

The dynamics and merging scenario of the galaxy cluster ACT-CL J0102-4915, El Gordo

K. Y. Ng¹, W. A. Dawson², D. Wittman¹, J. Jee¹, J. Hughes³, F. Menanteau³, C. Sifón⁴
(temporary order)

¹Department of Physics, University of California Davis, One Shields Avenue, Davis, CA 95616, USA

²Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94551-0808, USA

³Department of Physics & Astronomy, Rutgers University, 136 Frelinghuysen Rd., Piscataway, NJ 08854, USA

⁴Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, Netherlands

arXiv 666

ABSTRACT

Merging galaxy clusters with radio relics provide rare insights to the merger dynamics as the relics are created by the violent merger process.

From the double radio relic observation and X-ray wake morphology, it is believed that El Gordo is observed shortly after the first passage before reaching apo

We demonstrate one of the first uses of the properties of the radio relic to reduce the uncertainties of the dynamical variables and 3D configurations of a cluster merger, ACT-CL J0102-4915, El Gordo. At a redshift of 0.87, El Gordo ($M_{200c} = 2.75 \times 10^{15} \pm 7.4 M_\odot$) is one of the most massive clusters discovered in the early universe. The two subclusters of El Gordo has a mass ratio of around 2:1. The X-ray and weak-lensing data of El Gordo show an offset of X kpc between the intercluster gas and the dark matter (DM) at $\sim 4\sigma$ level. All these features of El Gordo make it part of a valuable class of dissociative mergers that can probe the self-interaction of dark matter. We employ a Monte Carlo simulation to investigate the three-dimensional (3D) configuration and dynamics of El Gordo. We give a summary of the inferred dynamical variables. By making use the polarization, velocity and position of the radio relic, we are able to confirm at $X\sigma$ that the subclusters of El Gordo are moving away from each other. We find that the 3D merger speed of El Gordo to be $\sim 3000 \text{ km s}^{-1}$ (or in projected velocity =), which is still consistent with the low line-of-sight velocity of $\sim 600 \text{ km s}^{-1}$ based on the inferred time-since-collision ($TSC = \text{Gyrs}$) and the projection angle ($\alpha = 41^\circ \pm$). We put our estimates of TSC and α into context by relating them to existing observations of El Gordo. Finally, we compare our simulation result of El Gordo to the simulation result of the Bullet Cluster, and show that El Gordo is a very promising candidate for giving tighter constraint than the Bullet Cluster on the self-interaction of dark matter. (200 words) (check against astro-ph word limit)

Key words: gravitational lensing – dark matter – cosmology: observations – X-rays: galaxies: clusters – galaxies: clusters: individual (ACT-CL J0102-4915) – galaxies: high redshift

1 INTRODUCTION

Mergers of dark-matter-dominated galaxy clusters probes properties of the cluster components like no other systems. Clusters of galaxies are made up of 80% of dark matter in mass content, with a smaller portion of intercluster gas ($\sim 15\%$ in mass content), and sparsely spaced galaxies ($\sim 2\%$ in mass content) (REF). During a merger of clusters, the subclusters are accelerated to high speeds of several thousand km s^{-1} . The offsets of different components of

the subclusters dissociate show how various interactions of the different components are at work. Observables such as offset between dark matter and the other components may suggest dark matter self-interaction (REF). (The following sentence does not actually fit in this paragraph and I have to put it somewhere else) difference of the galaxy colors in a merging cluster from relaxed cluster can also verify effects of environment on galaxy evolution.

Ever since the discovery of El Gordo in the Atacama Camera Telescope (ACT) survey (REF), there is an ongoing

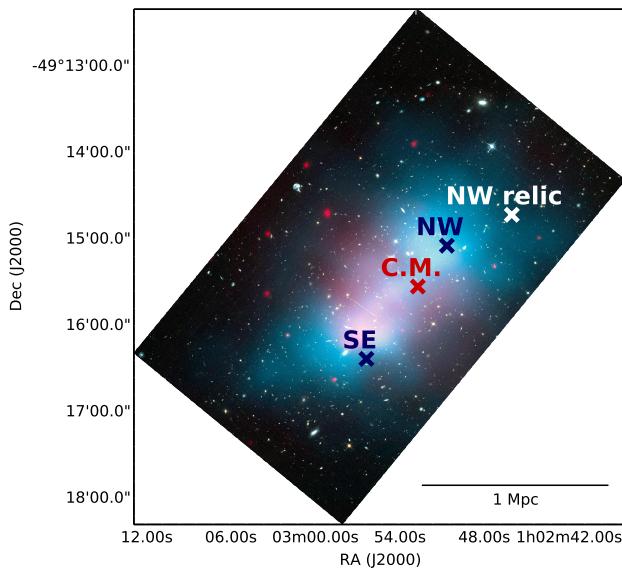


Figure 1. Configuration of El Gordo showing approximate overlaid of dark matter distribution in blue, and X-ray luminosity in red. The cross markers show the position of the northwest (NW) and southeast(SE) dark matter density peaks, and the center of mass (CM) location respectively.

effort for collecting comprehensive data for El Gordo. The presence of the radio relic, in conjunction with a depression in the X-ray surface brightness shown in M12, strongly suggest that El Gordo is a post-collision system so we limit our discussion to inferring the time-since-collision.

From the spectroscopy and Dressler-Schechter test for the member galaxies in Sifón et al. (2013), El Gordo is confirmed to be a binary merger without significant substructures. This picture is further supported by the weak lensing analysis by Jee et al. (2013). The weak lensing analysis shows a mass ratio of $\sim 2:1$ between the two main subclusters, named according to their location as the northeast (NW) and southeast (SE) subclusters respectively. (See Figure 1). El Gordo has interesting intracluster medium morphology as shown in the X-ray. In the northwest, it shows a wake feature, i.e., depression in the X-ray emissivity, while in the southeast, it shows highest X-ray emissivity indicative of a cold gas core southeast of the wake. The cold gas core may have passed from the northwest to the southwest to have caused this morphology (Menanteau et al. 2011, hereafter M12). The extended mass distribution of El Gordo also makes it a good gravitational lens. Zitrin et al. (2013) have found multiple strong gravitationally lensed images around the center region of El Gordo. On the outer skirt of El Gordo, strong radio emission is detected in the NW and the SE respectively. These radio emission has steep spectral index gradient and are identified as radio relic created from a merger.

El Gordo is one of small sample of galaxy clusters (~ 50) that have been associated with a radio relic. (This paragraph needs a lot more organization) Even fewer of them have been studied in great details, making El Gordo a valuable candidate for further analysis. Furthermore, El Gordo satisfies the four criteria for being a dissociative merger which are proposed to be excellent probes of self-interacting dark matter (Dawson et al. 2012) . (1) The subclusters of El Gordo has

a small ratio of mass, i.e. $\sim 2 : 1$ (Jee et al. 2013, hereafter J13). (2) The merger axis, the line joining the two subclusters, coincides with the alignment of the double radio relic propagating outward at the periphery of the cluster (Menanteau et al. 2012, hereafter M12). This suggests a simple merger configuration with small impact variables. (3) The X-ray luminosity peak is shown to be offset from the weak-lensing peak by X kpc at $X \sigma$ level (J13). (4) The observation of the double radio relic suggests that the angle between the merger axis and the plane of the sky has to be reasonably small (M12, Lindner et al. 2013), or else the relic may appear as a halo instead. (Skillman et al. 2013)

In this paper, we perform results of simulations for modeling the time evolution of the mergers. Determining the time-since-collision of mergers of similar clusters helps us reconstruct different stages of a cluster merger. Mergers of clusters proceed on the time-scale of millions of year, observations of each cluster only provides a snapshot of a particular type of merger. In order to understand the merger process observationally, we need to capture and identify different stages of similar dissociative mergers.

Another crucial piece of missing information is the 3D configuration, i.e. the projection angle α , which contributes the largest amount of uncertainties to the dynamical variables (Dawson 2013). With a large projection angle α , the radio emission may appear as a radio halo instead. (Skillman et al. 2013)

This work is particularly important since it is forbiddingly expensive to simulate clusters similar to El Gordo in high resolution. The probability for finding an analog of El Gordo in a cosmological simulation is as low as % (M12). A realistic cosmological simulation of El Gordo is thus computationally expensive. Under the hierarchical picture of structure formation in the Λ CDM model, there is a rare chance for massive clusters like El Gordo to have formed at a redshift of $z = 0.87$. Staged simulation would not be able to probe the angular dependence. Both weak lensing analysis and BLAH DATA of El Gordo (Jee et al. 2014) has revealed a relatively simple bimodal mass distribution. The lack of complex substructures makes modeling of El Gordo with only two subclusters possible.

In this paper, we adopt the following conventions: (1) we assume the standard Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$. (2) All confidence intervals are quoted at the 68% level unless otherwise stated. (3) All credible intervals (a.k.a. Bayesian confidence intervals that also takes into account prior probability) are also quoted at the 68% level unless otherwise stated and are central credible intervals. (4) All quoted masses (M_{200c}) are based on mass contained within r_{200} where the mass density is 200 times the critical density of the universe (ρ_{crit}) at the redshift of $z = 0.87$.

2 DATA

In this paper we made use of data from multiple sources for different purposes. For assigning membership of galaxies to the two identified subclusters, we examined the spectroscopic data obtained from the Very Large Telescope (VLT) and Gemini South as described in M12 and Sifón et al. (2013). For the weak-lensing mass estimation, we used the published Monte Carlo Markov Chains (MCMC) mass esti-

mates from J13. See Table 1 for descriptions of the probability density functions (PDFs) of the input variables.

In order to constrain our parameter space, we referred to the properties of the radio relic from Lindner et al. (2014). El Gordo shows radio relics on the periphery of both subclusters (M12). Two radio relics, the northwest (NW) relic and the southeast (SE) relic, of El Gordo were first discovered in the Sydney University Molonglo Sky Survey (SUMSS) data in low resolution at 843 MHz (Mauch et al. 2003) as shown in M12. The higher resolution radio observation conducted by Lindner et al. (2014) at 610 MHz and 2.1 GHz confirms that the identity of the NW radio relic after removing effects of radio point sources. The NW radio relic, which possesses the most extended geometry (Mpc in length) among all radio source, was identified to be associated with the merger. We do not refer to the SE radio relic in this paper since the SE radio relic is reported to be caused by a compact radio source and the association of the E radio relic to the merger is unclear (Lindner et al. 2014).

3 METHOD – MONTE CARLO SIMULATION

For this analysis, we used the collisionless dark-matter-only Monte Carlo modeling code written by Dawson (2013), hereafter (D13). In the D13 code, the time evolution of the head-on merger was computed based on an analytical model assuming that the only dominant force is the gravitational attraction from the masses of two truncated Navarro-Frenk-White (hereafter NFW) DM halos. Other major assumptions for modeling systems with this code include negligible impact parameter (cite M11???) and no self-interaction of dark matter. No mass accretion during the merger.

In the Monte Carlo simulation, many realizations of the collision is computed from the inputs of each realization, including the data (\vec{D}) and the model variable (α). In particular, the standard required data, which were in the form of samples of the probability density functions (PDFs), included the masses ($M_{200_{NW}}, M_{200_{SE}}$) the redshifts (z_{NW}, z_{SE}) and the projected separation of the two subclusters (d_{proj}). In each realization, we randomly drew samples of the PDFs. These inputs are then used for computing the output variables ($\vec{\theta}'$) by making use of conservation of energy to describe their collision due to the mutual gravitational attraction. (See Table 1 for quantitative descriptions of the sample PDFs and we outline how those PDFs are obtained in the following subsections.) To ensure convergence of the output PDFs, in total, 2 million (to be confirmed) realizations were computed. The results, however, are consistent up to a fraction of a percent just from 20 000 runs (D13).

We adapt a Bayesian interpretation of the outputs of the Monte Carlo simulation. The Bayes chain rule underlies the simulation can be written as:

$$P(M|\vec{D}) \propto P(\vec{D}|M)P(M) \quad (1)$$

where the likelihood is defined to be the PDF of \vec{D} given our physical model M which we parametrize using variables in Table 1 ($\vec{\theta}$) by assuming conservation of energy, which is represented by a suitable functional form f below. For example, the calculation of the output variables of the j -th

Table 1. Properties of the sampling PDFs of the Monte Carlo simulation

| Data | Units | Location | Scale | Ref |
|----------------|-------------------------------|----------|---------|-----------------|
| $M_{200_{NW}}$ | $10^{14} h_{70}^{-1} M_\odot$ | 13.0 | 1.6 | J13 |
| c_{NW} | | 2.50 | 0.02 | J13 |
| $M_{200_{SE}}$ | $10^{14} h_{70}^{-1} M_\odot$ | 7.6 | 1.2 | J13 |
| c_{SE} | | 2.70 | 0.04 | J13 |
| z_{NW} | | 0.86842 | 0.00109 | M12, Sifón 2013 |
| z_{SE} | | 0.87110 | 0.00117 | M12, Sifón 2013 |
| d_{proj} | Mpc | 0.74 | 0.007 | J13 |

realization can be denoted as:

$$(\vec{\theta}')^{(j)} = M(\vec{\theta}^{(j)}, \vec{D}) \quad (2)$$

and computed over all j realizations. Finally, we took the physical constraints on the dynamical variables into account by examining the resulting physical variables against the physical limits and excluding realizations that would produce impossible values. We refer to this process of excluding unphysical realizations as applying priors. Even though we denote the priors for one dimension at a time (See Appendix A), the correlations between different variables are properly taken into account due to how we throw away all the variables of problematic realizations.

3.1 Inputs of the Monte Carlo simulation

3.1.1 Membership selection and redshift estimation of subclusters

We adopt the identification of galaxy membership of El Gordo given by M12 with a total count of 89 galaxies. To further distinguish member galaxies of each subcluster, we adopt a redshift cut of $z < 0.886$, and a spatial cut along (01:03:22.0, -49:12:32.9) and (01:02:35.1, -49:18:09.8). The spatial cut is approximately perpendicular to the 2D merger axis and are consistent with the bimodal number density contours (See Figure 2). We determined that there are 51 members in the NW subclusters and 35 members in the SE subclusters. The spatial cut indicated by the black line was done after mapping the world coordinates to pixel coordinates to avoid anamorphic distortion.

After identifying members of each subcluster, we performed 10, 000 bootstrap realizations to estimate the biweight locations of the spectroscopic redshifts of the respective members in order to obtain the samples of the PDFs of the redshifts of each subcluster. The spectroscopic redshift of the subclusters were determined to be $z_{NW} = 0.86842 \pm 0.0011$ and $z_{SE} = 0.87131 \pm 0.0012$, where the quoted numbers represent the biweight location and 1σ biased corrected confidence level respectively (Beers et al. 1990). Both the estimated redshifts of the subclusters and the uncertainties are consistent with estimates of $z = 0.8701 \pm 0.0009$ for El Gordo given by Sifón et al. 2013, and the fact that the member galaxies of El Gordo shows large velocity dispersion and has the largest velocity dispersion among all the ACT galaxy clusters as reported by M12.

We estimated the radial velocity differences of the subclusters by first calculating the velocity of each subcluster

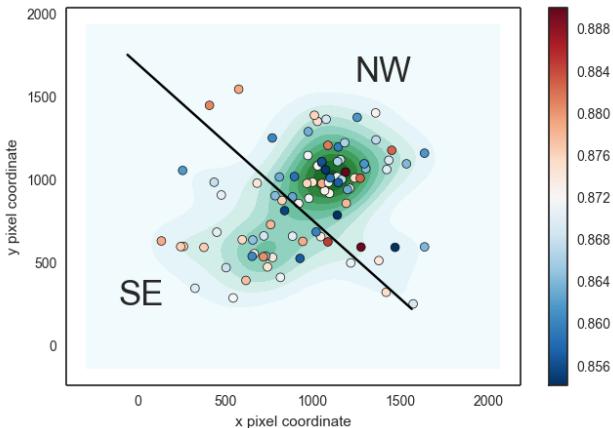


Figure 2. Points showing the locations of the member galaxies and the division of the member galaxies among the two subclusters of El Gordo by a spatial cut (black line). The color of the points shows the corresponding spectroscopic redshift of the member galaxies (see color bar for matching of spectroscopic values), with the redder end indicating higher redshift. The background number density contours in green indicate a bimodal distribution.

with respective to us, using

$$v_i = \left[\frac{(1+z_i)^2 - 1}{(1+z_i)^2 + 1} \right] c \quad (3)$$

where $i = 1, 2$ represents the two subclusters, and c is the speed of light. The radial velocity was calculated by:

$$\Delta v_{rad}(t_{obs}) = \frac{|v_2 - v_1|}{1 - \frac{v_1 v_2}{c^2}} \quad (4)$$

Due to the estimates of the subcluster redshifts are close to one another with overlapping confidence intervals, we obtained a low radial velocity difference of the two subclusters to be 476 ± 242 km / s. This result is lower than the quoted radial velocity differences of 586 km/s reported by M12 from the approximation of $\Delta v_{rad} = c(z_1 - z_2)/(1+z_1)$ in the frame of the NW subcluster.

3.1.2 Weak lensing mass estimation

We obtained the PDFs of the masses of the subclusters by doing a Monte Carlo Markov Chain (MCMC) analysis of the reduced shear from the weakly lensed background galaxies similar to Dawson et al. (2012). We computed the reduced shear signal generated by two NFW halos according to Umetsu (2010) (See Appendix ?? for details of implementation and output diagnostics). At each step we followed the procedure of a Metropolis algorithm. The transition kernel was set to be the log likelihood of fit of the model shear to the reduced shear of the data (??). In total, eight MCMC chains were used. After every 5000 MCMC steps for all the chains, we computed the R coefficient (Gelman & Rubin 1992) to check for convergence. We performed more MCMC steps as long as convergence was not achieved. After convergence was achieved, we removed the burn-in portions of the MCMC chains and used the resulting MCMC chains as samples of the PDFs of the masses.

We make use of an effective redshift of $z_{\text{eff}} = 1.37$ or $D_{LS}/D_L = 0.276$ (J13) and $g' \approx (1 + 0.79\kappa)g$ (J13, Seitz

& Schneider 1997). On the other hand, we fixed the position of the centers of the NFW halos to be the luminosity peaks of the respective galaxy populations of the two subclusters, which are at R.A. = 01:02:51.68, Decl. = $-49:15:04.40$ and R.A. = 01:02:38.38, Decl. = $-49:16:37.64$ for the NW and SE subclusters respectively (J13). The separations between the luminosity peaks and the estimated mass centroids of the subclusters are with several arcseconds, or around 30 kpc, for the both NW and SE subcluster (J13). The agreement between our analysis and (J13) to within the 68% credible interval serves as a sanity check on the estimated masses. The mass estimates from Jee et al. 2014 are $M_{200c} = 13.8 \pm 2.2 \times 10^{14} h_{70}^{-1} M_\odot$ for the NW subcluster and $M_{200c} = 7.8 \pm 2.0 \times 10^{14} h_{70}^{-1} M_\odot$ for the SE subcluster. We note that the results of the Monte Carlo simulation is not very sensitive to the estimated masses of the subclusters as discussed in section 4.2.

3.1.3 Estimation of projected separation (d_{proj})

To be consistent with our MCMC mass inference, our Monte Carlo simulation takes the projected separation of the NFW halos to be those of the two aforementioned luminosity peaks.

3.2 Outputs of the Monte Carlo simulation

We outline the outputs of the simulation here to facilitate the discussion of the design of the priors used in the simulation. The simulation provides PDF estimates for 10 output variables. Variables of the most interest include the time dependence and the angle α , which is defined to be the projection angle between the plane of the sky and the merger axis. Other output variables are dependent on α and the time dependence. Specifically, the simulation denotes the time dependence by providing several characteristic timescales, including the time elapsed between consecutive collisions (T) and the time-since-collision of the observed state (TSC).

We provide two versions of the time-since-collision variables TSC_0 and TSC_1 to denote different possible merger scenarios. 1) We call the scenario for which the subclusters are moving apart after collision to be “outgoing” and it corresponds to the smaller TSC_0 value, and 2) we call the alternative scenario “returning” for which the subclusters are approaching each other after turning around from the apoapsis for the first time and it corresponds to TSC_1 . We describe how we make use of properties of the radio relic to evaluate which scenario is more likely in section 3.4.

The simulation also output estimates of variables that describe the dynamics and the characteristic distances of the merger. The relative 3D velocities of the subclusters, both at the time of the collision ($v_{3D}(t_{col})$) and at the time of observation ($v_{3D}(t_{obs})$) are provided. The characteristic distances included in the outputs are the maximum 3D separation (d_{max}), which is the distance between the position of collision to the apoapsis; the projected separation d_{proj} and the 3D separation of the subclusters at observation.

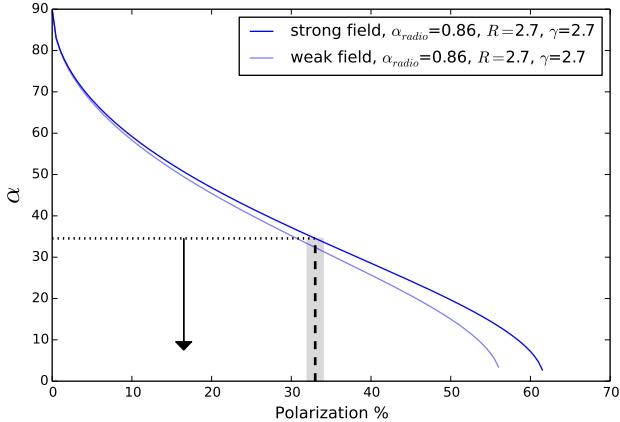


Figure 3. Predictions of polarization percentage of the radio relic at a given projection angle from different models, reproduced from Ensslin et al. 1998. Each model assumes electrons producing the radio emission to be accelerated inside uniform magnetic field of various strengths (*strong* or *weak*). The curves are plotted with spectral index of the radio emission (α_{radio}), spectral index of the electrons (γ) and compression ratio of the magnetic field (R) corresponding to the estimated values from Lindner et al. (2014). We highlight the observed polarization percentage of the main NW radio relic of El Gordo by the dotted vertical line with the greyed out region indicating the uncertainty (Lindner et al. 2014).

3.3 Design and application of priors

One of the biggest strengths of the Monte Carlo simulation by D13 is its ability to detect and rule out extreme input values that would result in unphysical realizations via the application of prior probability. Our default priors are described in D13 and we include them in Appendix B for the convenience of the readers. In addition, we have come up with a prior on the projection angle α based on the polarization fraction of the radio relic.

3.3.1 Monte Carlo priors based on the polarization fraction of the radio relic

We can relate the polarization fraction of the radio relic to the projection angle by examining the generating mechanism of the radio relic. The observed radio relic is due to synchrotron emission of free electrons in a magnetic field. If the magnetic field is uniform, the observed polarization fraction of the synchrotron emission of the electrons depends on the viewing angle (or equivalently the projection angle) with respect to the alignment of the magnetic field. Synchrotron emission from electrons inside unorganized magnetic field are randomly polarized. The high reported integrated polarization fraction from Lindner et al. (2014) can be explained by a highly aligned magnetic field, created by the compressed intracluster medium during a merger (Ensslin et al. 1998, van Weeren et al. 2010, Feretti et al. 2012). This picture is consistent with a high polarization fraction perpendicular to this magnetic field along the relic.

We designed our prior to reflect how α decreases monotonically as the maximum observable integrated polarization fraction. This assumption is based on the class of models given by Ensslin et al. (1998)(See Figure 3). In particular, we refer to a model from Ensslin et al. (1998) that would

give the most conservative estimate on the upper bound of α :

$$\alpha = 90^\circ - \arcsin \left(\sqrt{\frac{\frac{2}{15} \frac{13R-7}{R-1} \frac{\gamma+7/3}{\gamma+1} \langle P_{strong} \rangle}{1 + \frac{\gamma+7/3}{\gamma+1} \langle P_{strong} \rangle}} \right) \quad (5)$$

This model corresponds to the strong field case with the relic being supported by magnetic pressure only, with $\alpha_{radio} = 0.86$, compression ratio $R = 2.7$ and $\gamma = 2.7$. This model predicts a maximum integrated polarization fraction of $\sim 60\%$ when $\alpha \rightarrow 0$. From this model, the observed integrated polarization fraction of $33\% \pm 1\%$ corresponds to an estimated value of $\hat{\alpha} = 35^\circ$. This polarization fraction of $\sim 60\%$ predicted by (Ensslin et al. 1998) is consistent with the upper bound of relic polarization fraction in cosmological simulations (Skillman et al. 2013). No other model of the magnetic field should predict a higher polarization fraction, thus it is highly unlikely that we see 33% integrated polarization at $\alpha > 35^\circ$.

We cannot rule out $\alpha \leq 35^\circ$ as a result of possible variations in the magnetic field. Ensslin et al. (1998) assumes an isotropic distribution of electrons in an isotropic magnetic field. Cosmological simulations of radio relics from Skillman et al. (2013) show varying polarization fraction across and along the relic assuming $\alpha = 0$, resulting in a lower integrated polarization fraction. For example, it is possible to see a edge-on radio relic ($\alpha = 0$) with integrated polarization fraction of 33% . Furthermore, Skillman et al. (2013) shows that after convolving the simulated polarization signal with a Gaussian kernel of $4'$ to illustrate effects of non-zero beam size, the polarization fraction drops to between 30% to 65% even when $\alpha = 0$. Other uncertainties come from the fact that the inferred spectral indices differ between the two observed frequencies and vary between the three identified relic sources (Lindner et al. 2014). We examine the effects of changing the cutoff value of this prior to ensure the uncertainties do not introduce significant bias in the estimated output variables and we present the results in Appendix B. To summarize, we adopt a conservative uniform prior to encapsulate the information from the polarization fraction of the radio relic as:

$$P(\alpha) = \begin{cases} \text{const.} > 0 \text{ for } \alpha < 35^\circ \\ 0 \text{ otherwise} \end{cases} \quad (6)$$

3.4 Extension to the Monte Carlo simulation - Determining merger scenario with radio relic position

One of the biggest question involving the merger is if El Gordo was observed to be in a returning or outgoing phase. We perform a qualitative comparison of the observed and the simulated position of the NW radio relic in each scenario. We considered different possibilities of the time evolution of the shock velocity due to lack of knowledge of the details of the evolution. We drew physical insight from the simulations of the merger shock of the staged numerical simulation of the Bullet cluster from Springel & Farrar (2007) and the cosmological simulation from Paul et al. (2011). Right after the collision of the subclusters, Springel & Farrar (2007) shows that the shock speed is comparable to the merger speed of the two subclusters; the shock speed dropped only

by $\sim 14\%$ in the 300 Myr period while the speed of the main subcluster in the simulation dropped by $\sim 65\%$ in the center of mass frame. On the other hand, Paul et al. (2011) showed that the shock speed was ~ 1.5 times the relative collisional speed of the subcluster shortly after the collision and the shock speed decreases only slightly as it propagates away from the center of mass.

We approximated the upper and lower bounds of the NW relic speed with the simulated speeds of the NW subcluster. We worked in the center of mass frame where the shock speed is expected to drop slightly with time. The projected separation of the shock is approximated as:

$$s_{proj} = \langle v_{relic} \rangle (\hat{t}_{obs} - \hat{t}_{col}) \cos(\hat{\alpha}) \quad (7)$$

where the quantities with hats on the right hand side of the equations were inferred from the simulation, and s_{proj} is the estimated projected separation and we estimated the upper and lower bounds of the time-averaged velocity $\langle v_{relic} \rangle$ of the shock between the collision of the subclusters and the observed time as:

$$\langle v_{relic} \rangle = \beta v_{3D,1}(t_{col}) \quad (8)$$

where $0.8 \leq \beta \leq 1.2$ is a factor that we introduce to represent the uncertainty of the speed of the relic and $v_{3D,1}(t_{col})$ refers to the collisional velocity of the NW relic in the center-of-mass frame. The upper bound can be approximated as the collisional speed of the NW subclusters due to how the shock is powered by the collision. After the collision, it is unlikely that there would be significant energy injected into the shock to speed up the shock such that the shock travels much faster than the collision speed of the subcluster. While the shock does not experience gravitational deceleration as a pressure wave, some dissipative processes may have slowed down the shock wave slightly as it propagated. By making use of a range of β values, we examine how the rate of slow down would give a different lower bound of the projected separation of the relic.

4 RESULTS

We present an overview of the estimated physical parameters can be viewed in table 2. We estimated the relative merger speed of the two subclusters to be 2400 ± 400 km s $^{-1}$ at the time of collision. There have been several estimates of the relative merger speed based on different physics. For example, Lindner et al. (2014) inferred the upper bound of the relative collisional velocity from the Mach number of the NW radio relic, giving an upper bound estimate of 2500 ± 400 km s $^{-1}$. This speed is compatible with our simulated collisional speed within the uncertainties. However, we note that the conversion of the frame of reference is non-trivial since the speed is measured with respect to the turbulent intracluster gas as a Mach number. On the other hand Molnar and Broadhurst found the scenario where 2400 km s $^{-1}$ to be the most likely velocity for matching the morphology of the X-ray wake feature. Based on the observed peculiar velocity difference between the galaxies of the two subclusters, $v_{3D}(t_{col}) \approx v_{pec}(t_{obs})/\sin(\alpha)$, M12 gave

two estimates:

$$v_{3D}(t_{col}) = 586/\sin(\alpha) \text{ km s}^{-1} = \begin{cases} 2300 \text{ km s}^{-1}, & \alpha = 15^\circ \\ 1200 \text{ km s}^{-1}, & \alpha = 30^\circ \end{cases} \quad (9)$$

If we use our estimate of $\alpha = 41.7^\circ$ in (9), we obtain an estimate of $v_{3D}(t_{obs}) \approx 940$ km s $^{-1}$ instead. This is consistent with our simulation.

4.1 Time-since-collision

* estimate and compare to sound crossing time!

The simulation gives two estimates for the time-since-collision, with $TSC_0 = 0.61$ Gyr and $TSC_1 = 1.0$ Gyr. Based on section 3.4, we present estimates for the position of the NW radio relic based on the two PDFs of inferred TSC as shown in Fig. 4. We plotted the PDF of the simulated positions of the relic for the outgoing scenario (blue) and the returning scenario (green) against the observed position of the relic (light red), with the width of the relic accounting for the uncertainty due to the location of the center of mass, this uncertainty is greater than the width of the relic (~ 23 kpc) along merger axis (Lindner et al. 2014). When we assume $\langle v_{relic} \rangle/v_{3D,1}(t_{col}) = 1.5$ (uppermost panel of Figure 4), $\langle v_{relic} \rangle$ corresponds to an time-averaged relative velocity of ~ 3800 km s $^{-1}$ or, ~ 1400 km s $^{-1}$ in the center of mass frame, the outgoing scenario is more favored. As we examine a decreased ratio of $\langle v_{relic} \rangle/v_{3D,1}(t_{col})$, we probe how much the shock could have slowed down and still be consistent with the outgoing scenario. The second (relative velocity of 2800 km s $^{-1}$) and third row of Fig. 4 show that if $\langle v_{relic} \rangle \lesssim 1.1$, the returning scenario would be favored instead. Estimates of Lindner et al. (2014) of the shock velocity corresponds to 2500 ± 400 km s $^{-1}$. A definite conclusion about the best fit scenario would require better knowledge of the time evolution of the shock velocities.

Hydrodynamical simulations from Donnert et al. and Molnar and Broadhurst also give estimates of the TSC and the phase of collision.

4.2 Effects of applied prior on output variables

We performed tests of how each of the output variables vary according to the choice of the cutoff of the polarization prior between $\alpha_{cutoff} = 29^\circ$ to 49° instead of 35° . We found that in the most extreme case, choosing the cutoff values as 29° (-6°), the location of the $v_{3D}(t_{obs})$, is increased by 16%. While the upper 95% CI of d_{max} is the most sensitive to the prior and it changes by $\sim 20\%$ when $\alpha_{cutoff} = 49^\circ$. This shows that the exact choice of the cut off value for α does not affect our estimates significantly.

4.3 Three-dimensional (3D) configuration of El Gordo

Many previous studies (van Weeren et al. CITATIONS) have suggested that α should be small for the detection of double radio relics to be possible but did not provide quantitative constraints.

Most of the uncertainty of the parameter estimates come from the projection angle as shown in Fig (X). Without

Table 2. Table of the output PDF properties of the model variables and output variables from Monte Carlo simulation

| Variables | Units | Default priors | | | Default + position priors | | |
|--------------------|--------------------|----------------|-----------|-----------|---------------------------|-----------|-----------|
| | | Location | 68% CI † | 95% CI | Location | 68% CI | 95% CI |
| α | (degree) | 43 | 19-69 | 6-80 | 21 | 10-30 | 3-34 |
| d_{proj} | Mpc | 0.74 | 0.74-0.75 | 0.73-0.76 | 0.74 | 0.74-0.75 | 0.73-0.76 |
| d_{max} | Mpc | 1.2 | 0.9-2.2 | 0.77-4.6 | 0.93 | 0.81-1.2 | 0.75-1.9 |
| d_{3D} | Mpc | 1 | 0.79-2.1 | 0.75-4.3 | 0.8 | 0.76-0.88 | 0.74-0.91 |
| TSC_0 | Gyr | 0.61 | 0.4-0.95 | 0.26-2.4 | 0.46 | 0.3-0.55 | 0.21-0.64 |
| TSC_1 | Gyr | 1 | 0.77-1.7 | 0.63-4.4 | 0.91 | 0.69-1.3 | 0.59-2.3 |
| T | Gyr | 1.6 | 1.3-2.6 | 1.2-7.1 | 1.4 | 1.2-1.6 | 1.2-2.4 |
| $v_{3D}(t_{obs})$ | km s ⁻¹ | 580 | 260-1200 | 59-2400 | 940 | 360-1800 | 62-2900 |
| $v_{rad}(t_{obs})$ | km s ⁻¹ | 360 | 140-630 | 27-880 | 310 | 110-590 | 8-840 |
| $v_{3D}(t_{col})$ | km s ⁻¹ | 2800 | 2400-3700 | 2100-4200 | 2400 | 2200-2800 | 2100-3500 |

† CI stands for credible interval

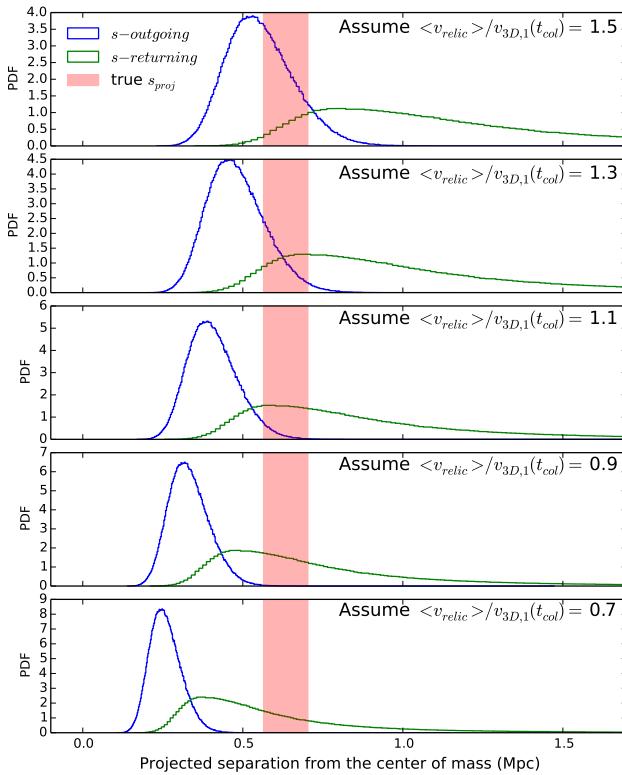


Figure 4. Comparison of the observed position of the relic (red bar includes uncertainty) with the predicted position from the two simulated merger scenarios (blue for outgoing and green for the returning scenario). The outgoing scenario is more favored than the returning scenario since the shock speed is unlikely to travel at much less than $0.6 v_{3D,1}(t_{col})$ for a significant period of time. We made use of the polarization prior for this figure. Alternate version of this figure with the default prior applied can be found in Appendix B.

the default priors, the Monte Carlo simulation gives the estimate of the projection angle as 41.7° , with the CI = 22.7° , 61.14° . After applying the radio relic polarization prior, the CI shrinks to $^\circ$ and $^\circ$. Both estimates are consistent with the estimate of $\alpha > 7.8^\circ$ from Lindner et al. (2014) based on the dynamics of the radio relic.

More work from N-body magnetodynamical simulations is needed to better understand about the physics and the

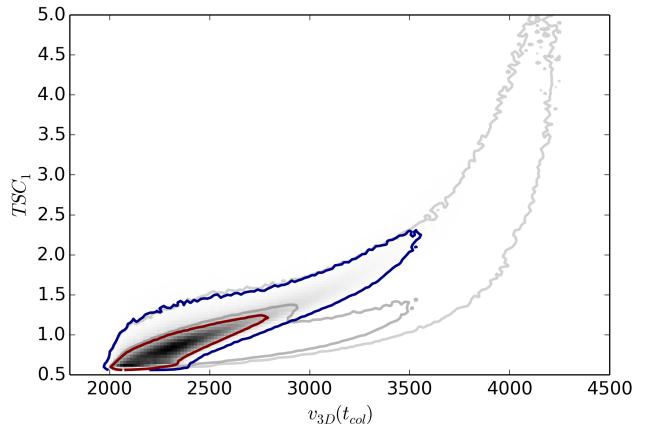


Figure 5. The marginalized output PDF of the outgoing time-since-collision (TSC_0) vs. the 3D velocity at the time of collision for El Gordo. The grey set of contours show the credible regions before applying the polarization prior and the colored contours correspond to the credible regions after applying the priors. The contours represents the 95% and 68% credible regions respectively.

observable constraints of radio relics on merger dynamics from cosmological simulations such as, More observational constraints

(More speculative stuff should be put at the end.) From this simulation we have shown that it is possible to detect the double radio relics with α being as big as 61.14° . More concrete conclusions can only be drawn with better understanding of the radio relic properties from simulations and observations.

- explains that there hasn't been quantitative constraints on the angle for which double radio relic can be observed, even though that many studies have suggested that the detection of radio relic should imply that α should be small. From this simulation we have shown that it is possible to detect the double radio relics with α being as big as 61.14° .

- James' paper did mention how the mass estimation depends on α , with the estimated mass being a lot smaller if $\alpha \geq 65^\circ$. However, since we did use the larger mass estimate as the input of this simulation, we can only say that the inferred α is consistent with the mass estimation.

- discussion of the different scenarios mention in M12:

- 1) we are viewing after core passage, but before first turn around, and the merger speed is low"
- 2) the merger speed is high, but we are viewing after the first turn around as the two components come together for a second core passage
 - discuss the inclination angle estimate from M12
 - Dave: explain where the limits of the projection angle comes from. what observational evidence contradicts the low velocity scenario the most

With this new piece of evidence, we find that the absence of an X-ray shock feature from El Gordo, may not be due to the merger speed being low, as suggested by (J13), but due to the long TSC. In particular, taking into account that the estimated projection angle of $\sim 41.7^\circ$, we estimate the projected relative velocity to be 597 km s^{-1} , which is consistent with the estimated line-of-sight velocity differences of $586 \pm 96 \text{ km s}^{-1}$ in M12.

Furthermore, the study from Lindner et al. (2014) Lindner et al. has come up with an estimation of the shock velocity of the radio relic of El Gordo as $\sim 4000 \text{ km s}^{-1}$. While this shock velocity is not the same as the merger velocity, they should be of similar magnitude. Indeed our simulation found that a merger velocity of 4000 km s^{-1} is within the 95% credible interval.

5 DISCUSSION

5.1 Our finding in the context of other studies of El Gordo

Compare to Lindner et al. (2014). Compare to Donnert (2014) for their best fit scenario.

Initial velocity is higher for the hydrodynamical simulations than our Monte Carlo simulations.

This is reflected by the low probability for our Monte Carlo simulations to reach d_{max} . This is an aspect that should be further investigated in future cosmological simulations.

5.2 Comparison to other merger clusters of galaxies

Estimates of the period of El Gordo is smaller than both Musketball and the Bullet Cluster due to the larger masses of its subclusters.

Talks about how El Gordo is more massive and collided at higher speed than both the Bullet and the Musketball, so El Gordo is probably a better probe of SIDM properties.

With this new piece of evidence, we find that the absence of an X-ray shock feature from El Gordo, may not be due to the merger speed being low, as suggested by J13. In particular, taking into account that the estimated projection angle of $\sim 41.7^\circ$, we estimate the projected relative velocity to be 597 km s^{-1} , which is consistent with the estimated line-of-sight velocity differences of $586 \pm 96 \text{ km s}^{-1}$ in M12.

Furthermore, the study from Lindner et al. (2014) Lindner et al. has come up with an estimation of the shock velocity of the radio relic of El Gordo as $\sim 4000 \text{ km s}^{-1}$. While this shock velocity is not the same as the merger velocity, they should be of similar magnitude. Indeed our simulation

found that a merger velocity of 4000 km s^{-1} is within the 95% credible interval.

5.3 Possible improvements for the radio relic prior

needs better simulation to understand the physical properties of radio relic.

position information can be used better to constraint the angle / the TSC, projection effects - can be degenerate according to Skillman et al. (2013)

5.4 Limitations of our model and future work

Impact parameter of El Gordo may not be negligible. Simulations from Ricker & Sarazin (2001) showed that cool-core is not disrupted when the impact parameters of mergers are of the order of $\sim 500 \text{ kpc}$.

6 SUMMARY & CONCLUSION

(Maybe not needed if there is a discussion) This paper presents one of the first examples of using the observed radio relic emission to constrain cluster merger properties. While we have demonstrated how to use the physical properties of the radio relic emission to constrain merger dynamics and configurations, many improvements can still be made as more studies of radio relic are being done from both cosmological simulations and observations.

Currently, there are only a few studies of radio relic available for a range of viewing angles ((alias?), one of Brüggen's paper). As more cosmological simulations inform us if the relic is observable at certain viewing angles will help us come up with better Monte Carlo filters.

7 ACKNOWLEDGEMENTS

We thank Franco Vazza and Marcus Brüggen for sharing their knowledge on the simulated properties of radio relic. We extend our gratitude to Reinout Van Weeren for first proposing the use of radio relic as prior. We appreciate the comments from Maruša Bradač about using the position of the relic to break degeneracy of the merger scenario.

REFERENCES

- Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32, doi:10.1086/115487
- Dawson W. A., 2013, ApJ, 772, 131, 1210.0014, doi:10.1088/0004-637X/772/2/131
- Dawson W. A. et al., 2012, ApJ, 747, L42, 1110.4391, doi:10.1088/2041-8205/747/2/L42
- Donnert J. M. F., 2014, MNRAS, 438, 1971, 1311.7066, doi:10.1093/mnras/stt2291
- Ensslin T. A., Biermann P. L., Klein U., Kohle S., 1998, A&A, 332, 395, 9712293
- Feretti L., Giovannini G., Govoni F., Murgia M., 2012, aap, 558, 1205.1919
- Gelman A., Rubin D. B., 1992, Stat. Sci., 7, 457

- Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, ApJ, 785, 20, 1309.5097, doi:10.1088/0004-637X/785/1/20
- Lindner R. R. et al., 2014, ApJ, 786, 49, 1310.6786, doi:10.1088/0004-637X/786/1/49
- Mauch T., Murphy T., Buttery H. J., Curran J., Hunstead R. W., Piestrzynski B., Robertson J. G., Sadler E. M., 2003, MNRAS, 342, 1117, doi:10.1046/j.1365-8711.2003.06605.x
- Menanteau F. et al., 2012, ApJ, 748, 7, 1109.0953, doi:10.1088/0004-637X/748/1/7
- Paul S., Iapichino L., Miniati F., Bagchi J., Mannheim K., 2011, ApJ, 726, 17, doi:10.1088/0004-637X/726/1/17
- Seitz C., Schneider P., 1997, A&A, 318, 687
- Sifón C. et al., 2013, ApJ, 772, 25, 1201.0991, doi:10.1088/0004-637X/772/1/25
- Skillman S. W., Xu H., Hallman E. J., O’Shea B. W., Burns J. O., Li H., Collins D. C., Norman M. L., 2013, ApJ, 765, 21, doi:10.1088/0004-637X/765/1/21
- Springel V., Farrar G. R., 2007, MNRAS, 380, 911, doi:10.1111/j.1365-2966.2007.12159.x
- Umetsu K., 2010, p. 30, 1002.3952, doi:10.3254/978-1-60750-819-9-269
- van Weeren R. J., Röttgering H. J. a., Brüggen M., Hoeft M., 2010, Science, 330, 347, doi:10.1126/science.1194293

APPENDIX A: BAYESIAN PRIORS USED FOR DAWSON’S MONTE CARLO SIMULATION

The default prior probabilities that we employed can be summarized as follows for most of the output variables:

$$P(TSC_0) = \begin{cases} \text{const if } TSC_0 < \text{age of universe at } z = 0.87 \\ 0 \text{ otherwise} \end{cases} \quad (\text{A1})$$

In addition, we apply the following prior on TSC_1 only when evaluating the statistics of TSC_1 , thus allowing realizations with a valid outgoing TSC but an invalid returning TSC_1 .

$$P(TSC_1) = \begin{cases} \text{const if } TSC_1 < \text{age of universe at } z = 0.87 \\ 0 \text{ otherwise} \end{cases} \quad (\text{A2})$$

To correct for observational limitations, we further convolve the posterior probabilities of the different realizations with

$$P(TSC_0|T) = 2 \frac{TSC_0}{T} \quad (\text{A3})$$

to account for how the subclusters move faster at lower TSC and thus it is more probable to observe the subclusters at a stage with a larger TSC .

APPENDIX B: PLOTS OF OUTPUTS OF THE MONTE CARLO SIMULATION

We present the PDFs of the inputs of the dynamical simulation and the output PDFs of the outputs after applying the polarization prior in addition to the default priors.

This paper has been typeset from a T_EX/ L^AT_EX file prepared
by the author.

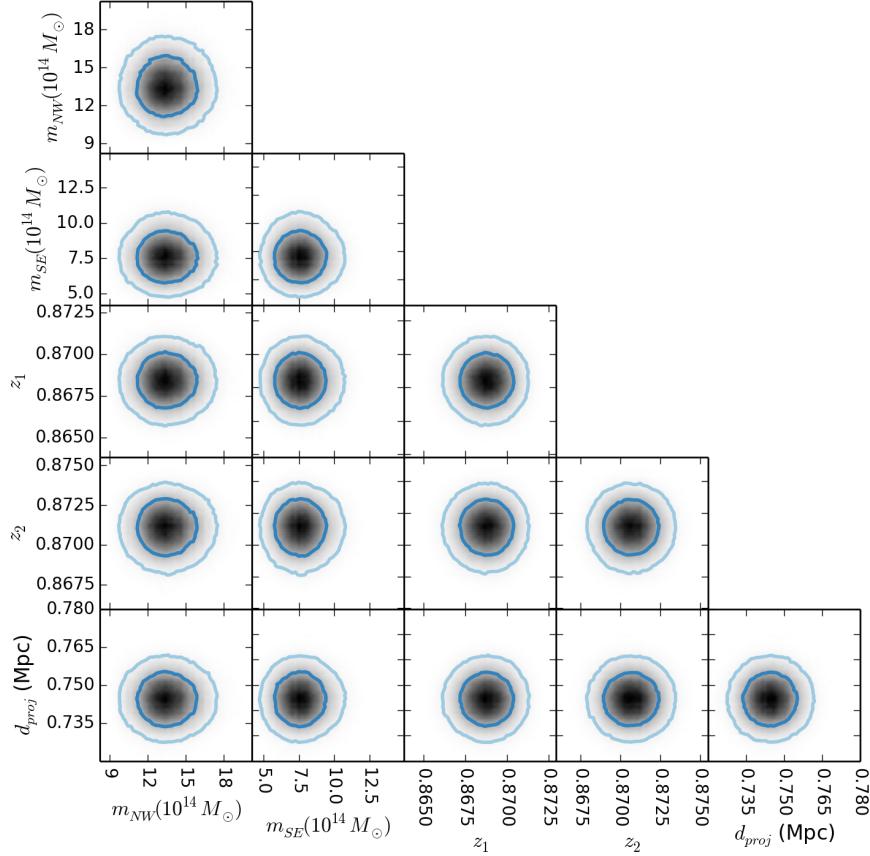


Figure B1. Marginalized PDFs of original inputs (vertical axis) and the inputs after applying polarization prior and default priors (horizontal axis). The inner and outer contour denote the central 68% and 95% credible regions respectively. The circular contours show that the application of priors did not introduce uneven sampling of inputs.

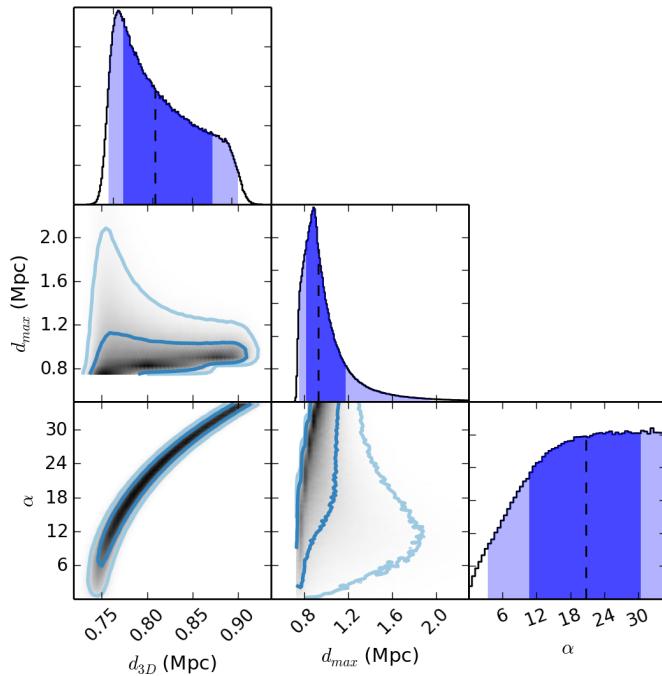


Figure B2. One-dimensional marginalized PDFs (panels on the diagonal) and two-dimensional marginalized PDFs of variables denoting characteristic distances and projection angle of the mergers.

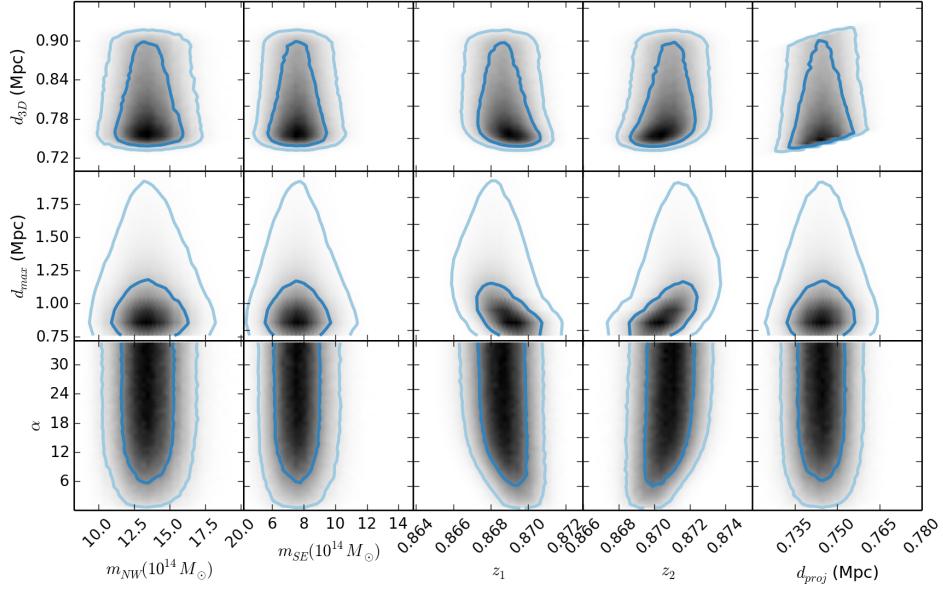


Figure B3. Marginalized PDFs of characteristic distances and projection angle of the merger and the inputs of the simulation.

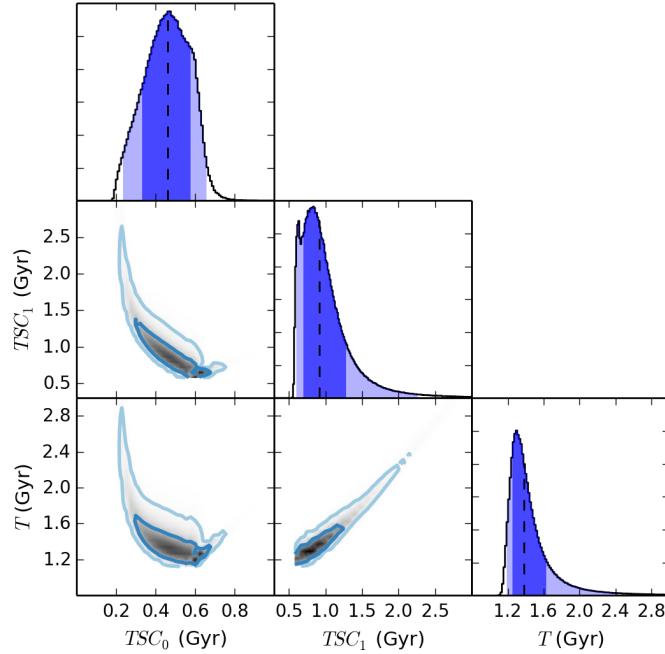


Figure B4. One-dimensional PDFs of characteristic timescales of the simulation (panels on the diagonal) and the marginalized PDFs of different timescales.

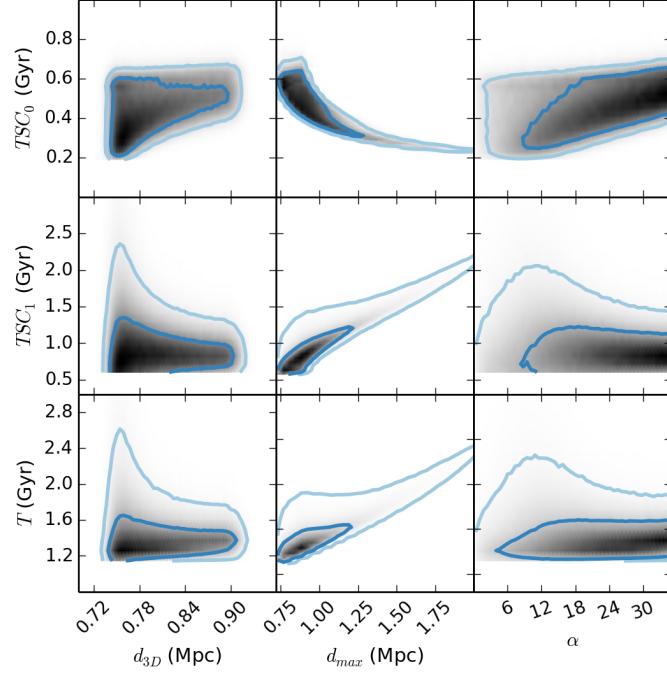


Figure B5. Marginalized PDFs of characteristic timescales of the simulation and the characteristic distances and the projection angle of the merger.

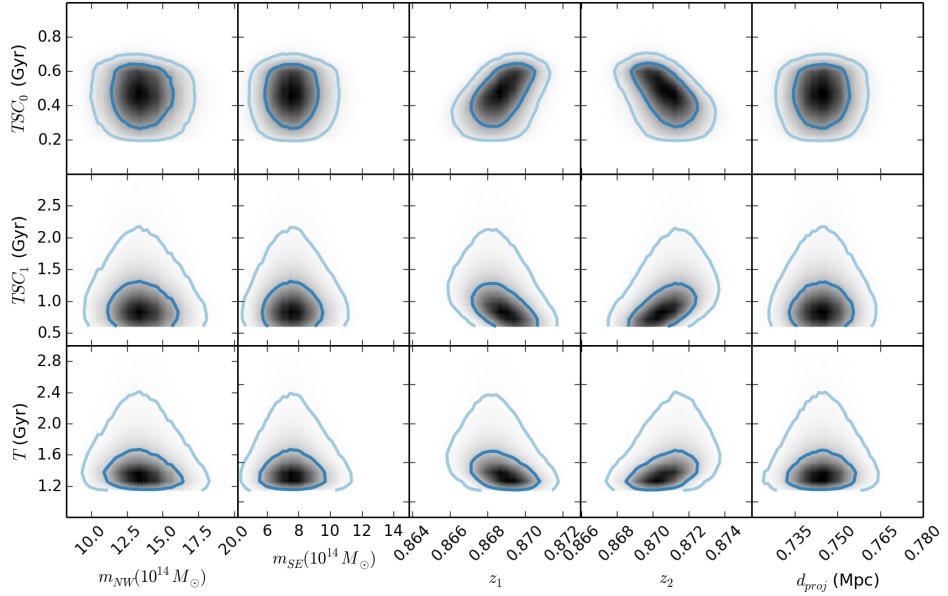


Figure B6. Marginalized PDFs of characteristic timescales of the simulation and the inputs.

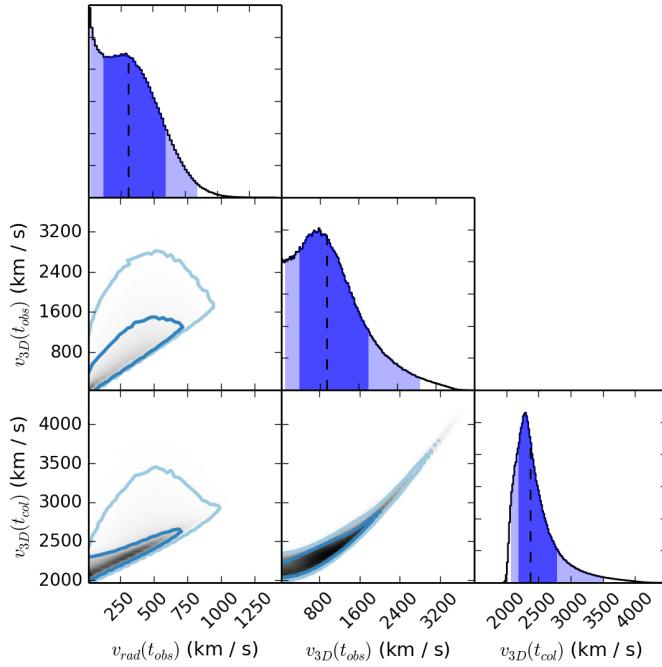


Figure B7. One-dimensional marginalized PDFs of velocities at characteristic times (panels on the diagonal) and marginalized PDFs of different velocities.

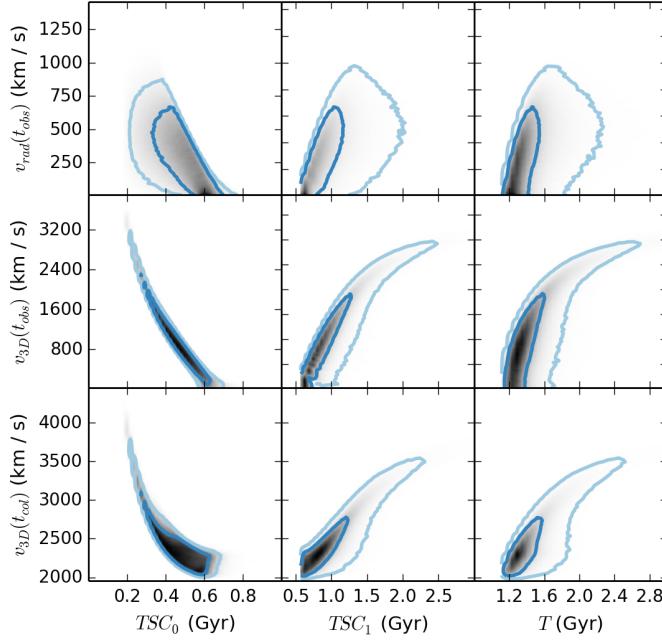


Figure B8. Marginalized PDFs velocities and the characteristic timescales of the simulation against the inputs.

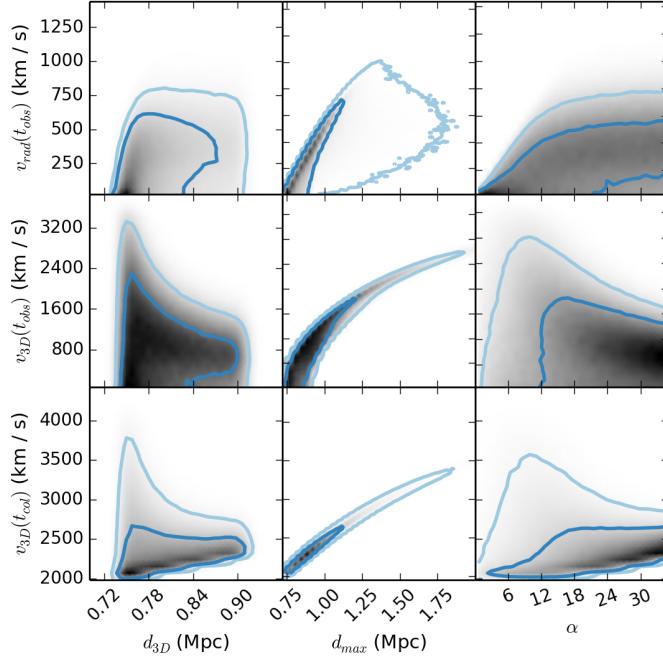


Figure B9. Marginalized PDFs of the velocities at characteristic timescales and the characteristic distances and the projection angle of the merger.

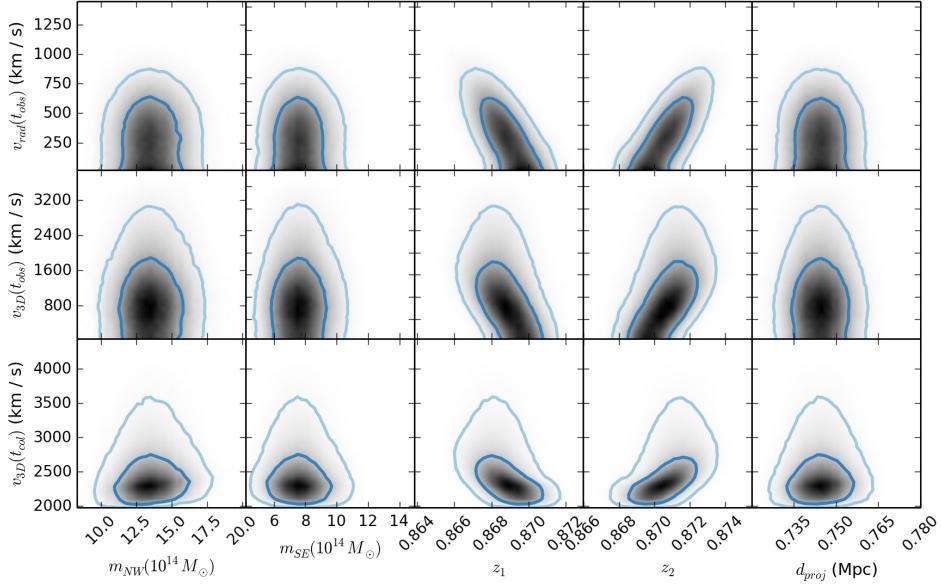


Figure B10. Marginalized PDFs of relative velocities characteristic timescales of the simulation and the inputs.