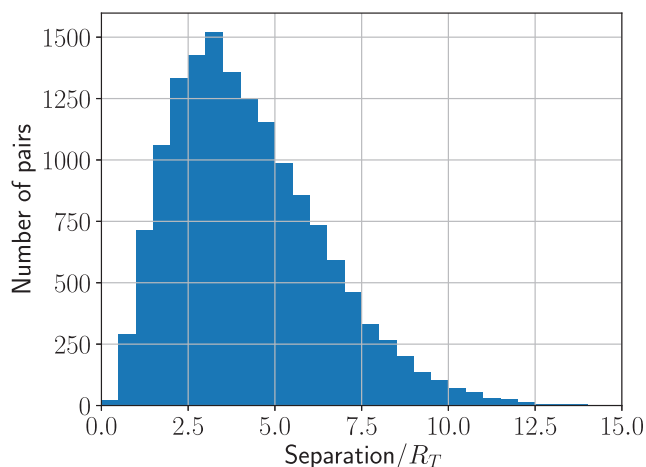


Figure 2. Histogram of the ratio \tilde{d} between the separation of candidate El Gordo analogue pairs and the sum of their virial radii. The pairs shown here are identified at $z = 1$ and satisfy the turnaround, mass ratio, and \tilde{v} conditions, but have a smaller total mass than El Gordo (no pairs would be left otherwise). Most pairs have $\tilde{d} > 4$. There are almost no pairs with $\tilde{d} < 0.5$, and only 314/15035 have $\tilde{d} < 1$.

of $t_{\text{ret}} = 910$ Myr. This exceeds our estimated Δt even without any time lag before pericentre.

Leaving a longer time lag between the Jubilee and observed configurations would mean choosing a higher redshift for our analysis. This lowers the chance of finding an El Gordo analogue as the number of fast collisions of massive objects decreases with increasing redshift. Therefore, we conservatively adopt $z = 1$ as the redshift at which we search for El Gordo analogues. In reality, the time lag could be significantly larger, so a higher z may be more appropriate.

2.3.5 Impact parameter

To get a sufficiently strong interaction between the clusters, their impact parameter should be rather small. To calculate this for cluster pairs in the Jubilee simulation, we begin by using energy conservation to estimate their relative velocity v_∞ at large separation. For pairs without sufficient energy to escape, we record $\tilde{b} = 0$. Otherwise, we calculate

$$\tilde{b} \equiv \frac{|\mathbf{r}_{\text{rel}} \times \mathbf{v}_{\text{rel}}|}{v_\infty R_T}. \quad (13)$$

Model B of Zhang et al. (2015) implies that $\tilde{b}_{\text{EG}} = 3.42$, which is appropriate for an off-centre collision. We do not directly include the requirement of a lower \tilde{b} to be more conservative. This condition should exclude only a small proportion of pairs (Section 3.2).

3 STATISTICAL ANALYSIS AND RESULTS

We fit the mass distribution of pairs analogous to El Gordo according to some subset of the criteria discussed in Section 2.3. The fitting procedure is described in Section 3.1. These fits are used in two methods to quantify the likelihood of observing El Gordo within the surveyed region: the power-law method (Section 3.2) and light-cone tomography (Section 3.3). In both cases, the use of fitting functions is necessary because the mass of El Gordo is larger than that of the most massive pair to satisfy the other conditions. However, we will see that it is only necessary to extrapolate \tilde{M} by ≈ 0.6 dex. Section 3.4

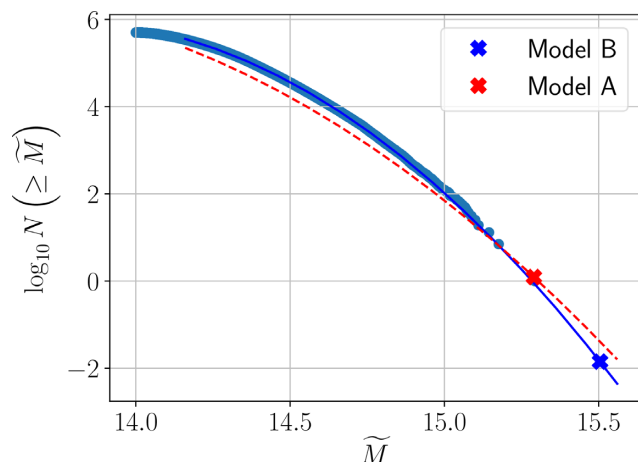


Figure 3. The blue data points show the cumulative \tilde{M} distribution for haloes in the Jubilee simulation at $z = 1$ that have turned around from the cosmic expansion and have a mass ratio ≤ 3.6 (Model B). The blue solid line is the quadratic fit to this distribution. The red dashed line is the analogous fit when we instead require a mass ratio ≤ 2 (Model A). For clarity, individual points are not shown in this case. The crosses indicate the \tilde{M} of each model.

discusses the implications of our results in combination with the Bullet Cluster.

3.1 Fitting the mass function of El Gordo-like pairs

With the 499 285 pairs that pass the mass ratio and turnaround conditions, we sort the pairs into ascending order of total mass as quantified by \tilde{M} (equation 10). This lets us obtain the cumulative distribution of \tilde{M} , i.e. the number of pairs $N(\geq \tilde{M})$ whose \tilde{M} equals or exceeds some value. We show this distribution in Fig. 3 for both Model A and Model B parameters. There is a slight difference caused by the smaller mass ratio in Model A, which in the previous step reduces the number of candidates to 379 612. We adopt the Model B values as nominal, but it is also helpful to consider Model A because its lower mass for El Gordo is less problematic. We discuss Model A in more detail in Section 4.2.

By fitting the mass distribution analytically, we can extrapolate the number of pairs analogous to El Gordo that have $\tilde{M} > \tilde{M}_{\text{EG}}$, which is the more problematic side for Λ CDM. Therefore, choosing the correct fitting function for the mass distribution is critical to achieving an accurate estimate of how unlikely the observation of El Gordo is in a Λ CDM context. We initially choose a quadratic fit of $\log_{10} N(\geq \tilde{M})$ against \tilde{M} since this seems better suited to the distribution. This provides a very good fit (Fig. 3). To check if it can be improved with a higher order fit, we compare the residuals to those of a cubic fit (Fig. 4). It is clear that a cubic would overfit the distribution and provide a poorer match to the Jubilee data at the critically important high-mass end. Therefore, we use a quadratic fitting function in the remainder of this article. The cumulative \tilde{M} distribution of pairs that pass the mass ratio, turnaround, and \tilde{v} conditions is

$$\log_{10} N(\geq \tilde{M}) = c_0 + c_1 \tilde{M} + c_2 \tilde{M}^2, \quad (14)$$

with the fitting coefficients listed in Appendix A for the Jubilee snapshots at $z = 0, 0.509, \text{ and } 1$. We use these fits to quantify the probability of finding analogues to the presumed El Gordo progenitors at $z = 1$. This is done using two analysis techniques, which we describe next.

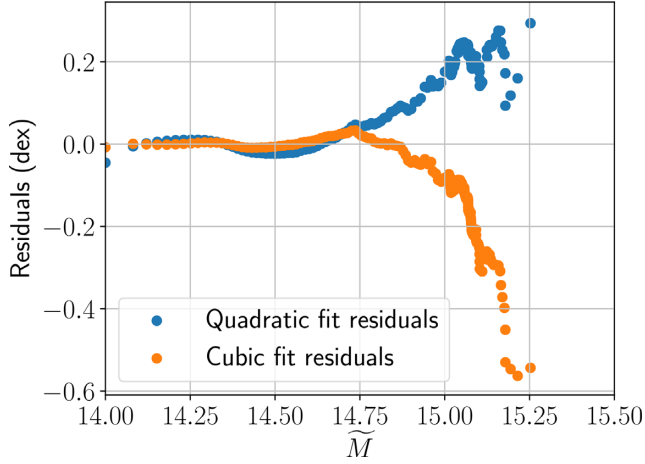


Figure 4. Residuals of the quadratic and cubic fits to the cumulative mass distribution for turned around halo pairs in the Jubilee simulation at $z = 1$ whose mass ratio is ≤ 3.6 .

3.2 Power-law method

Once we know the effective volume V_{eff} of the survey that discovered El Gordo, we can compare this to the Jubilee simulation volume to determine the expected number of El Gordo analogues in the survey volume. From this, we can quantify the P -value for a Λ CDM cosmology of the fact that El Gordo was discovered with its observed properties in the surveyed region. The basic principle is to find the chance that this contains a cluster collision with equally or even more extreme properties.

We extrapolate our fit to the cumulative pairwise mass distribution to get the number of El Gordo analogues in the $(6 h^{-1} \text{cGpc})^3$ simulation volume at $z = 1$. By repeating this at $z = 0.509$, we obtain the growth index

$$k \equiv \frac{\Delta \ln n}{\Delta \ln a}, \quad (15)$$

where n is the number of El Gordo analogues in the simulation volume, and the finite difference is taken between the Jubilee snapshots at $z = 1$ and 0.509 . We avoid using the $z = 0$ data because structure formation is slowed down at low z by the effect of dark energy, but its impact should not be too significant at $z \gtrsim 0.5$. When fitting the mass function using the widest available range ($\tilde{M} \geq 14.0$), we obtain that $k = 24.81$ for the $\tilde{v} > 0$ condition and $k = 35.55$ for the more realistic $\tilde{v} > \tilde{v}_{\text{EG}}$ condition.

We then use k to determine V_{eff} . The main idea is that since $k \gg 1$, the effective survey depth is limited by the fact that analogues to El Gordo rapidly become very rare with increasing z . To find the number of El Gordo analogues at $a < 0.5$ along our past light-cone, we should ideally integrate n over $a \leq 0.5$ while accounting for the variation of n with a . For $k \gg 1$, this is approximately equivalent to neglecting the redshift dependence of n and only integrating out to, e.g. $a = 0.45$. Since the co-moving number density of El Gordo analogues increases with a as $n \propto a^k$ with k given by equation (15), the effective ‘depth’ of the survey is

$$\Delta a \approx \frac{a}{k+1}, \quad (16)$$

where $a = 0.5$. In practice, this is not exactly correct since the co-moving volume V_c per unit a along our past light-cone satisfies

$$\frac{dV_c}{da} = \frac{cAd_c^2}{aH}, \quad (17)$$

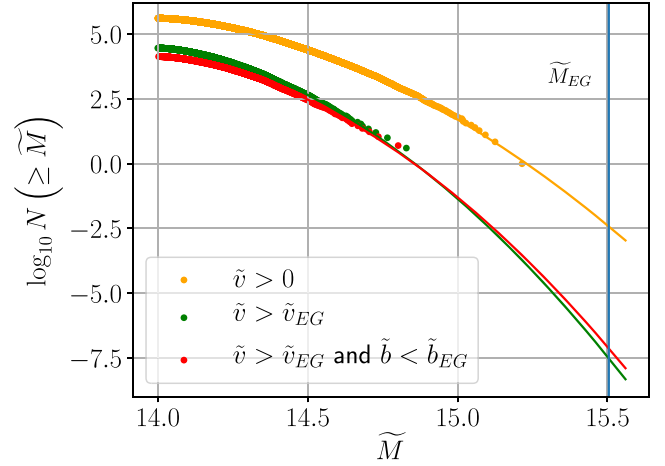


Figure 5. The cumulative \tilde{M} distribution with the conditions $\tilde{v} > 0$ (orange), $\tilde{v} > \tilde{v}_{\text{EG}}$ (green), and $\tilde{v} > \tilde{v}_{\text{EG}}$ plus $\tilde{b} < \tilde{b}_{\text{EG}}$ (red). The blue vertical line marks the El Gordo \tilde{M} .

where A is the sky area of the survey in natural units (steradians), H is the Hubble constant at epoch a , and d_c is the co-moving radial distance to an object at that epoch.

$$d_c \equiv \int_{t_i}^{t_0} \frac{c dt}{a}. \quad (18)$$

The integral must be taken up to the present time t_0 from the time t_i corresponding to the epoch at which we wish to know d_c . If this is the epoch when $a = 0.5$, we have that $H \propto a^{-1.16}$ while $d_c \propto a^{-1.50}$, so $dV_c/da \propto a^{-2.83}$. This means that the number of analogues per unit a along the past light-cone is $\propto a^{k-2.83}$, so the effective survey depth is

$$\Delta a = \frac{a}{k-1.83}. \quad (19)$$

Since we know the sky coverage of the survey, its effective co-moving volume is then

$$V_{\text{eff}} = Ad_c^2 \Delta d_c, \quad (20)$$

which gives $V_{\text{eff}} = 6.33 \times 10^8 \text{cMpc}^3$ ($4.17 \times 10^8 \text{cMpc}^3$) when we require $\tilde{v} > 0$ ($\tilde{v} \geq \tilde{v}_{\text{EG}}$). To get the co-moving volume in the past light-cone between a and $a - \Delta a$, we need the difference in d_c between these epochs, which we call Δd_c . We have simplified our equations by exploiting the fact that the angular diameter distance is d_c/a .

The basic characteristics of El Gordo were constrained for the first time in a survey with an area $A = 755 \text{deg}^2$ (Menanteau et al. 2012). As discussed in Section 2.3, we also add several conditions to obtain a cumulative mass distribution for El Gordo-like pairs. We add these in stages to see how the number of analogues decreases as each condition is imposed. The conditions are: $\tilde{v} > 0$, $\tilde{v} \geq \tilde{v}_{\text{EG}}$, and $\tilde{v} \geq \tilde{v}_{\text{EG}}$ plus $\tilde{b} \leq \tilde{b}_{\text{EG}}$. Our results are shown in Fig. 5.

Since we only consider clusters with $m/M_\odot \geq 3.5 \times 10^{13} \approx 10^{13.5}$, there are no pairs less massive than $\tilde{M} = 13.5 + \log_{10}(2) = 13.80$. Edge effects are thus very significant at this mass. To avoid edge effects at the low-mass end, from now on we only consider pairs with $\tilde{M} > 13.5 + \log_{10}(4.6) = 14.16$ unless explicitly stated otherwise, allowing for the maximum allowed mass ratio of 3.6 (equation 9). Ideally, we would impose a much higher floor and consider only those cluster pairs with similar mass to El Gordo. However, such objects are very rare, so fitting only the very high mass end can

lead to an inaccurate estimation of n (see Fig. 5). This effect is more pronounced when we add the $\tilde{v} \geq \tilde{v}_{\text{EG}}$ condition because in the high mass bin most relevant to El Gordo, the number of pairs starts decreasing very rapidly with \tilde{v} once $\tilde{v} \gtrsim 0.6$ (Section 4.2). Therefore, our main results rely on all pairs with $\tilde{M} > 14.16$, maximizing the amount of Jubilee data used. This means that our results rely on the top 40–50 per cent of our pairwise mass distribution.

We use the survey volume V to find the expected number of analogues N_{EG} within it. Since $N_{\text{EG}} < 1$, a situation ‘as bad as or worse than’ reality for Λ CDM implies in this case that the survey region contains ≥ 1 El Gordo-like system. In a Λ CDM context, the probability of this occurring is given by standard Poisson statistics (Watson et al. 2014a).

$$P = 1 - \exp(-N_{\text{EG}}). \quad (21)$$

Once we have the P -value, we can express it as an equivalent number of standard deviations for a Gaussian distribution, which we call the χ value. By definition, P and χ are related to each other by the Gaussian distribution.

$$1 - \frac{1}{\sqrt{2\pi}} \int_{-\chi}^{\chi} \exp\left(-\frac{x^2}{2}\right) dx \equiv P. \quad (22)$$

In this contribution, we will usually convert a P -value into the statistical significance χ by solving this equation using the Newton–Raphson algorithm. The nominal χ values that we obtain from this analysis are $\chi = 4.61$ for the $\tilde{v} > 0$ condition and $\chi = 6.69$ for the more realistic $\tilde{v} \geq \tilde{v}_{\text{EG}}$ condition. Since this already exceeds the 5σ threshold, we do not show results for imposing both $\tilde{v} \geq \tilde{v}_{\text{EG}}$ and $\tilde{b} \leq \tilde{b}_{\text{EG}}$. To illustrate the effect of the \tilde{b} condition, we show how the $z = 1$ mass function changes if we additionally impose this condition (Fig. 5). Requiring that $\tilde{b} < \tilde{b}_{\text{EG}}$ only slightly reduces the number of analogue pairs. There is a small effect in the low-middle mass range, but not so much at the high-mass end. This is because orbits are more radial at high mass as there is less likely to be a third massive object that imposed a significant tidal torque. Because of this, adding the \tilde{b} condition makes the mass function flatter. Extrapolating this to higher masses can lead to nonsensical results such as obtaining a lower χ value *after* imposing the $\tilde{b} \leq \tilde{b}_{\text{EG}}$ condition. Thus, we do not show χ values for the case of imposing both $\tilde{v} \geq \tilde{v}_{\text{EG}}$ and $\tilde{b} \leq \tilde{b}_{\text{EG}}$, except to note that the latter would not by itself make matters much worse for Λ CDM because radial orbits are expected at the high mass of El Gordo. We also do not consider results for the $\tilde{v} > 0$ case to be particularly useful because requiring $\tilde{v} > 0$ is not sufficient for finding analogues to a cluster merger that clearly requires a high infall speed to reproduce the observed shock features (Section 2.3.3). Similarly, imposing no condition on \tilde{v} is not sufficient to match El Gordo – this is only shown for illustrative purposes.

Our quadratic fits to the mass function may not be perfectly accurate as there could be a weak cubic dependence. Since the mass function declines rather steeply, our fits to it are dominated by pairs near the low-mass end, causing the fits to prioritize very modest improvements here over a better fit at the high-mass end critical to our analysis. One way to check for the impact of any cubic or higher order trend is to restrict the range of \tilde{M} used in our fit. By focusing on the high-mass region, it is possible to minimize such systematic effects at the cost of higher random errors from Poisson noise. Therefore, we try several different values for \tilde{M}_{min} , the lowest value of \tilde{M} used in our quadratic fit. The idea is to check if our results differ much when fitting the whole data set or fitting towards the high-mass end. We have summarized the results of all these different tests in Fig. 6, where we have expressed the number of analogues expected

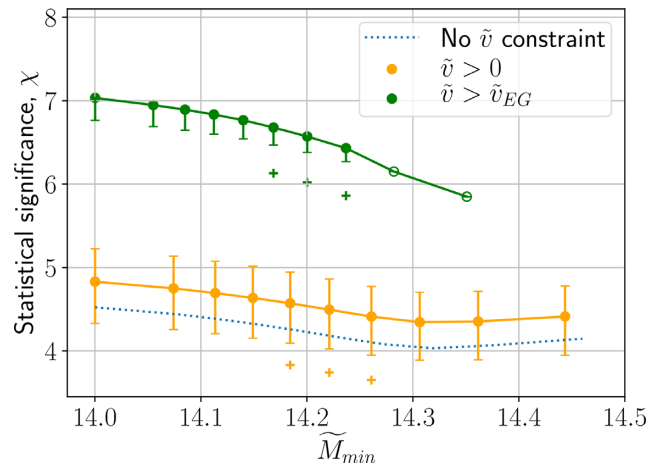


Figure 6. The statistical significance χ (equation 22) of the observation of an El Gordo-like object as defined in different ways, in terms of the minimum mass beyond which we fit the cumulative \tilde{M} distribution. Each line corresponds to a different set of conditions, which in increasing order of restrictiveness are: turnaround and mass ratio (blue dotted), $\tilde{v} > 0$ (orange), and $\tilde{v} > \tilde{v}_{\text{EG}}$ (green). Error bars on the $\tilde{v} > 0$ line show the effect of varying the El Gordo mass by ± 20 per cent, while those on the $\tilde{v} > \tilde{v}_{\text{EG}}$ line correspond to a ± 10 per cent uncertainty on \tilde{v}_{EG} . Along each line, each successive data point is based on 10 per cent fewer pairs at $z = 1$ in terms of the leftmost point; e.g. if this is based on 100 pairs, the second is based on 90 pairs and the third on 80 pairs, etc. The hollow circles indicate somewhat questionable results – error bars could not be reliably calculated as there were too few pairs. This is also the reason for missing upper error bars on the line for $\tilde{v} > \tilde{v}_{\text{EG}}$. The cross-shaped symbols represent the results of our light-cone tomography (Section 3.3), with the same relation between colour and the imposed conditions.

to be found in the El Gordo survey volume in terms of the χ value. This does not depend very much on \tilde{M}_{min} .

3.3 Light-cone tomography

To check that the previous analysis provides an accurate estimate of the probability that Λ CDM yields an El Gordo analogue in the surveyed volume, we compare the results with a different type of statistical analysis that we term light-cone tomography. The basic idea is to consider the distribution of pairs with El Gordo-like properties along our entire past light-cone, not just at $a \approx 0.5$. In particular, after requiring that pairs have similar dimensionless parameters to El Gordo, we consider to what extent its main dimensionful numbers (mass and redshift) are outliers to the distribution expected in Λ CDM.

The procedure can be described as follows: We apply our usual quadratic fit to the mass function (equation 14) in the Jubilee snapshots at $z = 0, 0.509$, and 1. The fit coefficients for the $\tilde{v} > \tilde{v}_{\text{EG}}$ condition are shown in Appendix A. Since we have three snapshots, we use a quadratic to fit the values for each coefficient against $\ln a$, allowing us to extrapolate the value of this coefficient to any other a . In other words, c_0 in equation (14) is now treated as a function of a , with its value $c_0(a)$ found by a quadratic interpolation from the three snapshots where we know c_0 . This tells us the co-moving number density of El Gordo-like pairs as a function of both \tilde{M} and a . For each bin in a along our past light-cone, the co-moving volume is given by integrating equation (20).

We use this information to obtain the number of pairs in the past light-cone of a $z = 0$ observer that lie within any given range of

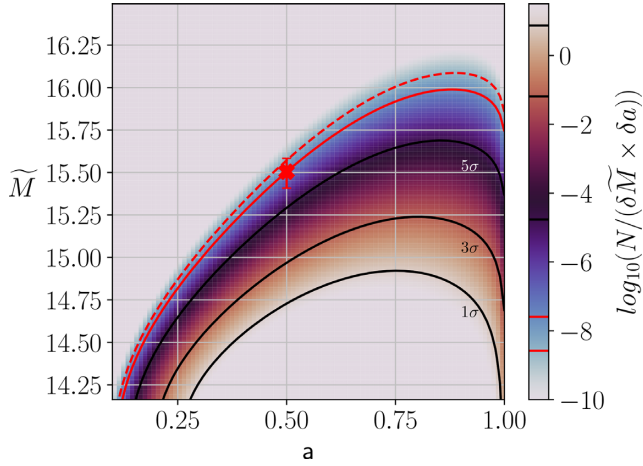


Figure 7. The number of analogues to El Gordo according to the turnaround, mass ratio, and $\tilde{v} \geq \tilde{v}_{\text{EG}}$ conditions per unit \tilde{M} per unit a along our past light-cone for the 755 deg^2 sky area in which it was discovered. The black contour lines show the 1σ , 3σ , and 5σ confidence regions, while the red contour at 6.16 σ goes through El Gordo (red cross, with error bar indicating ± 20 per cent uncertainty on the mass). Of the N_{EG} pairs outside this contour, only $0.1 N_{\text{EG}}$ pairs lie beyond the red dashed contour; i.e. most of the probability of observing El Gordo comes from between the red contours. The levels of all these contours are shown on the colour bar. The 5σ contour is at $\tilde{M} = 15.293$ when $a = 0.5$.

pair total mass and redshift. Using a 2D contour plot showing this quantity as a function of (\tilde{M}, a) , we get the contour level for El Gordo. We then infer the expected number (N_{EG}) of El Gordo-like objects in our past light-cone with (\tilde{M}, a) outside this contour. This lets us infer the corresponding P and χ values. Fig. 7 shows this contour plot for the $\tilde{v} > \tilde{v}_{\text{EG}}$ condition using a mass function fit with $\tilde{M}_{\text{min}} = 14.16$. The solid red line shows the critical contour through El Gordo, beyond which we expect $N_{\text{EG}} = 7.51 \times 10^{-10}$ pairs. For ease of reference, we call this the N_{EG} contour.

As El Gordo is clearly a significant outlier, most of these pairs lie only a little outside the N_{EG} contour. This can be seen with the red dashed line, which shows the contour beyond which there are $0.1 N_{\text{EG}}$ pairs. Thus, fully 90 per cent of N_{EG} arises from the thin strip between the red contours. Since the number density of pairs with respect to (\tilde{M}, a) is the same along each contour, the total number of pairs contributed by each a is roughly proportional to the gap between the N_{EG} and $0.1 N_{\text{EG}}$ contours – the wider the gap, the more important that a is. Thus, we can gain an idea of which cosmological epochs are most relevant to the probability of observing El Gordo in a Λ CDM universe. Fig. 7 shows that the gap between the $0.1 N_{\text{EG}}$ and N_{EG} contours is widest for $a \gtrsim 0.5$, which is also the range covered by the Jubilee snapshots we analyse. Therefore, even if our extrapolation to lower a is not perfectly accurate, this should not significantly affect our results – the total probability receives little contribution from $a \lesssim 0.5$.

Our light-cone tomography analysis indicates that when requiring $\tilde{v} > \tilde{v}_{\text{EG}}$, the statistical significance $\chi = 6.16$. This agrees fairly well with our power-law analysis (Fig. 6). Importantly, *both techniques agree that the Λ CDM model is falsified at $>6\sigma$ significance based on El Gordo alone.*

For our light-cone tomography, we consider cluster pairs all the way up to $a = 1$ in order to be conservative. This is not very realistic as there must have been some time lag between the observed state of El Gordo and the pre-merger stage at which we seek to capture El Gordo

analogues in the Jubilee simulation. As explained in Section 2.3.4, we estimate that this time lag is 559 Myr. Subtracting this from the 13.47 Gyr age of the universe in Λ CDM yields $a = 0.96$. In the real world, any El Gordo-like pair at $a > 0.96$ would simply not have enough time to evolve into the observed state. Thus, it may be more reliable to restrict our light-cone tomography to $a < 0.96$. This would slightly reduce N_{EG} and increase the significance χ . We found that χ rises by only ≈ 0.01 , so the time lag effect is not a major source of uncertainty at the present epoch. It may be more significant at $a \approx 0.5$, but we have already accounted for it by placing El Gordo at $z = 1$ instead of the observed $z = 0.87$ (Section 2.3.4). As discussed there, the time lag could be much longer than our assumed 560 Myr since most of the candidate pairs we identify have a rather large separation, so they would need a significant amount of time to reach pericentre (Fig. 2). Moreover, the time after pericentre could by itself contribute 910 Myr if El Gordo is observed after second turnaround, which better accounts for some observables (Ng et al. 2015).

3.4 Combined implications with the Bullet Cluster

The Bullet Cluster was discovered by Tucker et al. (1995) and is listed in their table 1. In their section 3, they indicate that this and the other clusters they discovered were identified by searching through 1435 fields of view of the Imaging Proportional Counter on the Einstein X-ray observatory (Giacconi et al. 1979). According to table 1 of the latter work, each field of view is a square with 1.25° sides. Thus, the Bullet Cluster was discovered in only 5.4 per cent of the sky, with the small sky coverage being due to poor spatial resolution that meant much of the sky was obscured by emission from foreground supernova remnants. Despite these difficulties, the extreme properties of the Bullet Cluster were already apparent in Tucker et al. (1998).

A detailed analysis of the Bullet Cluster is beyond the scope of this contribution, but we utilize the result of Kraljic & Sarkar (2015) that 0.1 pairs analogous to it should be observed out to its redshift of $z = 0.296$. They implicitly assumed that observing a cluster pair ‘as bad as or worse than’ the Bullet involved a lower redshift, which is possible for a nearby object since the co-moving volume per unit a declines rapidly with a when $a \approx 1$. We have made the opposite assumption since El Gordo is at a much higher redshift, so it is clear that the situation will be worsened for Λ CDM by going to higher z . The situation may be different for the Bullet, so we assume that the choice adopted by Kraljic & Sarkar (2015) is reasonable. However, their result of 0.1 analogous pairs is only valid for a survey that covers the full sky. This is apparent in their equation (4.5), which is directly analogous to our equation (20) – but with 4π instead of the sky area A . Accounting for the sky coverage of 5.4 per cent, we get that the surveyed region is expected to have $N_{\text{BC}} = 5.4 \times 10^{-3}$ pairs analogous to the Bullet Cluster in a Λ CDM context. This makes the Bullet Cluster a $\chi_{\text{BC}} = 2.78\sigma$ outlier.⁶

To approximately combine this with El Gordo, we add the squares of the individual χ values.

$$\chi_{\text{tot}}^2 = \chi_{\text{BC}}^2 + \chi_{\text{EG}}^2. \quad (23)$$

⁶This result could differ somewhat if Kraljic & Sarkar (2015) had used the dimensionless \tilde{v} instead of the infall velocity, since cluster pairs of lower mass (which are more common) can more easily match \tilde{v} than the actual infall velocity.

Table 1. Results of the power-law (Section 3.2) and light-cone tomography (Section 3.3) analyses for different \tilde{v} conditions. We show the expected number of analogues N in the surveyed region as a P -value [$P = 1 - \exp(-N)$] and the corresponding statistical significance χ (equation 22). Results for El Gordo (the Bullet Cluster) have a subscript EG (BC). The row $P_{\text{EG} + \text{BC}}$ shows the effect of combining our results with the Bullet Cluster, which is a $\chi_{\text{BC}} = 2.78\sigma$ outlier according to the results of Kraljic & Sarkar (2015) when accounting for the sky footprint of the discovery survey (see the text). We also show the impact of scaling the number of El Gordo and Bullet Cluster analogues by assuming that the surveys in which they were found covered the full sky (last row). This is a conservative estimate of how our conclusions would be weakened if the entire sky contains no additional problematic objects for Λ CDM.

| | $\tilde{v} > 0$ | | $\tilde{v} > \tilde{v}_{\text{EG}}$ | |
|--|--|--|---|---|
| | Power-law | Light-cone tomography | Power-law | Light-cone tomography |
| P_{EG} | 4.08×10^{-6} (4.61 σ) | 1.12×10^{-4} (3.86 σ) | 2.23×10^{-11} (6.69 σ) | 7.51×10^{-10} (6.16 σ) |
| $P_{\text{EG} + \text{BC}}$ | 5.13×10^{-7} (5.02 σ) | 1.20×10^{-5} (4.38 σ) | 3.98×10^{-12} (6.94 σ) | 1.24×10^{-10} (6.43 σ) |
| $P_{\text{EG} + \text{BC}}$ (full sky) | 6.09×10^{-5} (4.01 σ) | 1.36×10^{-3} (3.20 σ) | 5.04×10^{-10} (6.22 σ) | 1.55×10^{-8} (5.66 σ) |

The probability of a higher χ_{tot}^2 is then found using the standard formula for two degrees of freedom.

$$P = \exp\left(-\frac{\chi_{\text{tot}}^2}{2}\right). \quad (24)$$

This can be converted into a statistical significance χ for a single variable using our usual approach of applying equation (22).

Table 1 summarizes the main results of the power-law and light-cone tomography analyses, both when considering El Gordo alone and in combination with the Bullet Cluster. We see that the power-law and light-cone tomography analyses give similar results. The Bullet Cluster and El Gordo were discovered in a small fraction of the entire sky. To be conservative, we can assume that no additional cluster pairs are found in the rest of the sky that pose a problem to Λ CDM. We illustrate what effect this could have by scaling up N_{BC} and N_{EG} under the assumption that their discovery surveys covered the whole sky. This reduces χ_{BC} and χ_{EG} , but we are still left with an overall significance $\chi > 5$. Therefore, the power-law and light-cone tomography methods agree that *the Λ CDM model must be rejected at $>5\sigma$ confidence even if no additional problematic objects exist in the rest of the sky.*

In reality, full sky surveys such as the Planck survey (Planck Collaboration VIII 2011) have already found more clusters that could further increase the tension. One example is PLCK G214.6+37.0, a triple system of clusters at $z \approx 0.47$ that appear to be at an early stage of interaction. The total mass $M_{200} \approx 1.17 \times 10^{15} M_{\odot}$, while the mass ratio between its components is ≤ 1.4 . Another example is PLCK G287.0+32.9 at $z = 0.39$ (Planck Collaboration VI 2013), a quadruple system of clusters with a total mass of $M_{200} = 2.04_{-0.21}^{+0.20} \times 10^{15} M_{\odot}$ in which the largest cluster (10 \times more massive than the other components of the system) is undergoing a complex merger with one or more of the other clusters (Bagchi et al. 2011; Finner et al. 2017). We discussed a few other potentially problematic cases in Section 1. It is therefore unrealistic to assume that there are no other objects besides El Gordo and the Bullet Cluster that could entail a problem for Λ CDM. Moreover, even this very conservative assumption is insufficient to solve the tension (Table 1).

4 POSSIBLE EXPLANATIONS FOR EL GORDO

Table 1 shows that the results from our power-law and light-cone tomography analyses are in agreement, with a difference in statistical significance χ of ≈ 0.7 when requiring $\tilde{v} > 0$ and ≈ 0.5 for $\tilde{v} > \tilde{v}_{\text{EG}}$. Our most conservative estimate of the tension is $\chi_{\text{EG}} = 6.16$ based on light-cone tomography, so an uncertainty of 0.5 would not be

enough to reconcile Λ CDM with observations at 5σ confidence.⁷ Moreover, if χ_{EG} is combined with the 2.78σ tension caused by the Bullet Cluster, the combined significance rises to 6.43σ . These results account for the sky coverage of the respective surveys in which the objects were discovered. If we adjust our analysis by pretending that the surveyed area in each case was the whole sky, the significance is still close to 6σ . Thus, future surveys will not be able to alleviate the tension for Λ CDM even if they uncover no additional problematic objects.

Even with the more relaxed condition ($\tilde{v} > 0$) in which the high infall velocity of El Gordo is not taken into account, we still get a statistical significance close to 4σ , which makes the El Gordo mass at $z = 1$ an extremely unlikely feature on its own. Isolated clusters of this mass are not catastrophic for Λ CDM (e.g. fig. 17 of Jee et al. 2014). However, the paired nature of El Gordo substantially reduces the odds of finding such a system in the surveyed region. This very important effect was not considered in previous works that treated El Gordo as one object when quantifying the tension with Λ CDM (Menanteau et al. 2012; Jee et al. 2014; Sahlén et al. 2016). Of course, these analyses could not consider the high required infall velocity. Our analysis is the first to consider both the \tilde{M} and \tilde{v} of El Gordo in a standard context.

Even if El Gordo is treated as one object, the statistical methods used differ between studies. Unlike Jee et al. (2014), Buddendiek et al. (2015) showed in their fig. 8 that even a single $2 \times 10^{15} M_{\odot}$ cluster at $a = 0.5$ would be quite problematic for Λ CDM. This is also evident in fig. 5 of Watson et al. (2014a). Regardless of the exact frequency of such clusters at this redshift, a major merger of two such clusters would be even less frequent, since this can only occur if two individually rare objects formed close enough to turn around from the cosmic expansion within a few Gyr. It is thus not very surprising that the existence of El Gordo significantly challenges Λ CDM.

Beyond this qualitative agreement with previous studies, we are unable to compare our results with other similar works since they do not conduct an analysis nearly as detailed as ours. However, we can compare the results of our two statistical analyses among themselves. Both give $\chi > 6$, so we can be fairly confident in setting the lower limit of the χ value to 6. Therefore, the expected number of analogues to El Gordo in the surveyed region must be increased by >3 orders of magnitude just to reach the plausibility threshold of $\chi = 5$, which

⁷If we assign our calculated χ a Gaussian uncertainty of 0.5, then we can suppose that we overestimate χ by 0.5 σ . The overall likelihood of any y is then that of $\chi_{\text{tot}}^2 = y^2 + (6.16 - 0.5y)^2$ for 2 degrees of freedom. We never get $\chi_{\text{tot}}^2 < 30.4$, but a 5σ event corresponds to $\chi_{\text{tot}}^2 = 28.7$ (equation 24).

would make the observation of El Gordo as probable as a fair coin landing heads 21 times in a row.

The rather low P -values we obtain are almost certainly overestimates for various reasons. As discussed in Section 2.3.4, many of the Jubilee pairs we call ‘analogues’ are actually still rather widely separated at $z = 1$, so they would not evolve into the observed configuration by $z = 0.87$. Moreover, we have only searched for pairs whose orbit is similar to that of the presumed El Gordo progenitors. We have not allowed for the fact that only a small fraction of the orbit is spent close to pericentre, but we must be observing El Gordo at this phase. The P -values are also reduced by the requirement to have a sufficiently small impact parameter, though this by itself is not very problematic (Fig. 5).

4.1 Reliability of the Jubilee simulation

The length and mass scales relevant to El Gordo should be very well resolved in the Jubilee simulation: The minimum resolved halo mass is $1.49 \times 10^{12} h^{-1} M_{\odot}$ (20 particles; see section 2 of Watson et al. 2014b), and the minimum spatial resolution is $0.5 h^{-1}$ cMpc (mesh number of 12000^3 for a simulation box with a side length of $6 h^{-1}$ cGpc; see table 1 of Watson et al. 2013). At $z = 1$, these correspond to a virial mass and radius of $M_{200} = 1.88 \times 10^{12} M_{\odot}$ and $r_{200} = 0.3$ Mpc, respectively. Since we select only those haloes that have $M_{200} \geq 3.5 \times 10^{13} M_{\odot}$, we can be confident that the haloes we use are all well resolved.

The 8.57 cGpc side length of the Jubilee simulation significantly exceeds the co-moving Hubble radius $c/(aH) = 5.04$ cGpc at $a = 0.5$ (c is the speed of light in vacuum), so the simulation is about as large as the entire observable Universe at that epoch. As a result, the Jubilee simulation volume is $\approx 1380\times$ larger than that of the El Gordo discovery survey. Therefore, the severe tension between Λ CDM and the existence of El Gordo is evident already in the disparity between its mass and that of the most massive Jubilee pair with similar dimensionless parameters (Fig. 5). Importantly for the accuracy of our results, they are subject to only a very small amount of cosmic variance due to the large simulation volume – if the root mean square matter density fluctuation between spheres of radius $8 h^{-1}$ Mpc is $\sigma_8 = 0.811 \pm 0.006$ today (Planck Collaboration VI 2020), then for the Jubilee volume at $z = 1$, it should be only 5.4×10^{-4} as the density fluctuations scale inversely with size (Harrison 1970; Zeldovich 1972) and grow $\propto a$. Consequently, the density fluctuations on a mass scale of $10^{15} M_{\odot}$ are accurately handled by a numerical simulation with a side length of just $0.1 h^{-1}$ cGpc (fig. 2 of Watson et al. 2013).

Our own statistical analysis at $z = 1$ (from which we obtain our nominal result for P_{EG}) is based on fitting to 15 035 pairs that pass the $\tilde{v} \geq \tilde{v}_{\text{EG}}$ condition and whose total mass $\tilde{M} \geq 14.16$ to ensure only clusters are considered. The highest mass pair has $\tilde{M} = 14.91$. Due to the large number of objects, the Poisson noise should be very small over the majority of this range, allowing for an accurate analytic fit to the pairwise mass function. This is evident from the fractional uncertainty in the number of pairs with $\tilde{M} \geq 14.16$: Our estimate of $N = 15\,035$ in the simulation volume has a fractional Poisson uncertainty of only

$$\text{Poisson noise} = \frac{1}{\sqrt{N}} = 8.16 \times 10^{-3}. \quad (25)$$

This explains the very good analytic fit in Fig. 5.

The uncertainty increases at higher mass – due to the extreme rarity of objects like El Gordo in Jubilee, it is not possible to simply count the number of analogue pairs with $\tilde{M} > \tilde{M}_{\text{EG}}$. Instead, we

must rely on an extrapolation of our quadratic mass function. Since we identified low mass analogues to El Gordo with \tilde{M} as high as 14.91, we only need to extrapolate the mass function by 0.6 dex, less than the range of \tilde{M} used to define the mass function. Thus, the extrapolated number of analogues with $\tilde{M} > \tilde{M}_{\text{EG}}$ should be quite reliable. The uncertainties would be even lower when $z = 0.509$ and 0 due to an increasing number of haloes above a fixed mass (Fig. 7; see also Appendix A).

The majority of the pairs we find are separated by much more than the sum of their virial radii (Fig. 2). As a result, the halo finder used (AHF) would only very rarely misclassify a pair as one object, incorrectly removing it from the statistics. Different halo finders also yield similar cluster mass functions (fig. 3 of Watson et al. 2014a). Therefore, we can be confident that simulated galaxy cluster pairs similar to the El Gordo progenitors are reliably identified by AHF.

For these reasons, we expect that the pairwise mass function of the Jubilee simulation over the fitted range should be a very accurate representation of Λ CDM. Indeed, Watson et al. (2014a) noted that the Jubilee simulation has proven capable of reproducing the distribution of massive individual clusters according to both observations and theoretical predictions. It has also complied with most of the Λ CDM predictions for the distribution of structures.

Currently, it would be hard to compare if other cosmological simulations give a similar occurrence rate for extreme objects like El Gordo, since cosmological simulations as large as Jubilee are not that common. Nevertheless, there is no reason to believe that the Jubilee simulation might not accurately represent Λ CDM since, up to now, it has been shown to work correctly in accordance with the Λ CDM cosmological model for which it was designed. Moreover, the use of Poisson statistics is justified for rare objects like El Gordo (Watson et al. 2014a).

4.2 A lower mass

Our adopted mass of $3.2 \times 10^{15} M_{\odot}$ exceeds the $(2.16 \pm 0.32) \times 10^{15} M_{\odot}$ estimated in table 2 of Menanteau et al. (2012) by combining various techniques. However, the large uncertainties given there indicate that these earlier estimates are not very reliable. In addition, the methods used to obtain the El Gordo mass did not account for the fact that it is an interacting cluster. The X-ray temperature and luminosity-based estimates of $\approx 2.6 \times 10^{15} M_{\odot}$ should have been multiplied by a factor of 1.17 to account for this (section 4.1.3 of Vikhlinin et al. 2009). In addition, the SZ signal mass estimate of $1.6^{+0.6}_{-0.4} \times 10^{15} M_{\odot}$ is biased low by $\approx 10\text{--}20$ per cent in the case of interacting clusters (fig. 7 of Krause et al. 2012), while the estimate from the cluster’s velocity dispersion should have taken into account that dynamical friction slows down a substructure before it can suffer mass loss due to tidal stripping (Munari et al. 2013). In fact, Menanteau et al. (2012) mention in their section 3.3 that due to considering El Gordo a virialized cluster, their nominal value for its mass is likely to be an underestimate. This is supported by the fact that summing their dynamical masses for its two subcomponents gives $(2.8 \pm 0.9) \times 10^{15} M_{\odot}$, which is closer to the Jee et al. (2014) weak lensing estimate. Clearly, reliable estimates of the El Gordo mass require a hydrodynamical model designed to reproduce the X-ray, weak lensing, and SZ maps in detail. This is precisely what was done by Zhang et al. (2015), as discussed in their section 2.2.

So far, we have focused on their fiducial Model B for reasons discussed in Section 2.1. Fig. 3 shows that the lower mass Model A may reduce the tension with Λ CDM. We therefore discuss whether it provides a better explanation of the observations in a Λ CDM

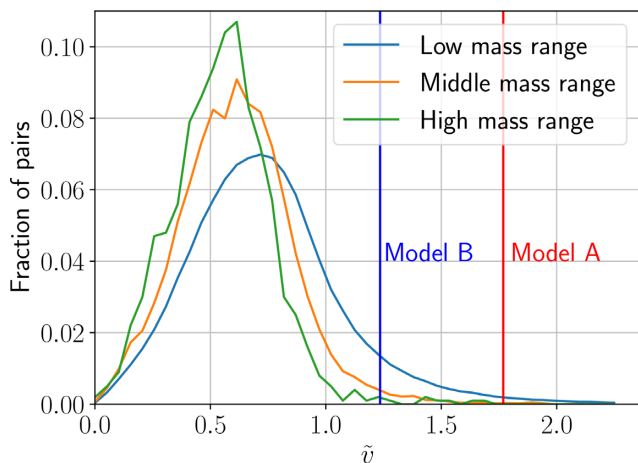


Figure 8. The \tilde{v} distribution for different mass ranges (in units of $10^{15} M_{\odot}$: 1–3.6, 3.6–6.2, and >6.2). The ‘low’ and ‘middle’ mass ranges each have 396350 pairs, while the ‘high’ mass range has only 1000 pairs. All curves are normalized to a sum of 1. The vertical lines show the \tilde{v} of each model. Notice the narrowing of the \tilde{v} distribution with pair total mass.

context. This entails considering the likelihood of the pre-merger configuration and how well Model A reproduces the observed morphology.

Although Model A requires a lower mass for El Gordo ($\tilde{M} = 15.29$), the progenitors should have a more nearly equal mass ratio (≤ 2 rather than ≤ 3.6) and be infalling on to each other ≈ 20 per cent faster (table 2 of Zhang et al. 2015). Thus, Model A requires $\tilde{v} = 1.77$ while model B manages with a lower $\tilde{v} = 1.24$. To explore what effect this could have, we use Fig. 8 to show the distribution of \tilde{v} in three different \tilde{M} bins. It is clear that the higher \tilde{v} will significantly reduce the number of analogues if the Model A parameters are adopted. This effect is especially pronounced at the high mass of El Gordo because the \tilde{v} distribution becomes narrower at high masses. The most likely explanation is that such massive cluster pairs are probably quite isolated – it would be very unlikely to find a third massive object in the vicinity of the pair. Without the third object, it is very difficult for the pair to have $\tilde{v} > 1$ as this entails more kinetic energy than is required to escape. This is why the observed combination of \tilde{M} and \tilde{v} is so problematic for Λ CDM.

Model A thus achieves a different trade-off between how difficult it is for Λ CDM to explain different aspects of the pre-merger configuration. To better understand if it helps overall, we fit the cumulative \tilde{v} distribution $\log_{10} N(\geq \tilde{v})$ using a cubic (Fig. 9). It is evident that the higher \tilde{v} reduces the number of \tilde{v} analogues by ≈ 1 order of magnitude, as also noted in Table 2. This is less significant than the ≈ 2 orders of magnitude increase expected from the use of a lower mass (Fig. 3), suggesting that Model A is more likely from a cosmological perspective.

However, the last column of Table 2 shows that it is still very difficult to simultaneously explain the mass and collision velocity of El Gordo, even with the Model A parameters. The very low number of analogues listed there can be understood as follows: Of the 1.23 analogues in \tilde{M} for Model A, we expect a fraction $0.753/1000$ to pass the $\tilde{v} > \tilde{v}_{\text{EG}}$ condition, implying that 9.26×10^{-4} systems pass both conditions if they are independent. However, the results in Fig. 8 show that the \tilde{v} distribution gets narrower at high \tilde{M} . We implicitly account for this narrowing by only considering pairs with $\tilde{v} > \tilde{v}_{\text{EG}}$ before extrapolating in \tilde{M} . Thus, we expect even fewer pairs with

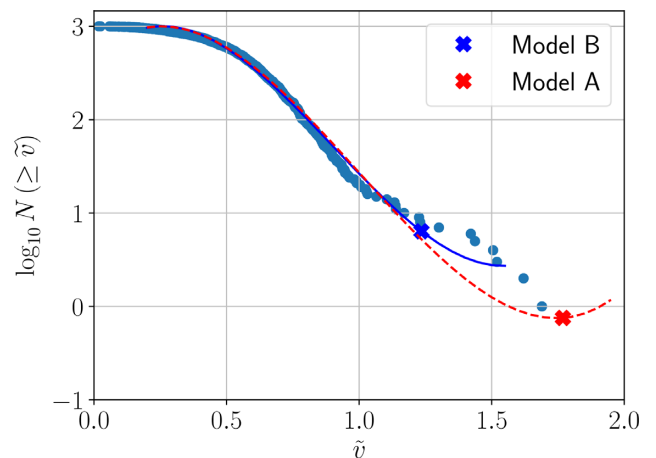


Figure 9. The cumulative \tilde{v} distribution for the 1000 most massive candidate El Gordo analogues. The dotted red (solid blue) line shows a cubic fit for Model A (B). For clarity, individual points are only shown for the more realistic Model B (see the text). The crosses show \tilde{v} for each model. Notice that Model A is more of an outlier here, opposite to the situation in Fig. 3.

Table 2. The number of pairs with larger \tilde{M} and/or \tilde{v} than El Gordo in the Jubilee simulation volume of $(6 h^{-1} \text{ cGpc})^3$, which is much larger than the surveyed region (Section 3.2). The properties of El Gordo are from Models A and B in Zhang et al. (2015). Since the \tilde{v} distribution narrows at high masses, the results for \tilde{v} are based on cubic fits to the cumulative \tilde{v} distribution of the 1000 most massive pairs (Fig. 9). The last column corresponds to imposing the \tilde{v} condition before performing a quadratic extrapolation in \tilde{M} . Since each condition is individually rather problematic, requiring both simultaneously leads to very few analogues (see the text).

| | \tilde{M} | Number of pairs with higher \tilde{v} | Both |
|---------|-----------------------|---|-----------------------|
| Model A | 1.23 | 7.53×10^{-1} | 1.00×10^{-5} |
| Model B | 1.42×10^{-2} | 6.38 | 3.08×10^{-8} |

both a higher \tilde{M} and a higher \tilde{v} than El Gordo. This probably explains why we get only 1.00×10^{-5} analogues instead of 9.26×10^{-4} .

To quantify the overall effect of using Model A instead of Model B, we repeat our light-cone tomography analysis for the Model A parameters. This method suggests less tension with Λ CDM, so we use it to be more conservative. We find that the P -value is 2.73×10^{-7} , representing 5.14σ tension. The $360\times$ enhancement to the Model B probability of 7.51×10^{-10} (Section 3.3) is very much in line with the $320\times$ enhancement suggested by the last column of Table 2. When combined with the Bullet Cluster observations using equation (23), we get an overall tension of 5.50σ . Thus, it is clear that the more likely pre-merger configuration in Model A is still inconsistent with Λ CDM expectations.

So far, we have only discussed whether the Model A pre-merger configuration is likely to arise in Λ CDM. For a holistic discussion of whether Model A is a valid solution in this context, we also need to consider whether it reproduces observations of El Gordo (second step in Fig. 1). The fact that it does not do so nearly as well as Model B explains why Zhang et al. (2015) adopted the latter as their nominal model and optimized it further in Extended Model B. The main problem with Model A is that it does not reproduce the twin-tailed morphology of El Gordo, as can be seen in their fig. 6. This is a very important feature of the real El Gordo (fig. 1 of Menanteau et al. 2012). In addition, the implied mass of $1.95 \times 10^{15} M_{\odot}$ is in

tension with the weak lensing mass of $(3.13 \pm 0.56) \times 10^{15} M_{\odot}$ (Jee et al. 2014). This estimate is for a slightly different definition of the mass: The radius used is such that the enclosed density is $200\times$ the average matter density rather than the critical density. However, the Universe was matter dominated at $z = 1$, so the difference is small. This is evident in fig. 14 of Jee et al. (2014), which shows that the quoted uncertainty is significantly larger.

Returning to the Model B case, we need to consider whether a lower mass version of Model B might plausibly agree with detailed observations, even if a higher mass yields better agreement. Clearly, if Λ CDM is correct, a compromise is needed between the likelihood of the initial conditions of a hydrodynamical merger simulation and the extent to which it matches observations in detail (Fig. 1). We can address this by considering the variations tried by Zhang et al. (2015), exploiting the fact that they ran 123 models. At the end of their section 3.2, they discuss a modification with respect to fiducial Model B involving a reduced main cluster mass of $\tilde{M}_1 = 15.20$ and, because of the smaller size of the system, a lower impact parameter of $b = 600$ kpc. The mass ratio and infall velocity remain the same as in fiducial Model B, so the total mass is $\tilde{M} = 15.31$ and $\tilde{v} = 1.43$. Since this model has an X-ray luminosity significantly smaller than that of fiducial Model B, we can consider $\tilde{M} = 15.31$ a secure lower limit without some other compensatory adjustment. Repeating our light-cone tomography analysis for these ‘low-mass Model B’ parameters, we obtain a 5.17σ tension ($P = 2.35 \times 10^{-7}$).

Zhang et al. (2015) concluded that Model B mergers require a higher mass of $\tilde{M}_1 \approx 15.40$ in order to reproduce an X-ray luminosity consistent with observations. The X-ray luminosity can also be increased by considering a higher velocity ($\approx 3500 \text{ km s}^{-1}$), but this would simply lead to a configuration very similar to Model A – which as mentioned above also does not have plausible initial conditions, and is problematic for other reasons. In any case, the higher resulting \tilde{v} would worsen the tension for Λ CDM. Furthermore, Model A and low-mass Model B require a smaller impact parameter than fiducial Model B. This is geometrically less likely and ought to reduce the probability further by a factor of $\approx(3/8)^2$ in the Model A case and $\approx(6/8)^2$ in the low-mass Model B case. Therefore, a lower mass does not alleviate the tension between Λ CDM and the observed properties of El Gordo. While the pre-merger configuration becomes somewhat more likely in terms of mass, several additional tensions emerge when switching to either model.

Zhang et al. (2015) were already aware that the high mass and collision velocity of their best-fitting model would be problematic for Λ CDM (see their section 4). Given also their thorough exploration of parameter space, we assume that further hydrodynamical Λ CDM simulations with significantly different initial conditions would not reproduce the observed properties of El Gordo. In particular, the tension cannot be reduced below 5σ by lowering the El Gordo mass, either by switching to Model A or while other parameters are fixed to Model B values. This is because substantially different pre-merger configurations fail to reproduce key aspects of the observations.

Recent surveys reveal an increasing number of high- z massive objects that together could pose a problem for Λ CDM (e.g. Kang & Im 2015) in addition to individual objects that by themselves are already on the verge of falsifying it. Thus, several authors considered the Λ CDM model to be the actual problem, and tried to explain the presence of these objects using a different model: Angus & McGaugh (2008) and Katz et al. (2013) in Milgromian dynamics (MOND); Milgrom 1983, and Brownstein & Moffat (2007) in Modified Gravity (MOG; Moffat 2006). MOG is a covariant scalar–tensor–vector gravity theory that allows the gravitational constant G , a vector field coupling ω , and the vector field mass μ to vary with

space and time. This model was recently ruled out at 5.49σ using the velocity dispersion profile of the ultra-diffuse galaxy Dragonfly 44 (Haghi et al. 2019). While perhaps the basic ideas of MOG can still be saved by applying some corrections or modifications, its current formulation seems to be in conflict with observations. Therefore, we do not investigate if MOG could work as a solution to the presence of massive, high- z colliding clusters.

4.3 Milgromian dynamics (MOND)

The MOND model generalizes gravity at low accelerations in such a way that, in its regime of action ($g \ll a_0$), the gravitational field strength g behaves as

$$g = \frac{\sqrt{GMa_0}}{r} \quad (26)$$

at distance r from an isolated point mass M generating the gravitational field. MOND adds a fundamental new acceleration scale a_0 below which the deviation from Newtonian dynamics becomes significant. To match observed galaxy rotation curves, $a_0 \approx 1.2 \times 10^{-10} \text{ m s}^{-2}$ (Begeman, Broeils & Sanders 1991; Gentile, Famaey & de Blok 2011). This generalization of gravity has a direct implication on the measurement of masses both by dynamical methods and by lensing, which works similarly to General Relativity in that both have the same relation between the non-relativistic \mathbf{g} and light deflection (Milgrom 2013, and references therein). The Newtonian dynamical mass M_N is related to the MOND dynamical mass M_M by (Katz et al. 2013)

$$M_M = M_N^2 \times \left(\frac{a_0 r^2}{G} + M_N \right)^{-1}. \quad (27)$$

This was derived by Angus & Diaferio (2011) based on the simple interpolating function (Famaey & Binney 2005). This states that in spherical symmetry, the true gravity \mathbf{g} and the Newtonian gravity \mathbf{g}_N are related by

$$\mathbf{g} = \nu \mathbf{g}_N, \text{ where} \quad (28)$$

$$\nu = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{a_0}{g_N}}. \quad (29)$$

In the quasi-linear formulation of MOND (Milgrom 2010), equation (28) is generalized to more complicated geometries by setting $\nabla \cdot \mathbf{g} = \nabla \cdot (\nu \mathbf{g}_N)$. To recover the correct asymptotic limits, ν should be 1 for $g_N \gg a_0$ and $\sqrt{a_0/g_N}$ for $g_N \ll a_0$. The transition between these limits is a free function in MOND. While several interpolating functions have been developed for this purpose (Kent 1987; McGaugh, Lelli & Schombert 2016), the simple interpolating function seems to work better with recent observations (Iocco, Pato & Bertone 2015; Banik & Zhao 2018; Chae, Bernardi & Sheth 2018). In particular, it provides a good fit to the relation between the radial components of \mathbf{g} and \mathbf{g}_N in rotating galaxies (McGaugh et al. 2016; Lelli et al. 2017). For recent reviews of MOND, we refer the reader to Famaey & McGaugh (2012) and Milgrom (2014).

Equation (27) shows that dynamical masses in MOND should be lower than those inferred from Newtonian dynamics. The extent to which this is true depends on what value we adopt for r . Since the El Gordo clusters are caught close to pericentre, we take M to be their combined mass and r to be the virial radius of the more massive component. Thus, we use $M_N = 3.13 \times 10^{15} M_{\odot}$ ($\tilde{M}_N = 15.50$) and $r = 1.65$ Mpc, with the latter coming from table 2 of Jee et al. (2014). Their fig. 10 shows that the mass is mostly concentrated within a

projected radius of 1 Mpc, so 1.65 Mpc seems like a reasonable choice for r where the assumption of spherical symmetry would approximately hold. Much larger values would not be appropriate as the observations do not go out that far. With these values, we get that the weak lensing mass of El Gordo would become $M_M = 1.8 \times 10^{15} M_\odot$ ($\tilde{M}_M = 15.26$) in MOND, about half the Newtonian value.

MOND was originally developed to explain the flat rotation curves of galaxies without resorting to the presence of dark matter (Milgrom 1983). It also proved useful at explaining other phenomena such as the satellite planes of the Milky Way and Andromeda (Banik, O’Ryan & Zhao 2018; Bílek et al. 2018). However, MOND has faced serious problems as a model that could single-handedly explain all physical phenomena in the Universe. The most well-known problem is that some undetected mass is needed to explain the velocity dispersions of galaxy clusters (Sanders 1999). Although MOND greatly reduces the discrepancy between the observed and dynamically inferred mass compared to the Newtonian case (Ettori et al. 2019), a significant mismatch remains at small radii. The different distributions of dynamical and baryonic mass is most evident in the Bullet Cluster, where the weak lensing and X-ray peaks are offset at high significance (Clowe et al. 2006). Therefore, the additional mass needed by MOND should be collisionless.

It is important to realize that this extra mass is evident on a much larger scale than individual galaxies, so the Bullet Cluster does not require *cold* dark matter (Angus et al. 2007). In MOND, it is of course not possible to add dark matter particles that would significantly cluster on galaxy scales. Instead, some form of hot dark matter (HDM) is required. Assuming that the missing mass is not just baryons in some hard to detect form (Milgrom 2008),⁸ a popular candidate is sterile neutrinos (Angus 2009) since these do not require any significant modifications to the standard model of particle physics (Merle 2017; Boyarsky et al. 2019). Neutrinos are also the only known massive particles that are electrically neutral and long-lived. In the MOND context, an undiscovered species of sterile neutrino would be required with a mass of $m_{\nu_s} = 11 \text{ eV}/c^2$ because this is the mass at which thermally produced sterile neutrinos have the same average mass density as the dark matter in Λ CDM. Sterile neutrinos with $m_{\nu_s} > 10 \text{ eV}/c^2$ have such a short free streaming length as to be consistent with the Planck results (section 6.4.3 of Planck Collaboration XIII 2016). 11 eV/c^2 sterile neutrinos thus provide a plausible explanation for the observed anisotropies in the CMB (Angus 2009). This marriage of MOND + sterile neutrinos (which we term ν HDM; discussed further in section 3.1 of Haslbauer et al. 2020) can also explain the internal dynamics of 30 virialized galaxy clusters, with the implied neutrino density marginally reaching the Tremaine–Gunn limit (Tremaine & Gunn 1979) at the centre once allowance is made for the brightest cluster galaxy (Angus, Famaey & Diaferio 2010). An extra sterile neutrino could plausibly have evaded direct detection with current technology, though there are some hints for it in terrestrial experiments (e.g. MiniBooNE Collaboration 2018).

Structure formation in MOND is expected to be much more efficient than in Λ CDM (Sanders 1998). In Angus & Diaferio (2011) and Angus et al. (2013), the ν HDM model was explored with cosmological N -body simulations that assumed a standard expansion history (Skordis et al. 2006) and applied MOND only to the density

perturbations (Sanders 2001; Nusser 2002). The main conclusion of Angus et al. (2013) was that massive galaxy clusters are overproduced in ν HDM. However, this can be attributed to several factors:

(i) The resolution of their $(256 h^{-1} \text{ cMpc})^3$ simulation was very low, so it could barely account for the presence of smaller, less massive structures. This makes it less likely to end up having small structures in the simulation, and more likely to have a few very massive objects.

(ii) A small box like the one they used cannot account for the external field from distant background objects. The external field effect is a physical consequence of the non-linear gravity law in MOND (Milgrom 1986). It implies that the internal gravity of a system is weakened by a constant gravitational field from the external environment, even in the absence of tides. In this case, the presence of large background accelerations would have made the gravitational potential of the system more Newtonian, possibly suppressing the overproduction of massive structures.

(iii) The mass function produced by the simulation was compared to the cluster mass functions of Reiprich & Böhringer (2002) and Rines, Diaferio & Natarajan (2008), which were developed for galaxy clusters that mostly lie at $z \lesssim 0.1$. According to Keenan et al. (2013), our Galaxy lies inside an ≈ 300 Mpc void with a density contrast of ≈ -0.5 . Galaxy clusters at low z would be inside this void (hereafter the KBC void). Therefore, the $z \lesssim 0.1$ cluster mass function is very likely not representative of the Universe as a whole. This is also apparent in that El Gordo-like objects should be far more common at $z = 0$ than at $z = 1$ regardless of the cosmological model. Thus, it is surprising to not see similarly extreme objects at low z (except perhaps the Bullet Cluster at $z = 0.296$). This is most likely due to the KBC void out to $z \approx 0.15$ (Keenan et al. 2013), which is also apparent in X-ray surveys of galaxy clusters (Böhringer et al. 2015; Böhringer, Chon & Collins 2020). We point out that the Angus & Diaferio (2011) simulations already predicted the presence of large voids with $250 h^{-1}$ Mpc diameter, but attributed this to a flaw in their model. There is strong evidence that we are living inside such a supervoid (i.e. the KBC void; see section 1 of Haslbauer et al. 2020, and references therein). By modelling its detailed dynamics in the ν HDM framework, they showed that it can explain the Hubble tension and the curvature of the low- z distance–redshift relation, which are otherwise difficult to understand in Λ CDM.

Katz et al. (2013) explored the ν HDM model using N -body simulations in a box whose linear dimensions were twice as large as those of Angus et al. (2013). The conclusion of Katz et al. (2013) was that both the velocities and masses of massive clusters are larger at late times in MOND compared to Λ CDM. Clusters become more massive at an earlier epoch in MOND, which would help to explain the large number of massive high- z clusters that have been found in the last few years (e.g. Foley et al. 2011). As with the earlier ν HDM simulations, Katz et al. (2013) also reported an overabundance of high-mass clusters. However, they again used the Reiprich & Böhringer (2002) observational data for comparison. This catalogue reaches clusters up to $z \approx 0.3$ but is mostly composed of lower redshift clusters, whose abundance would be significantly affected by the KBC void.

Katz et al. (2013) found that among their simulated cluster pairs with a similar redshift, mass, and mass ratio to the Bullet Cluster, 13 per cent reach its 3000 km s^{-1} infall velocity in MOND, but only 2 per cent do so in their larger Λ CDM simulation volume (see their fig. 8). Remarkably, they managed to identify a few pairs in their MOND simulation that match the redshift, collision velocity, and main cluster mass of El Gordo (see their section 5.1). Their adopted

⁸This would involve only a small fraction of the baryons expected from big bang nucleosynthesis.

pre-merger configuration had an infall velocity of 2300 km s^{-1} and a main cluster mass of $2 \times 10^{15} M_{\odot}$, which as discussed above are reasonable in a MOND context. Pairs like El Gordo are quite rare at its high redshift in ν HDM – but they do occur. No such pairs were found in their Λ CDM simulation. Therefore, the presence of El Gordo-like objects is not as unusual in the ν HDM model compared to Λ CDM.

To quantify whether the occurrence rate of El Gordo analogues in ν HDM is similar to observations, we perform a simple calculation of how many El Gordo-like objects are expected in the surveyed region. For this, we assume that one pair analogous to El Gordo was discovered by Katz et al. (2013) in their $(512 h^{-1} \text{ cMpc})^3$ simulation volume. Given the effective survey volume calculated with equations (19) and (20), we obtain that $N_{\text{EG}} = 1.16$. This assumes that the growth parameter k in equation (15) ($k = 32.75$ for a fit with $\tilde{M}_{\text{min}} = 14.16$ after requiring $\tilde{v} > \tilde{v}_{\text{EG}}$) is the same in MOND as in Λ CDM, which is not necessarily true. While a more accurate calculation goes beyond the scope of this contribution, we can assume that k should not change by more than order unity (i.e. by $\lesssim 1$ dex) regardless of whether we are using MOND or Λ CDM. Since $N_{\text{EG}} \approx 1/k$ for $k \gg 1$ (equation 19), order unity changes to k should not significantly affect N_{EG} . It would also not make much difference if instead of one pair analogous to El Gordo, Katz et al. (2013) had found, e.g. three pairs – they certainly did not find very many, but they did find ≥ 1 (see their section 5.1). Thus, *the ν HDM framework yields the correct order of magnitude for the number of El Gordo analogues in the surveyed region.*

Part of the reason for this is that the MOND dynamical mass is about half the Newtonian value. As discussed in Section 4.2, a lower mass helps to alleviate the tension with the simulated cluster mass function. If we repeat the redshift tomography analysis using instead M_{M} while still keeping the infall velocity and impact parameter of fiducial Model B, we obtain that the expected number of analogues rises from $N_{\text{EG}} = 7.51 \times 10^{-10}$ to 3.78×10^{-6} , reducing the tension from 6.16σ to 4.62σ . While this manages to get the tension below 5σ , it rises back to 5.04σ when combined with the Bullet Cluster. Since we mentioned above that $N_{\text{EG}} \approx 1$ in MOND, the lower mass is not the main reason why it could explain the existence of El Gordo. Rather, the MOND model manages to raise the number of El Gordo analogues mainly by enhancing structure growth.

The extent to which this occurs depends on how gravitational fields from inhomogeneities couple to g_{Hubble} , the acceleration required to maintain the time-dependent Hubble flow. We discuss this theoretical uncertainty below, and refer the reader to section 5.2.3 of Haslbauer et al. (2020) for a more detailed discussion. The ν HDM simulations of Katz et al. (2013) applied MOND only to the gravity sourced by inhomogeneities, thus assuming no coupling to g_{Hubble} . The possibility of a non-trivial coupling was discussed in Sanders (2001), but has rarely been considered since. If there is such a coupling, the resulting Hubble field effect (HFE) could raise the appropriate value for the gravity in equation (28), thereby suppressing the MOND boost to gravity. In other words, a large background acceleration could make the gravitational potential of the system more Newtonian, which can dampen the production of massive structures. Therefore, the Katz et al. (2013) simulation might have overestimated the number of very massive objects that we should find in a ν HDM universe. The assumption of no HFE may have affected the Katz et al. (2013) results only if:

- (i) the gravitational field g of structures that were already included in the simulation is smaller than a_0 , and
- (ii) g_{Hubble} dominates over g .

The first condition is required so there is a MOND enhancement to gravity at all – if not, then the HFE cannot further suppress the already non-existent MOND boost to gravity. The second condition is required for the HFE to be significant. To make further progress, we note that the g_{Hubble} term for a system with size r is defined as

$$g_{\text{Hubble}} \equiv \frac{\ddot{a}}{a} r = -\frac{4\pi G}{3} (\rho_{\text{m}} - 2\rho_{\Lambda}) r \quad (30)$$

$$= H_0^2 \left(-\frac{1}{2} a^{-3} \Omega_{\text{m},0} + \Omega_{\Lambda,0} \right) r_c a, \quad (31)$$

where we use the canonical $a_0 = 1.2 \times 10^{-10} \text{ m s}^{-2}$ (Begeman et al. 1991; Gentile et al. 2011) and the same values as the Jubilee simulation for the present cosmological parameters H_0 , $\Omega_{\text{m},0}$, and $\Omega_{\Lambda,0}$. The co-moving radius $r_c = 22.70 \text{ cMpc}$ is the radius of a sphere enclosing the mass of El Gordo at the cosmic mean density. This is redshift independent due to mass conservation. r_c is thus a typical co-moving length-scale for the problem, while r is the corresponding physical scale.

To estimate g sourced by inhomogeneities, we need to make a few assumptions. We take that perturbations grow $\propto a$, as occurs in Λ CDM during the matter-dominated era. This is conservative as it raises the relative importance of the HFE by suppressing g . With this assumption, the Newtonian gravity of the inhomogeneities is $g_{\text{N}} \propto a^{-1}$. In order to obtain the actual value of g_{N} , we need to know g_{N} at some epoch, which we take to be the epoch of recombination ($a_{\text{CMB}} \approx 1/1100 = 9.1 \times 10^{-4}$). We define that $g_{\text{N}} = g_{\text{CMB}}$ at that epoch, so in general,

$$g_{\text{N}} = g_{\text{CMB}} \times \frac{a_{\text{CMB}}}{a}. \quad (32)$$

Following section 3.1.3 of Haslbauer et al. (2020), we estimate the density fluctuations at recombination to be $\delta_{\text{CMB}} \approx 10^{-4}$, implying that the typical gravitational field is

$$g_{\text{CMB}} \approx \frac{c \delta_{\text{CMB}}}{t_{\text{CMB}}} \approx 21 a_0, \quad (33)$$

where $t_{\text{CMB}} = 380 \text{ kyr}$ is the time of recombination. We then assume that since El Gordo is a rare object even in the MOND context, it corresponds to at least a 2σ density fluctuation. The gravitational field at recombination would thus be $g_{\text{CMB}} \approx 42 a_0$. The MOND boost to this is negligible, but at later times the enhancement to g_{N} needs to be calculated using equation (28).

In the matter-dominated era, $g_{\text{Hubble}} \propto a^{-2}$ (equation 31) while $g_{\text{N}} \propto a^{-1}$. As a result, g_{Hubble} becomes sub-dominant to g after some epoch $a = a_{\text{HFE}}$. Meanwhile, the HFE plays no role if $g \gtrsim a_0$ because the behaviour is Newtonian regardless of the HFE. Thus, the HFE can also be neglected prior to the epoch when $g = a_0$, which we define as occurring when $a = a_{\text{MOND}}$. Therefore, the Katz et al. (2013) simulation might overestimate the MOND enhancement to gravity only in the period when $a_{\text{MOND}} \lesssim a \lesssim a_{\text{HFE}}$. Using equations (31) and (32) with the interpolating function in equation (29), we obtain that $a_{\text{HFE}} = 0.06$ while $a_{\text{MOND}} = 0.08$. Hence, there is never any era during which there is a significant enhancement to Newtonian gravity that might be overestimated by not considering a possible HFE – either the HFE would be a sub-dominant correction or the behaviour is Newtonian in any case, so the HFE could not make the growth of structure even more Newtonian. This means that theoretical uncertainties regarding the HFE are very small at the 22.7 cMpc scale relevant to the formation of El Gordo analogues in the ν HDM framework. As a result, their estimated frequency in the Katz et al. (2013) simulation should be a good representation of their frequency in this framework.

From the Katz et al. (2013) results, it appears that the ν HDM model, while still in need of improvement, could serve as a possible explanation for massive high-velocity galaxy cluster collisions at high redshift. In particular, we see that Angus et al. (2013) were not justified in their criticism that MOND does not get the correct cluster mass function at the high-mass end. Leaving aside other issues like resolution, if massive clusters were overproduced, then massive cluster pairs would certainly be overproduced. However, our calculations show that Katz et al. (2013) obtained approximately the correct number of El Gordo analogues in their simulation volume. This conclusion should remain valid despite uncertainty in how the Hubble acceleration couples to that sourced by inhomogeneities in a MOND universe. Therefore, it is important to revisit the ν HDM model.

5 CONCLUSIONS

We studied the probability of observing an El Gordo-like collision between two massive high- z galaxy clusters by searching for progenitor galaxy cluster pairs that have turned around from the cosmic expansion in the $(6h^{-1}\text{cGpc})^3$ Jubilee simulation box at $z = 1$. We used two different statistical analyses based on the number of analogue pairs to the El Gordo progenitor clusters with appropriate total mass, mass ratio, and pre-merger infall velocity. The initial conditions were obtained from the hydrodynamical simulations of Zhang et al. (2015), from which we adopt their nominal Model B. Our main result is that the discovery of an El Gordo-like collision within the surveyed region excludes the Λ CDM model at 6.16σ using the method that gives less tension (light-cone tomography). We also used the analysis of Kraljic & Sarkar (2015) to show that the Bullet Cluster is in 2.78σ tension with Λ CDM. In this framework, the combination of these observations is a highly unlikely 6.43σ event.

We considered whether El Gordo could have a lower mass (Section 4.2). However, this only reduces the tension to 5.14σ (5.50σ when combined with the Bullet Cluster). In reality, the tension should be even higher as it does not account for the following:

- (i) the poor fit to the observed twin-tailed morphology of El Gordo,
- (ii) the discrepancy between the assumed mass and the weak lensing observations (Jee et al. 2014), and
- (iii) the small impact parameter required in this scenario, which may be unlikely on geometrical grounds.

We also consider the possibility that El Gordo and the Bullet Cluster are extremely unique objects such that no other problematic objects like these will be found in the remaining observable sky. Increasing the surveyed area to the full sky reduces the combined tension of El Gordo and the Bullet Cluster to 5.66σ , which does not contribute that much to solve the problem. Besides, the assumption is not very realistic – full sky surveys like Planck have already found interacting clusters whose properties are unlikely to arise in a Λ CDM context (Section 3.4). Further work is required to quantify the extent to which the cases discussed there and in Section 1 are problematic for Λ CDM.

Explaining the Bullet Cluster, El Gordo, and a growing number of fast, massive, high-redshift galaxy clusters can be very difficult within the Λ CDM cosmology. An alternative model uses MOND gravity supplemented by $11\text{ eV}/c^2$ sterile neutrinos to explain various phenomena including the CMB anisotropies, cluster-scale problems for purely baryonic MOND, and the baryon-weak lensing offset in the Bullet Cluster (Section 4.3; see also section 3.1 of Haslbauer et al. 2020). This ν HDM model was previously explored in Katz et al. (2013), where a handful of objects analogous to El Gordo

were found in their simulation volume of $(512h^{-1}\text{cMpc})^3$ (see their section 5.1). This is very similar to the effective volume of the survey that discovered El Gordo. We estimate that the Katz et al. (2013) simulation implies 1.16 El Gordo analogues in the surveyed volume, implying good agreement with observations. Although this result could be off by a factor of a few, it is clear that the ν HDM model produces approximately the correct abundance of massive cluster pairs similar to the El Gordo progenitors. Thus, we argue against the conclusion of Angus et al. (2013) that MOND overproduces massive galaxy clusters. Section 4.3 discusses possible reasons for their erroneous conclusion.

We conclude that the El Gordo galaxy cluster collision rules out the Λ CDM cosmology at high significance, but can likely be explained in a MOND cosmology supplemented by $11\text{ eV}/c^2$ sterile neutrinos. While it is difficult to dispute the evidence for dark matter in galaxy clusters like the Bullet, it is also difficult to explain its properties and those of El Gordo without modifying gravity on large scales – or at low accelerations. Given also the galaxy-scale challenges for Λ CDM (e.g. Kroupa 2012, 2015) and the many successes of MOND on this scale (e.g. Famaey & McGaugh 2012; Li et al. 2018), the most likely scenario is that at present we have understood neither the full matter–energy content of the Universe nor the law of gravity that governs it.

ACKNOWLEDGEMENTS

EA is supported by a stipend from the Stellar Populations and Dynamics Research Group at the University of Bonn. IB is supported by an Alexander von Humboldt Foundation postdoctoral research fellowship. The authors would like to thank the Jubilee team for providing the data needed to carry out this project. They also thank Moti Milgrom for helpful discussions on MOND in galaxy clusters. The authors are grateful to the referee for her/his constructive comments.

DATA AVAILABILITY

The data underlying this article are available in the article.

REFERENCES

- Abazajian K. N., Horiuchi S., Kaplinghat M., Keeley R. E., Macias O., 2020, *Phys. Rev. D*, 102, 043012
- Aiola S. et al., 2020, preprint ([arXiv:2007.07288](https://arxiv.org/abs/2007.07288))
- Alcock C. et al., 2000, *ApJ*, 542, 281
- Angus G. W., 2009, *MNRAS*, 394, 527
- Angus G. W., Diaferio A., 2011, *MNRAS*, 417, 941
- Angus G. W., Diaferio A., Famaey B., van der Heyden K. J., 2013, *MNRAS*, 436, 202
- Angus G. W., Famaey B., Diaferio A., 2010, *MNRAS*, 402, 395
- Angus G. W., McGaugh S. S., 2008, *MNRAS*, 383, 417
- Angus G. W., Shan H. Y., Zhao H. S., Famaey B., 2007, *ApJ*, 654, L13
- Babcock H. W., 1939, *Lick Obs. Bull.*, 498, 41
- Bagchi J. et al., 2011, *ApJ*, 736, L8
- Banik I., O’Ryan D., Zhao H., 2018, *MNRAS*, 477, 4768
- Banik I., Zhao H., 2018, *MNRAS*, 480, 2660
- Bauer A., Springel V., 2012, *MNRAS*, 423, 2558
- Begeman K. G., Broeils A. H., Sanders R. H., 1991, *MNRAS*, 249, 523
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
- Bosma A., 1978, PhD thesis, Groningen Univ.
- Boyersky A., Drewes M., Lasserre T., Mertens S., Ruchayskiy O., 2019, *Prog. Part. Nucl. Phys.*, 104, 1
- Brownstein J. R., Moffat J. W., 2007, *MNRAS*, 382, 29

- Buddendiek A. et al., 2015, *MNRAS*, 450, 4248
- Burkert A., 1995, *ApJ*, 447, L25
- Břek M., Thies I., Kroupa P., Famaey B., 2018, *A&A*, 614, A59
- Böhringer H., Chon G., Bristow M., Collins C. A., 2015, *A&A*, 574, A26
- Böhringer H., Chon G., Collins C. A., 2020, *A&A*, 633, A19
- Cavaliere A., Fusco-Femiano R., 1978, *A&A*, 70, 677
- Chae K.-H., Bernardi M., Sheth R. K., 2018, *ApJ*, 860, 81
- Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, *ApJ*, 648, L109
- Cucciati O. et al., 2018, *A&A*, 619, A49
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Dolag K., Stasyszyn F., 2009, *MNRAS*, 398, 1678
- Donnert J. M. F., 2014, *MNRAS*, 438, 1971
- Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, *MNRAS*, 390, L64
- Ebeling H., Barrett E., Donovan D., Ma C. J., Edge A. C., van Speybroeck L., 2007, *ApJ*, 661, L33
- Ettori S. et al., 2019, *A&A*, 621, A39
- Famaey B., Binney J., 2005, *MNRAS*, 363, 603
- Famaey B., McGaugh S. S., 2012, *Living Rev. Relativ.*, 15, 10
- Finner K. et al., 2017, *ApJ*, 851, 46
- Foley R. J. et al., 2011, *ApJ*, 731, 86
- Forrest B. et al., 2020, *ApJ*, 890, L1
- Fryxell B. et al., 2000, *ApJS*, 131, 273
- Gentile G., Famaey B., de Blok W. J. G., 2011, *A&A*, 527, A76
- Giacconi R. et al., 1979, *ApJ*, 230, 540
- Gill S.P.D., Knebe A., Gibson B.K., 2004, *MNRAS*, 351, 399
- Girelli G., Bolzonella M., Cimatti A., 2019, *A&A*, 632, A80
- Haghi H., Amiri V., Hasani Zonoozi A., Banik I., Kroupa P., Haslbauer M., 2019, *ApJ*, 884, L25
- Harnois-Déraps J., Pen U.-L., Iliev I. T., Merz H., Emberson J. D., Desjacques V., 2013, *MNRAS*, 436, 540
- Harrison E. R., 1970, *Phys. Rev. D*, 1, 2726
- Haslbauer M., Banik I., Kroupa P., 2020, *MNRAS*, 499, 2845
- Hayashi E., White S. D. M., 2006, *MNRAS*, 370, L38
- Hayashi M. et al., 2019, *PASJ*, 71, 112
- Hernquist L., 1990, *ApJ*, 356, 359
- Hockney R. W., Eastwood J. W., 1988, *Computer Simulation Using Particles*. CRC Press, New York, USA and Oxford, UK
- Hoof S., Geringer-Sameth A., Trotta R., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 012
- Iocco F., Pato M., Bertone G., 2015, *Phys. Rev. D*, 92, 084046
- Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, *ApJ*, 785, 20
- Jee M. J. et al., 2009, *ApJ*, 704, 672
- Kang E., Im M., 2015, *J. Korean Astron. Soc.*, 48, 21
- Katz H., McGaugh S., Teuben P., Angus G. W., 2013, *ApJ*, 772, 10
- Keenan R. C., Barger A. J., Cowie L. L., 2013, *ApJ*, 775, 62
- Kent S. M., 1987, *AJ*, 93, 816
- Kim J., Jee M. J., Perlmutter S., Hayden B., Rubin D., Huang X., Aldering G., Ko J., 2019, *ApJ*, 887, 76
- Knebe A. et al., 2011, *MNRAS*, 415, 2293
- Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608
- Kraljic D., Sarkar S., 2015, *J. Cosmol. Astropart. Phys.*, 2015, 050
- Krause E., Pierpaoli E., Dolag K., Borgani S., 2012, *MNRAS*, 419, 1766
- Kroupa P., 2012, *PASA*, 29, 395
- Kroupa P., 2015, *Can. J. Phys.*, 93, 169
- Lage C., Farrar G., 2014, *ApJ*, 787, 144
- Lamer G., Hoefl M., Kohnert J., Schwöpe A., Storm J., 2008, *A&A*, 487, 33
- Lee J., Komatsu E., 2010, *ApJ*, 718, 60
- Lelli F., McGaugh S. S., Schombert J. M., Pawłowski M. S., 2017, *ApJ*, 836, 152
- Lim S., Lee J., 2014, *ApJ*, 783, 39
- Li P., Lelli F., McGaugh S., Schombert J., 2018, *A&A*, 615, A3
- Markevitch M., Gonzalez A. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, *ApJ*, 606, 819
- McGaugh S. S., Lelli F., Schombert J. M., 2016, *Phys. Rev. Lett.*, 117, 201101
- Menanteau F. et al., 2010, *ApJ*, 723, 1523
- Menanteau F. et al., 2012, *ApJ*, 748, 7
- Merle A., 2017, *Sterile Neutrino Dark Matter*. Morgan Claypool Publishers, Bristol, UK
- Milgrom M., 1983, *ApJ*, 270, 365
- Milgrom M., 1986, *ApJ*, 302, 617
- Milgrom M., 2008, *New Astron. Rev.*, 51, 906
- Milgrom M., 2010, *MNRAS*, 403, 886
- Milgrom M., 2013, *Phys. Rev. Lett.*, 111, 041105
- Milgrom M., 2014, *Scholarpedia*, 9, 31410
- Miller T. B. et al., 2018, *Nature*, 556, 469
- MiniBooNE Collaboration, 2018, *Phys. Rev. Lett.*, 121, 221801
- Moffat J. W., 2006, *J. Cosmol. Astropart. Phys.*, 2006, 004
- Molnar S. M., Broadhurst T., 2015, *ApJ*, 800, 37
- Molnar S. M., Hearn N. C., Stadel J. G., 2012, *ApJ*, 748, 45
- Mortonson M. J., Hu W., Huterer D., 2011, *Phys. Rev. D*, 83, 023015
- Mullis C. R., Rosati P., Lamer G., Böhringer H., Schwöpe A., Schuecker P., Fassbender R., 2005, *ApJ*, 623, 85
- Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, *MNRAS*, 430, 2638
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Ng K. Y., Dawson W. A., Wittman D., Jee M. J., Hughes J. P., Menanteau F., Sifón C., 2015, *MNRAS*, 453, 1531
- Nusser A., 2002, *MNRAS*, 331, 909
- Ostriker J. P., Peebles P. J. E., 1973, *ApJ*, 186, 467
- Ostriker J. P., Steinhardt P. J., 1995, *Nature*, 377, 600
- Planck Collaboration VI, 2013, *A&A*, 550, A132
- Planck Collaboration VI, 2020, *A&A*, 641, A6
- Planck Collaboration VIII, 2011, *A&A*, 536, A8
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Planck Collaboration XXVII, 2014, *A&A*, 571, A27
- Reiprich T. H., Böhringer H., 2002, *ApJ*, 567, 716
- Riess A. G., 2020, *Nat. Rev. Phys.*, 2, 10
- Riess A. G. et al., 1998, *AJ*, 116, 1009
- Rines K., Diaferio A., Natarajan P., 2008, *ApJ*, 679, L1
- Rogstad D. H., Shostak G. S., 1972, *ApJ*, 176, 315
- Rubin V. C., Ford W. K. J., 1970, *ApJ*, 159, 379
- Sahlén M., Zubeldía I., Silk J., 2016, *ApJ*, 820, L7
- Sanders R. H., 1998, *MNRAS*, 296, 1009
- Sanders R. H., 1999, *ApJ*, 512, L23
- Sanders R. H., 2001, *ApJ*, 560, 1
- Skordis C., Mota D. F., Ferreira P. G., Boehm C., 2006, *Phys. Rev. Lett.*, 96, 011301
- Springel V., 2005, *MNRAS*, 364, 1105
- Stalder B. et al., 2013, *ApJ*, 763, 93
- Stanford S. A. et al., 2012, *ApJ*, 753, 164
- Sunyaev R. A., Zeldovich Y. B., 1970, *Ap&SS*, 7, 3
- Thompson R., Nagamine K., 2012, *MNRAS*, 419, 3560
- Tisserand P. et al., 2007, *A&A*, 469, 387
- Tozzi P. et al., 2015, *ApJ*, 799, 93
- Tremaine S., Gunn J. E., 1979, *Phys. Rev. Lett.*, 42, 407
- Tucker W. et al., 1998, *ApJ*, 496, L5
- Tucker W. H., Tananbaum H., Remillard R. A., 1995, *ApJ*, 444, 532
- Vikhlinin A. et al., 2009, *ApJ*, 692, 1033
- Wang T. et al., 2016, *ApJ*, 828, 56
- Watson W. A., Iliev I. T., Diego J. M., Gottlöber S., Knebe A., Martínez-González E., Yepes G., 2014a, *MNRAS*, 437, 3776
- Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, *MNRAS*, 433, 1230
- Watson W. A. et al., 2014b, *MNRAS*, 438, 412
- Wen Z. L., Han J. L., 2018, *MNRAS*, 481, 4158
- Yang J. et al., 2020, *ApJ*, 897, L14
- Zeldovich Y. B., 1972, *MNRAS*, 160, 1
- Zhang C., Yu Q., Lu Y., 2015, *ApJ*, 813, 129
- Zitrin A., Menanteau F., Hughes J. P., Coe D., Barrientos L. F., Infante L., Mandelbaum R., 2013, *ApJ*, 770, L15

APPENDIX A: MASS FUNCTIONS FOR EL GORDO-LIKE CLUSTER PAIRS

The coefficients of our quadratic fit to the cumulative mass function (equation 14) are given in Table A1 based on the El Gordo parameters in Model B of Zhang et al. (2015). The corresponding results for their

Table A1. Coefficients of our quadratic fit to the cumulative mass distribution (equation 14) for pairs that satisfy the turnaround, mass ratio, and \tilde{v} conditions appropriate to Model B of Zhang et al. (2015). This is the nominal case considered in our work.

| Mass function | 0 | Redshift 0.509 | 1 |
|---------------|----------|-------------------|----------|
| c_0 | − 369.69 | − 544.03 | − 786.14 |
| c_1 | 55.06 | 80.05 | 114.75 |
| c_2 | − 2.02 | − 2.92 | − 4.16 |

Table A2. Similar to Table A1, but for Model A of Zhang et al. (2015).

| Mass function | 0 | Redshift 0.509 | 1 |
|---------------|----------|-------------------|----------|
| c_0 | − 160.13 | − 203.90 | − 347.02 |
| c_1 | 26.45 | 33.46 | 54.41 |
| c_2 | − 1.05 | − 1.33 | − 2.10 |

Model A are given in Table A2. As described in the text, both fits are done only for pairs with $\tilde{M} > 14.16$ to avoid edge effects at low masses. Since the Jubilee simulation lacks pairs as massive as El Gordo, our statistical analysis is based on extrapolating these fits to the El Gordo mass.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.