

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

K. Y. Ng<sup>1</sup>, W. A. Dawson<sup>2</sup>, D. Wittman<sup>1</sup>, J. Jee<sup>1</sup>, J. Hughes<sup>3</sup>, F. Menanteau<sup>3</sup>, C. Sifón<sup>4</sup>  
*(temporary order)*

<sup>1</sup>Department of Physics, University of California Davis, One Shields Avenue, Davis, CA 95616, USA

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

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

<sup>4</sup>Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, Netherlands

arXiv 666

## ABSTRACT

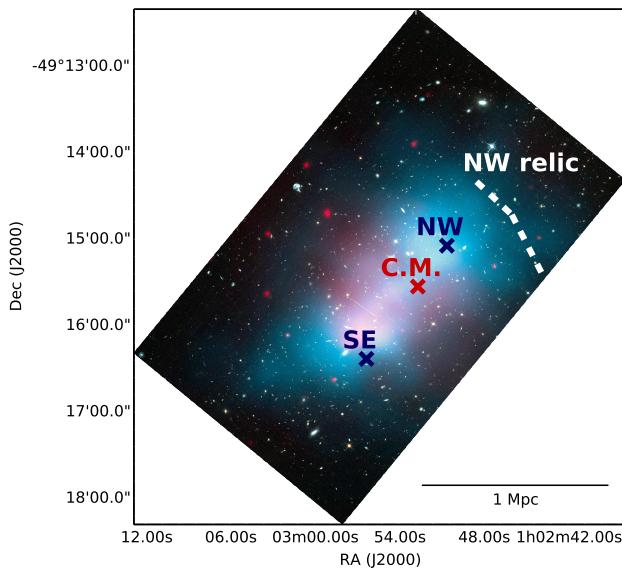
Merging galaxy clusters with radio relics provide rare insights to the merger dynamics as the relics are created by the violent merger process. We demonstrate one of the first uses of the properties of the radio relic to reduce the uncertainties of the dynamical variables and determine 3D configuration of a cluster merger, ACT-CL J0102-4915, El Gordo. From the double radio relic observation and X-ray wake morphology, it is believed that El Gordo is observed shortly after the first core-passage of the subclusters. We employ a Monte Carlo simulation to investigate the three-dimensional (3D) configuration and dynamics of El Gordo and provide a summary of the probability density functions of the inferred variables. By making use the polarization fraction of the radio relic, we are able to constrain the estimate of the angle between the sky and the merger axis to be  $\alpha = 21^\circ \pm 9^\circ$ . We find the relative 3D merger speed of El Gordo to be  $2400 \pm 400$  km s<sup>-1</sup>. We put our estimates of the time-since-collision into context by showing that if the time-averaged shock velocities  $\gg 1400$  km s<sup>-1</sup> in the center of mass frame, the two subclusters are more likely to be moving away from each other than moving towards each other. Finally, we compare our simulation results with other merging clusters and show that El Gordo has very similar dynamical properties as the Bullet Cluster.

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

## 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  $\sim 80\%$  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). During a merger of clusters, the subclusters are accelerated to high speeds of several thousand km s<sup>-1</sup>. The offsets of different components of the subclusters reflect the differences in the strengths of interactions between various components. Galaxies are expected to lead the gas due to its negligible interaction cross sections with other components. The intracluster medium (ICM) is expected to lose momentum through electromagnetic interactions. On the other hand, offsets between dark matter and galaxies may suggest dark matter self-interaction (Kahlhoefer et al. 2013, Randall et al. 2008).

El Gordo possesses a range of special features that allows us to constrain the merger dynamics in multiple ways. Ever since the discovery of El Gordo in the Atacama Camera Telescope survey (ACT; Marriage et al. 2011), there is an ongoing effort for collecting comprehensive data for El Gordo. 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. (2014). 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., turbulent flow due to object of higher density moving through fluids, while in the southeast, it shows highest X-ray emissivity indicative of a cold gas core south-



**Figure 1.** Configuration of El Gordo showing overlay of dark matter distribution in blue, and X-ray emissivity in red. The cross markers show the positions of the northwest (NW) and southeast (SE) dark matter density peaks, and the center of mass (CM) location respectively. Note that the mass ratio of the NW subcluster to the SE subcluster is  $\sim 2 : 1$ . The dashed white line indicates the approximate location of the northwest radio relic (NW relic). (Image Credit: NASA, ESA and Jee et al. 2014).

east of the wake. The cold gas core may have passed from the northwest to the southeast to have caused this morphology (Menanteau et al. 2012, 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 outskirt, 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 associated with shockwaves created from a merger (Lindner et al. 2014). El Gordo is one of small sample of galaxy clusters ( $\sim 50$ ) that have been associated with a radio relic and show dissociation between the X-ray gas and the DM subclusters. Even fewer of them have been studied in great details, making El Gordo a valuable candidate for further analysis.

In this paper, we combine most of the available information of El Gordo with the main goal of giving estimates of the associated dynamical parameters after taking into account all available constraints and reflecting uncertainties due to the missing variables. 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 identify different stages of similar dissociative mergers and gather statistics to understand the physics of the mergers. Another crucial piece of missing information is the 3D configuration, i.e. the angle between the plane of the sky and the merger axis called the projection angle  $\alpha$ . Since most of the dynamical observables

are projected quantities while the modelling of physics requires 3D variables, the deprojection based on  $\alpha$  contributes the largest amount of uncertainties to the dynamical variables (D13). From the morphology of the double relic of El Gordo, it is believed that  $\alpha$  should be small. For mergers with a large projection angle, the radio emission would be projected towards the center of the merger, which is hard to be detected (Vazza et al. 2012). However, the only quantitative constraints on  $\alpha$  is from Lindner et al. (2014) with a lower bound of  $\alpha \geq 11.6^\circ$ .

We employ a data-driven approach that thoroughly probes parameter space by directly drawing samples from the observables. This work based on Monte Carlo simulation is particularly important since it is forbiddingly expensive to simulate and analyze clusters similar to El Gordo in high resolution. Previous attempts at modeling El Gordo with hydrodynamical simulations such as Donnert (2014) and Molnar & Broadhurst (2014) provided only in total a dozen possible configurations of El Gordo, which do not reflect the input uncertainties. Another approach for estimating dynamical parameters would be to look for multiple analogs of El Gordo in cosmological simulations. However, under the hierarchical picture of structure formation in the  $\Lambda$ CDM model, there is a rare chance for massive clusters like El Gordo to have formed at a redshift of  $z = 0.87$ . The number density of analogs with mass comparable to El Gordo in a cosmological simulation is as low as  $10^{-11} \text{ Mpc}^{-3}$  (M12).

In the following sections, we adopt the following conventions: (1) we assume the standard  $\Lambda$ 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) are also quoted at the 68% level unless otherwise stated and are central credible intervals. We adopt this terminology to remind readers of our Bayesian interpretation of the probability density functions, and that we obtained the intervals / regions are estimated by integrating the posterior probability densities. (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 ( $\rho_{crit}$ ) at the redshift of  $z = 0.87$ .

## 2 DATA

We gathered and analyzed 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 estimates from J13. See Table 1 for descriptions of the probability density functions (PDFs) of the input variables.

In order to further 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 (0.56 Mpc in length) among all radio source, was identified to be associated with the merger. (See Fig. 1)

### 3 METHOD – MONTE CARLO SIMULATION

We used the collisionless dark-matter-only Monte Carlo modeling code written by Dawson (2013), hereafter (D13), to compute the physics of between the first and second core-passage of the DM subclusters. In the D13 code, the time evolution of the head-on merger was computed based on an analytical, dissipationless model assuming that the only dominant force is the gravitational attraction from the masses of two truncated Naverro-Frenk-White (hereafter NFW) DM halos. In the simulation, many realizations of the collision is computed by drawing random realizations of the probability density functions (PDFs) of the inputs, including the data ( $\vec{D}$ ) and one model variable, the projection angle between the plane of the sky and the merger axis,  $\alpha$ . In particular, the required data, included the masses ( $M_{200_{NW}}, M_{200_{SE}}$ ) the redshifts ( $z_{NW}, z_{SE}$ ) and the projected separation of the two subclusters ( $d_{proj}$ ). (See Table 1 for quantitative descriptions of the sample PDFs) Each set of inputs is 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. To ensure convergence of the output PDFs, in total, 2 million realizations were computed. The results, however, are consistent up to a fraction of a percent just from 20 000 runs (D13). The random sampling allows us to thoroughly explore the multidimensional input parameter space and account for the uncertainties of the inputs at the same time.

We adapt a Bayesian interpretation of the PDFs 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 ( $\vec{\theta}$ ). For example, the calculation of the output variables of the  $j$ -th realization can be denoted as:

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

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 since we discarded all the variables of the problematic realizations. The system of El Gordo satisfies several major assumptions in the Monte Carlo simulation. One of the strongest assumptions is that all realizations correspond to gravitationally bound systems. The simulation excludes all realizations that would result in relative collisional velocities of the subclusters higher than the free-fall velocity. We justify our assumption of only

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

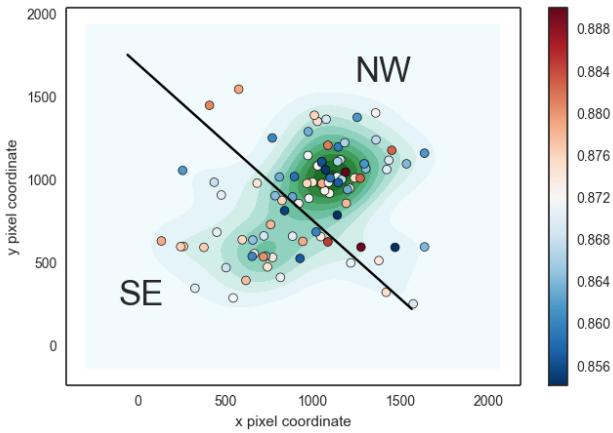
Data	Units	Location	Scale	Ref
$M_{200_{NW}}$	$10^{14} h_{70}^{-1} M_\odot$	13.0	1.6	J13
$c_{NW}$		2.50	0.02	J13
$M_{200_{SE}}$	$10^{14} h_{70}^{-1} M_\odot$	7.6	1.2	J13
$c_{SE}$		2.70	0.04	J13
$z_{NW}$		0.86842	0.00109	M12, Sifón 2013
$z_{SE}$		0.87110	0.00117	M12, Sifón 2013
$d_{proj}$	Mpc	0.74	0.007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

modeling gravitationally bound system by noting that the relative escape velocity of the subclusters for El Gordo is  $4500 \text{ km s}^{-1}$  (based on the mass estimates of Jee et al. (2014)). Studies from cosmological simulations giving the PDFs of the pairwise velocities of massive merging clusters ( $> 10^{15} M_\odot$ ) indicate that it is highly unlikely that the pairwise velocities would  $> 3000 \text{ km s}^{-1}$  under  $\Lambda\text{CDM}$ . (Thompson & Nagamine 2012, Lee & Komatsu 2010). Other major assumptions for modeling systems with this code include negligible impact parameter. Even though there are studies indicating that the impact parameter of El Gordo may be as large as  $300 \text{ kpc} \approx 40\% r_s$  (Molnar & Broadhurst 2014), where  $r_s$  is the characteristic core radius of the NFW halo with the mass of the SE subcluster. According to Ricker (1998), the resulting remnant of bimodal cluster mergers would only be different when the impact parameter  $> 10 r_{core}$ . Mastropietro & Burkert (2008) also reported that an impact parameter of  $0.1 r_{200}$  affected merger dynamics only at the  $\sim 10\%$  level. Other assumptions in this simulation include negligible dynamical friction during the merger, negligible mass accretion and negligible self-interaction of dark matter. Discussion of the effects of each of these assumptions are included in D13.

### 3.1 Inputs of the Monte Carlo simulation

#### 3.1.1 Membership selection and redshift estimation of subclusters

We adopted 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 first converted the coordinates (01:03:22.0, -49:12:32.9) and (01:02:35.1, -49:18:09.8) to pixel space to avoid anarmomorphic distortion. Then we performed a spatial cut using the aforementioned two points in pixel space as the two ends of the cut. The spatial cut is approximately perpendicular to the 2D merger axis and is consistent with the bimodal number density contours (See Figure 2). There are 51 members identified in the NW subclusters and 35 members in the SE subclusters. 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  $1\sigma$  bias-corrected confidence level respectively (Beers



**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.

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 with respect to us, using

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

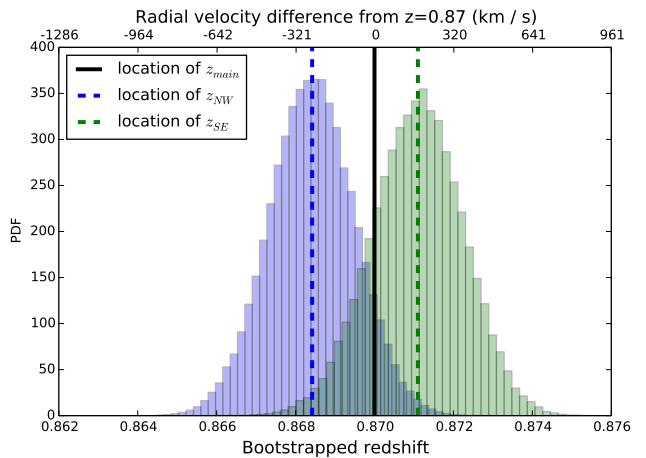
where  $i = 1, 2$  represents the two subclusters, and  $c$  is the speed of light. The radial velocity was calculated by:

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

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  $476 \pm 242$  km s $^{-1}$  (See Fig. 3). The quoted radial velocity differences of 586 km s $^{-1}$  reported by M12 is higher than our estimates but within the 68% bias-corrected confidence interval. Limitations and possible improvements of this analysis of  $v_{rad}$  are provided in the discussion.

### 3.1.2 Weak lensing mass estimation

We obtained 40,000 samples of the joint PDFs of the masses of the two dark matter halos as the outputs of the Monte Carlo Markov Chain (MCMC) procedure from Jee et al. 2014. Discussion of the handling of the weak lensing source galaxies and the details the MCMC procedure for mass estimation can be found in Jee et al. 2014.



**Figure 3.** Bootstrapped location of the redshift estimates and  $v_{rad}$  estimates for each subcluster using the selected spectroscopic members.

### 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 inferred DM centroid locations in Jee et al. 2014. We draw random samples of the location of centroids from two 2D Gaussians centered at RA=01:02:50.601, Decl.=−49:15:04.48 for the NW subcluster and RA = 01:02:56.312, Decl. = −49:16:23.15 for the SE subcluster, with a 1" standard deviation each as estimated from the convergence map of Jee et al. (2014). Or equivalently, the samples of the inferred centroid locations correspond to a projected separation of  $0.74 \pm 0.007$  Mpc.

## 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 8 output variables. Variables of highest interest include the time dependence and the angle  $\alpha$ , which is defined to be the projection angle between the plane of the sky and the merger axis. Other output variables are dependent on  $\alpha$  and the time dependence. Specifically, the simulation denotes the time dependence by providing several characteristic time-scales, 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. Evolution of the merger after the second passage is not considered. Outputs from our dissipationless simulation for a “second” passage will not differ from the first passage.

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 and the 3D separation of the subclusters at observation ( $d_{3D}$ ).

### 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 devised a new prior on the projection angle  $\alpha$  based on the polarization fraction of the radio relic.

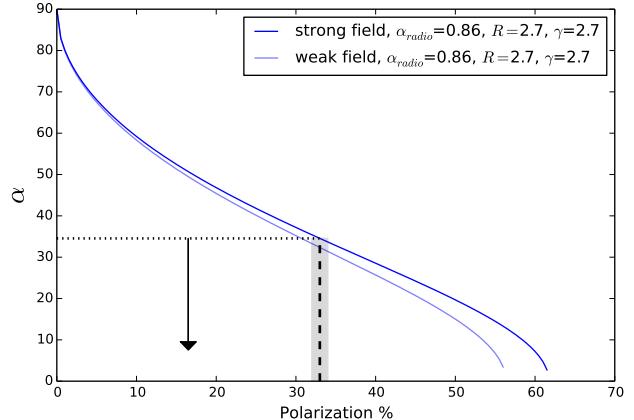
#### 3.3.1 Monte Carlo prior based on the polarization fraction of the radio relic

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, compressing the ICM 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  $\alpha$  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 4). In particular, we refer to a model from Ensslin et al. (1998) that would give the most conservative estimate on the upper bound of  $\alpha$ :

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

This model corresponds to the strong field case with the relic being supported by magnetic pressure only, with the spectral index of the radio emission being  $\alpha_{radio} = 0.86$ , the compression ratio of the magnetic field being  $R = 2.7$  and the spectral index of the electrons being  $\gamma = 2.7$ . 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



**Figure 4.** 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 ( $\alpha_{radio}$ ), spectral index of the electrons ( $\gamma$ ) 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).

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  $\alpha \leq 35^\circ$  because magnetic field nonuniformities can lower the polarization below the Ensslin model value. 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  $\alpha = 0$ . We examined the effects of perturbing 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:

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

We refer to 6 as the polarization prior. Unless otherwise stated, the main results of the paper are obtained after applying this polarization prior in addition to the default priors.

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

Variables	Units	Default priors			Default + polarization priors		
		Location	68% CI <sup>†</sup>	95% CI	Location	68% CI	95% CI
$\alpha$	(degree)	43	19-69	6-80	21	10-30	3-34
$d_{max}$	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 <sup>-1</sup>	580	260-1200	59-2400	940	360-1800	62-2900
$v_{rad}(t_{obs})$	km s <sup>-1</sup>	360	140-630	27-880	310	110-590	8-840
$v_{3D}(t_{col})$	km s <sup>-1</sup>	2800	2400-3700	2100-4200	2400	2200-2800	2100-3500

<sup>†</sup> CI stands for credible interval

### 3.4 Extension to the Monte Carlo simulation - Determining merger scenario with radio relic position by model comparison

One of the biggest questions involving the merger is whether El Gordo was observed to be in a returning or outgoing phase. We came up with an educated guess by making use of the observed projected separation of the relic from the center of mass. Simulations of cluster mergers such as the work of Paul et al. (2011), van Weeren et al. (2011), and Springel & Farrar (2007) showed that, merger shock fronts that may correspond to the radio relics 1) are generated near the center of mass of the subclusters close to the time of the first core-passage, 2) propagate outwards with the shock speed decreasing only slightly. The propagation speed of the shock wave *with respect to the center-of-mass* is reported to drop between 10% to 30% from private communication with Paul S. and  $\sim 10\%$  from Springel & Farrar (2007). To capture the monotonically decreasing trend of the propagation speed of the shock fronts with respect to the center of mass, we expressed the possible shock speeds as a factor of the inferred collisional speed of the corresponding subcluster in the center of mass (momentum) frame. Then we calculated how far the shock would have propagated for our inferred  $TSC_0$  and  $TSC_1$  values. We worked in the center of mass frame where the shock speed is expected to drop slightly with TSC. The projected separation of the shock is approximated as:

$$s_{proj}^j \approx \langle v_{relic} \rangle^j (t_{obs}^j - t_{col}^j) \cos(\alpha^j), \quad (7)$$

where the superscript  $j$  of any variable denotes that the value of the variable from the  $j$ -th realization of the simulation, and  $s_{proj}$  is the estimated projected separation. We estimated the upper and lower bounds of the time-averaged velocity  $\langle v_{NWrelic} \rangle$  of the shock between the collision of the subclusters and the observed time as:

$$\langle v_{NWrelic} \rangle^j = \beta v_{3D,NW}^j(t_{col}) \quad (8)$$

$$= \beta v_{3D}^j(t_{col}) \frac{m_{SE}^j}{m_{SE}^j + m_{NW}^j}, \quad (9)$$

where  $0.7 \leq \beta \leq 1.3$  is a factor that we introduce to represent the uncertainty of the velocity of the relic shockwave and  $v_{3D,NW}(t_{col})$  refers to the collisional velocity of the NW subcluster in the center-of-mass frame as a comparison. Likewise, we have also computed the expected projected separation of the SE relic using:

$$\langle v_{SErelic} \rangle^j = \beta v_{3D}^j(t_{col}) \frac{m_{NW}^j}{m_{SE}^j + m_{NW}^j}. \quad (10)$$

Simulation of the Bullet Cluster by Springel & Farrar (2007) indicates that the propagation velocity of the shock evolves such that  $\beta \approx 0.95$  within  $\sim 0.4$  Gyr after the collision. For the analysis of El Gordo,  $\beta \approx 0.9$  would be the most likely value given that the  $TSC$  of El Gordo is longer. The approximations in equations 8 and 10 reflect the physics that the shock is driven by the merger.

## 4 RESULTS

We found that the two subclusters collided with a relative velocity of  $2400 \pm^{900}_{400}$  km s<sup>-1</sup>, at an estimated projection angle of  $\alpha = 21^\circ \pm^{9}_{11}$ . From our analysis of the two scenarios, we found that El Gordo is more likely to be observed at a returning phase with a estimate of  $TSC_1 = 0.91 \pm^{0.22}_{0.39}$  Gyr. We present an overview of all the estimated variables in table 2, with results only applying the default priors on the left hand side of the table and those applied with the polarization prior on the right hand side. Furthermore, we include the plots of all the marginalized PDFs with the polarization prior in Appendix B.

Our estimates of  $v_{3D}(t_{col}) = 2400 \pm^{900}_{400}$  km s<sup>-1</sup> at the time of collision is compatible with independent estimates from Lindner et al. 2014. By making use of the Mach number of the NW radio relic, Lindner et al. (2014) reported an estimate of the upper bound of the relative collisional velocity to be  $2500 \pm^{400}_{300}$  km s<sup>-1</sup>. From the simulation of the Bullet Cluster, Springel & Farrar (2007) showed that there could be an induced flow of gas at a velocity as high as  $\sim 1100$  km s<sup>-1</sup> from the outer edge of the subclusters to the center during a cluster merger. Therefore, it is reasonable to assume the reported speed from Lindner et al. (2014) to be an estimate of the *relative collisional speed between the subclusters*, instead of the collision speed *relative to the center-of-mass frame* of the subcluster. Magnitude of the relative  $v_{3D}$  of the subclusters dropped as the subclusters climbed out of the gravitational potential of each other, and reduced to a relative  $v_{3D}(t_{obs}) = 940$  km s<sup>-1</sup> at the time of observation.

#### 4.1 Time-since-collision (TSC) and the merger scenario

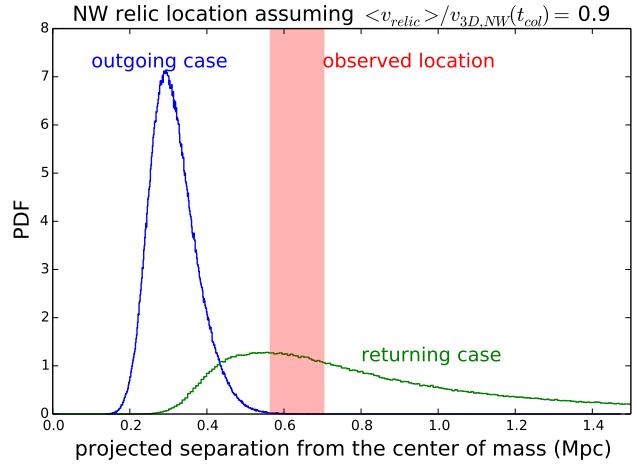
The simulation gives two estimates for the time-since-collision, with  $TSC_0 = 0.46 \pm 0.16$  Gyr and  $TSC_1 = 0.91 \pm 0.22$  Gyr. Both the estimates of  $TSC_0$  and  $TSC_1$  are compatible physical time-scales of observable features of El Gordo. Both estimates are lower than the approximate observable time-scale of the wake feature in the X-ray, i.e. the sound crossing time of  $\sim 2$  Gyr. The observable time scale of the radio relics is also on the scale of  $\sim 1$  Gyr. Based on section 3.4, we present the most likely value of  $\beta = 0.9$  in Fig. 5 to show that the returning case is preferred. This conclusion favoring the returning case holds true for  $\beta < 1.1$ , which corresponds to  $\langle v_{NW\text{relic}} \rangle < 1000 \text{ km s}^{-1}$  and  $\langle v_{SE\text{relic}} \rangle < 1800 \text{ km s}^{-1}$  in the center of mass frame. For comparison purpose, we found that an extreme, and unlikely range of  $\beta > 1.5$  would be needed for the outgoing scenario to be preferred. (See appendix C for plots of all the range of  $\beta$  that we examined).

When we assume  $\langle v_{\text{relic}} \rangle / v_{3D,1}(t_{\text{col}}) = 1.5$  (uppermost panel of Figure C1),  $\langle v_{\text{relic}} \rangle$  corresponds to an time-averaged velocity of  $\sim 3800 \text{ km s}^{-1}$  relative to SE subcluster, or, in the center of mass frame  $\sim 1400 \text{ km s}^{-1}$ , the outgoing scenario is more favored. As we examine a decreased ratio of  $\langle v_{\text{relic}} \rangle / v_{3D,1}(t_{\text{col}})$ , we probe how much the shock could have slowed down and still be consistent with the outgoing scenario. The second (relative velocity of  $2800 \text{ km s}^{-1}$ ) and third row of Fig. C1 show that if  $\langle v_{\text{relic}} \rangle \lesssim 1.1$ , the returning scenario would be favored instead.

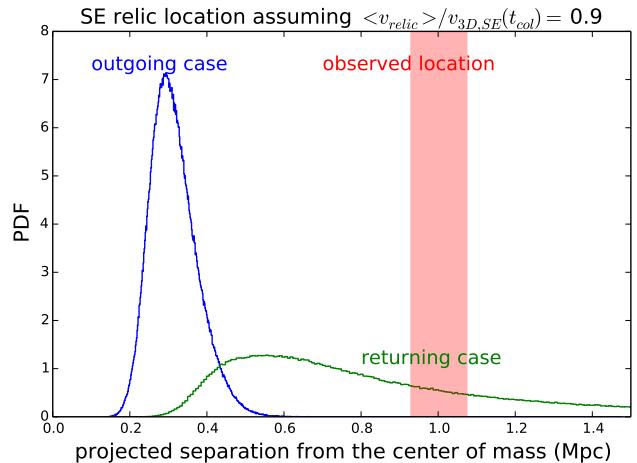
Finally, we note that the estimate of NW shock velocity at  $2500 \pm 400 \text{ km s}^{-1}$  by Lindner et al. (2014) was reported with respect to the turbulent ICM. It is unclear what the bulk motion of the ICM would be at the time of observation of El Gordo and how to ensure the frame of references of our simulation outputs and the observation are consistent for comparison. If we marginalize  $\beta$  to compute a Bayes factor to compare the ratio between the likelihood of the returning case and the outgoing case, we get a Bayes factor to be  $\approx 6$  from the NW relic and a Bayes factor of  $\approx 1700$  for the SE relic, showing substantial evidence to support the returning scenario (See Appendix C). A definite conclusion about the best fit scenario would require better knowledge of the formation location of the shock and the time evolution of the shock propagation velocity.

#### 4.2 Sensitivity analysis of the polarization prior

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



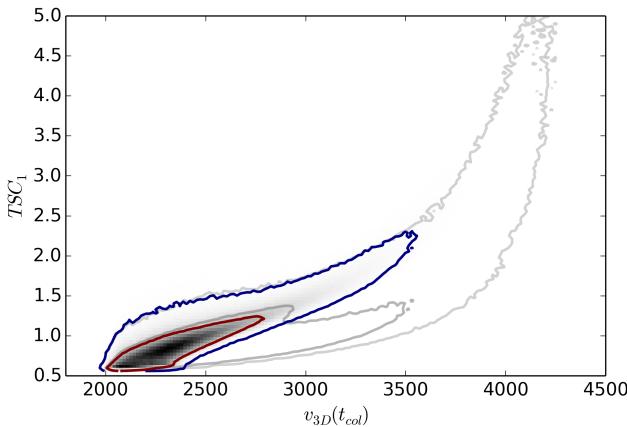
**Figure 5.** Comparison of the PDFs of the observed position of the NW relic (red bar includes the 95% confidence interval of location of the NW radio relic in the center of mass frame) with the predicted position from the two simulated merger scenarios (blue for outgoing and green for the returning scenario). We made use of the polarization prior for producing this figure.



**Figure 6.** Comparison of the PDFs of the observed position of the SE relic (red bar includes the 95% confidence interval of location of the radio relic in the center of mass frame)

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

The projection angle of the El Gordo is constrained by the 95% confidence interval of  $3^\circ$  and  $34^\circ$ , with an estimate of  $21^\circ$ . Many previous studies (van Weeren et al. CITATIONS) have suggested that  $\alpha$  should be small for the detection of double radio relics to be possible but did not provide quantitative constraints. Without the polarization priors, the Monte Carlo simulation gives the estimate of the projection angle as  $41.7^\circ$ , with the 95% CI =  $22.7^\circ$ ,  $61.14^\circ$ . After applying the radio relic polarization prior, the CI shrinks to  $^\circ$  and  $^\circ$ .



**Figure 7.** The marginalized output PDF of the returning time-since-collision ( $TSC_1$ ) 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.

## 5 DISCUSSION

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

We outline the qualitative agreement and disagreement between our simulations and other hydrodynamical simulations of El Gordo such as Donnert (2014) and Molnar & Broadhurst (2014). Our simulation focuses on giving PDF estimates of particular dynamical and kinematic variables, whereas the hydrodynamical simulations focused on understanding the detailed gas dynamics and the corresponding X-ray observables of El Gordo. The goals, assumptions, and initial conditions of Donnert (2014) and Molnar & Broadhurst (2014) are different from our simulations such that it is hard to come up with a fair comparison. For example, both hydrodynamical simulations were based on a few sets of initial conditions, instead of a thorough sampling of the inputs. The outputs of the hydrodynamical simulations are unable to convey uncertainties due to the inputs. Both hydrodynamical simulations made use of a higher  $m_{NW} = 1.9 \times 10^{15} M_\odot$  which is larger than the upper 95% CI of  $m_{NW}$  that we used. Furthermore, Molnar & Broadhurst (2014) initialized the relative infall velocity (velocity when separation of subclusters equal the sum of the two virial radii) to be  $> 2250 \text{ km s}^{-1}$ . This corresponds to  $v_{3D}(t_{col}) \gtrsim 4700 \text{ km s}^{-1}$ , which is close to the escape velocity of the subclusters. The range of projection angles suggested by Molnar & Broadhurst (2014) of  $\alpha \gtrsim 45^\circ$  are also excluded by our polarization prior. On the other hand, our simulations show qualitative agreement with Donnert (2014) on the duration between consecutive core-passages. With the time resolution of 0.25 Gyr, Donnert (2014) gave an qualitative estimate of  $T \approx 2$  Gyr between the first and second core-passage in Fig. 6 of their work, while our estimate gives  $T = 1.6$  Gyr. A better comparison to the hydrodynamical simulations may be achieved by performing Gaussian process emulation to obtain outputs with more similar inputs

as our simulations, but the work of fair, detailed comparison is out of the scope of this paper.

### 5.2 Comparison to other merging clusters of galaxies

Dynamical properties of El Gordo are more similar to the Bullet Cluster (Bradač et al. 2006, Springel & Farrar 2007, Mastropietro & Burkert 2008) than the Musketball Cluster (Dawson et al. 2012, D13). Both the Bullet Cluster and the Musketball were studied by D13 using the Monte Carlo simulation. The three output variables that we compare among the listed clusters include  $TSC$ ,  $v_{3D}$  and the masses. These variables can affect how strongly the self-interaction of dark matter would manifest itself in case the DM self-interaction cross section is not zero. A higher  $TSC$  might allow more time for DM to interact with itself; a higher  $v_{3D}$  might and a higher mass would correspond to higher central density for a higher self-interaction rate. Although El Gordo is the most massive among the listed clusters, the higher mass did not result in a higher  $v_{3D}(t_{col})$ . The estimate of  $v_{3D}(t_{col}) = 2400 \text{ km s}^{-1}$  is comparable to the inferred  $v_{3D}(t_{col}) = 2800 \text{ km s}^{-1}$  of the Bullet Cluster. Both the phase  $TSC_0 \approx 40\% T$  and the  $TSC_0 = 0.6$  Gyr are similar between El Gordo and the Bullet Cluster. It is remained to be investigated if our hypothesis about the relationship between how these variables would affect our chance of seeing effects from self-interaction of dark matter.

### 5.3 Limitations of our data analysis and future work

We note that our method of inferring the 3D velocities by deprojecting the radial velocities might not be optimal but is commonly used (M12, Dawson et al. 2012). The relative radial velocities are dependent on the identification of subcluster members. However, it is unclear that a spatial cut according to number density peaks would be optimal for selecting galaxies that represent the underlying velocities of the subclusters. Traditionally, grouping members may be treated as a (statistical) clustering problem with well known methods such as normal mixture model. However, our setting is complicated by the fact that each of the observed redshifts of galaxies is an unknown combination of cosmological redshift and redshift due to peculiar motion along the line-of-sight. Our study only made use of a limited number of galaxies with spectroscopy information. This could cause selection bias and incompleteness due to limitations of the of spectroscopic targets. Existing literature and tests mostly focus on detecting substructures in galaxy clusters through non-Gaussianity of the velocity dispersion (Dressler & Shectman 1988, Einasto et al. 2012). Few extensive studies have focused on finding out the uncertainties associated with this clustering procedure for inferring the dynamics of merging clusters. This problem of inferring galaxies to best reflect subcluster dynamics presents an interesting direction for investigation in the future.

Despite the limitations and simplifying assumptions, this work provides insight to the merger dynamics and configuration of El Gordo as a basis for further studies. Tests for signs of self-interaction of DM in El Gordo can be done by

comparing the offset between the member galaxies and the DM distribution of the subclusters. By understanding the merging scenario, we can infer if the observed galaxy-DM offset are in the expected direction such that the collisionless galaxies lead the DM subhalos along the direction of motion. In addition, our work has allowed us to examine what information would be needed to better understanding of the merger dynamics. Important pending questions concerning the radio relic and the merger include: What are the typical propagation velocities of the shockwave that corresponds to the radio relic in the center of mass frame of the cluster? What physical properties of the DM subclusters would correlate the best with the time-evolution of the propagation velocity of the shockwave? What are the duration after the merger for which radio relics are observable in terms of the merger core-passage time-scales? We urge simulators working on radio relic simulations to help come up with answers that would be directly comparable with observed data.

As a summary, this paper presents one of the first examples of using the observed radio relic emission to constrain cluster merger properties, the results

(i) favors an estimate of a small projection angle of  $21^\circ \pm 11^\circ$  (See Fig. B3)

(ii) shows a *relative* collisional velocity between the subclusters of El Gordo as  $2400 \text{ km s}^{-1}$

(iii) favors a returning scenario if the  $\langle v_{NW\text{relic}} \rangle \leq 1000 \text{ km s}^{-1}$  and  $\langle v_{SE\text{relic}} \rangle \leq 1800 \text{ km s}^{-1}$  between the collision and the observation in the center-of-mass frame. It takes an unlikely high speed of  $\langle v_{\text{relic}} \rangle \gg 1.5 v_{3D,\text{sub}}(t_{\text{col}})$  for the outgoing scenario to be favored.

## 6 ACKNOWLEDGEMENTS

We thank Franco Vazza, Marcus Brüggen and Surajit Paul 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. KN is grateful to Paul Baines and Tom Loredo for discussion of the use of Bayesian priors and sensitivity tests. Part of this work was performed under the auspices of the U.S. DOE by LLNL under Contract DE-AC52-07NA27344.

## REFERENCES

- Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32, doi:10.1086/115487
- Bradač M. et al., 2006, ApJ, 652, 937, 0608408, doi:10.1086/508601
- 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
- Dressler A., Shectman S. A., 1988, AJ, 95, 985, doi:10.1086/114694
- Einasto M. et al., 2012, A&A, 540, A123, doi:10.1051/0004-6361/201118697
- 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, aapr, 20, 54, 1205.1919, doi:10.1007/s00159-012-0054-z
- 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
- Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2013, MNRAS, 437, 2865, 1308.3419, doi:10.1093/mnras/stt2097
- Lee J., Komatsu E., 2010, ApJ, 718, 60, doi:10.1088/0004-637X/718/1/60
- Lindner R. R. et al., 2014, ApJ, 786, 49, 1310.6786, doi:10.1088/0004-637X/786/1/49
- Marriage T. a. et al., 2011, ApJ, 737, 61, doi:10.1088/0004-637X/737/2/61
- Mastropietro C., Burkert A., 2008, MNRAS, 389, 967, doi:10.1111/j.1365-2966.2008.13626.x
- 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
- Molnar S. M., Broadhurst T., 2014, 1, 9, 1405.2617
- Paul S., Iapichino L., Miniati F., Bagchi J., Mannheim K., 2011, ApJ, 726, 17, doi:10.1088/0004-637X/726/1/17
- Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173, doi:10.1086/587859
- Ricker P. M., 1998, ApJ, 496, 670, doi:10.1086/305393
- 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
- Thompson R., Nagamine K., 2012, MNRAS, 419, 3560, doi:10.1111/j.1365-2966.2011.20000.x
- van Weeren R. J., Brüggen M., Röttgering H. J. A., Hoeft M., 2011, MNRAS, 418, 230, 1108.1398v1, doi:10.1111/j.1365-2966.2011.19478.x
- van Weeren R. J., Röttgering H. J. a., Brüggen M., Hoeft M., 2010, Science, 330, 347, doi:10.1126/science.1194293
- Vazza F., Brüggen M., van Weeren R., Bonafede A., Dolag K., Brunetti G., 2012, MNRAS, 421, 1868, 1111.1720, doi:10.1111/j.1365-2966.2011.20160.x
- Zitrin A., Menanteau F., Hughes J. P., Coe D., Barrientos L. F., Infante L., Mandelbaum R., 2013, ApJ, 770, L15, arXiv:1304.0455v1, doi:10.1088/2041-8205/770/1/L15

## APPENDIX A: DEFAULT 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:

$$P(v_{3D}(t_{col})|\alpha, v_{proj}(t_{col})) = 0 \text{ if } v_{3D}(t_{col}) > v_{\text{free fall}}. \quad (\text{A1})$$

$$P(TSC_0) = \begin{cases} \text{const if } TSC_0 < \text{age of universe at } z = 0.87 \\ 0 \text{ otherwise.} \end{cases} \quad (\text{A2})$$

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  $TSC_1$ .

$$P(TSC_1) = \begin{cases} \text{const if } TSC_1 < \text{age of universe at } z = 0.87 \\ 0 \text{ otherwise.} \end{cases} \quad (\text{A3})$$

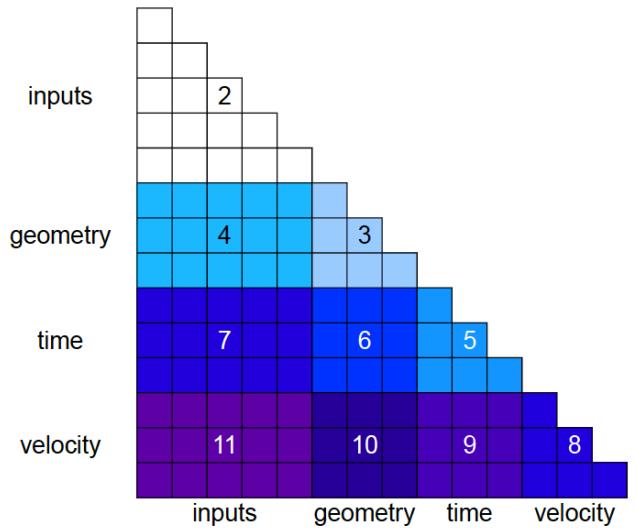
To correct for observational limitations, we further convolve the posterior probabilities of the different realizations with

$$P(TSC_0|T) = 2 \frac{TSC_0}{T}, \quad (\text{A4})$$

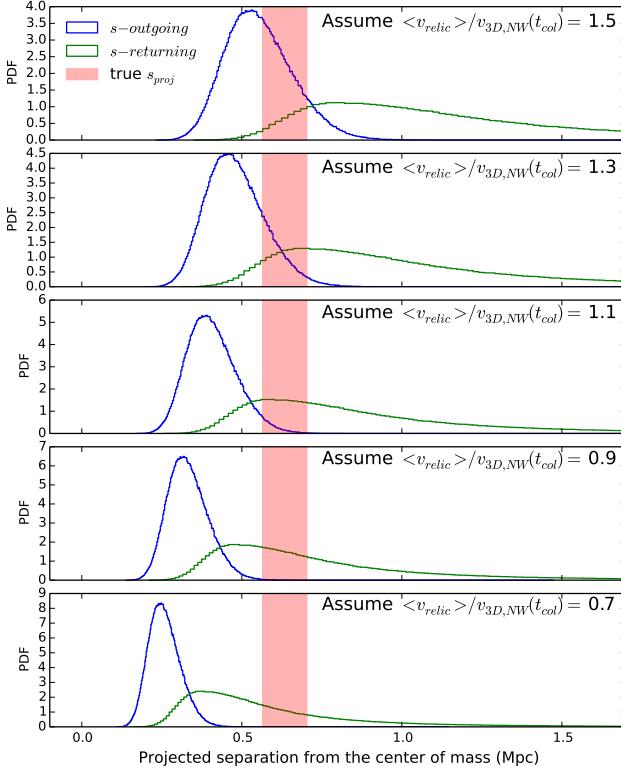
to account for how the subclusters move faster at lower  $TSC$  and thus it is more probable to observe the subclusters at a stage with a larger  $TSC$ .

## APPENDIX B: PLOTS OF OUTPUTS OF THE MONTE CARLO SIMULATION

We present the PDFs of the inputs of the dynamical simulation and the marginalized PDFs of the outputs after applying the polarization prior in addition to the default priors. See following figure for explanations of the order that we present the variables.



**Figure B1.** Matrix of variables used in the simulations. We present them in 4 categories, including, inputs, geometry, time and velocity. Regions of the same color represent one plot and the number indicates the corresponding figure number in this appendix.



**Figure C1.** Comparison of the PDFs of the observed position of the NW relic (red bar includes 95% confidence interval of location of relic in the center of mass frame) with the predicted position from the two simulated merger scenarios (blue for outgoing and green for the returning scenario). For the most likely value of  $\beta < 1.1$ , the returning scenario is preferred. For comparison purpose, we also show that the shock velocity has to be as extreme as  $1.5 v_{3D,NW}(t_{col})$  (top panel) for the outgoing scenario to be favored. Note that we made use of the polarization prior for producing this figure.

### APPENDIX C: MODEL COMPARISON FOR THE OUTGOING AND RETURNING SCENARIO

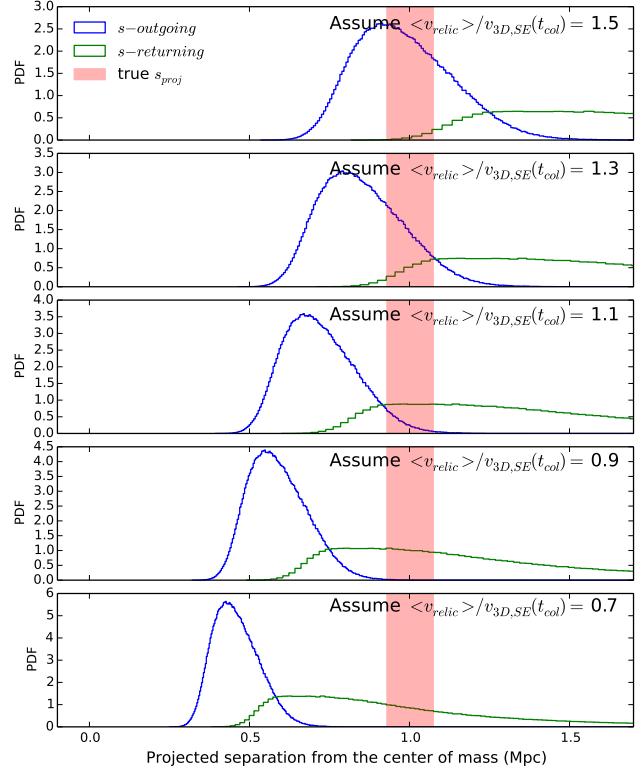
We can write down our model for the different scenarios to be  $M_{out}$  and  $M_{ret}$  for the outgoing and returning scenarios respectively. Both models are computed using the same variables except for TSC. The models are parametrized by  $\beta$ , which represents our uncertainty of  $\langle v_{relic} \rangle$  (See equation 7). Our data in this case is the observed projected location of the radio relic  $s_{proj}$ . The Bayes factor thus can be computed as:

$$K = \frac{P(s_{proj}|M_{ret})}{P(s_{proj}|M_{out})} \quad (\text{C1})$$

$$= \frac{\int P(\beta|M_{ret})P(s_{proj}|\beta, M_{ret})d\beta}{\int P(\beta|M_{out})P(s_{proj}|\beta, M_{out})d\beta} \quad (\text{C2})$$

$$\approx \begin{cases} 6 & \text{for the NW relic} \\ 1700 & \text{for the SE relic} \end{cases} \quad (\text{C3})$$

where  $P(s_{proj}|\beta, M_{ret})$  and  $P(s_{proj}|\beta, M_{out})$  can be viewed as the green and blue curves respectively for each given  $\beta$



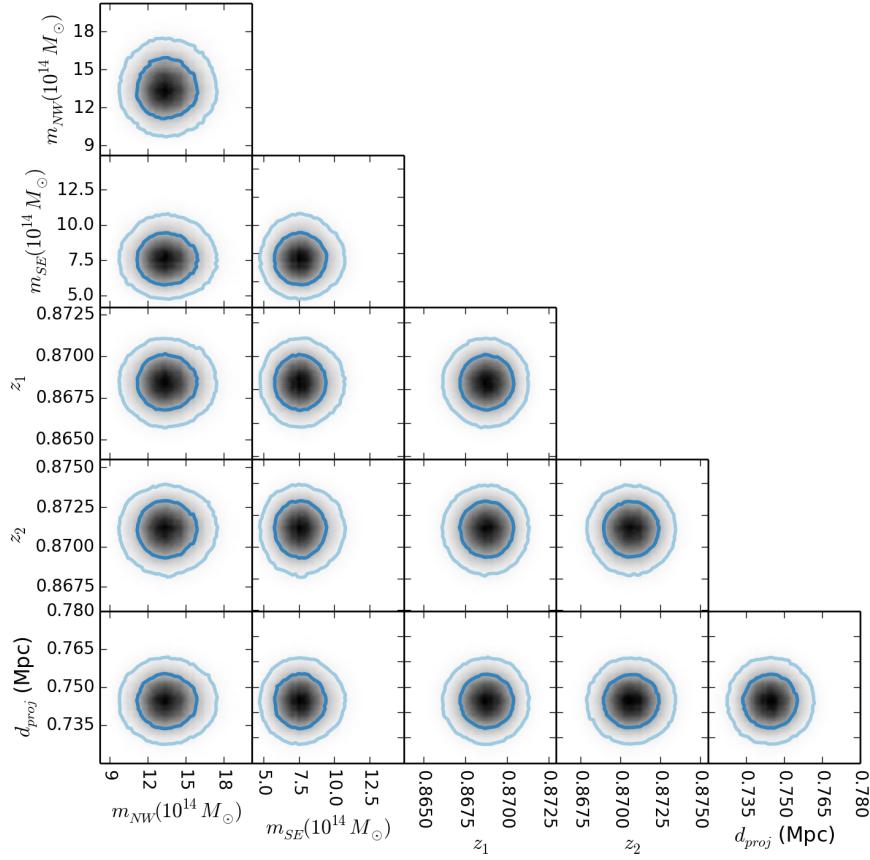
**Figure C2.** Comparison of the PDFs of the observed position of the SE relic (red bar includes 95% confidence interval of location of relic in the center of mass frame) with the predicted position from the two simulated merger scenarios (blue for outgoing and green for the returning scenario). For the most likely value of  $\beta < 1.1$ , the returning scenario is preferred. We obtained similar conclusion about the merger scenario as for the NW relic calculation.

value and we have used priors on  $\beta$  as:

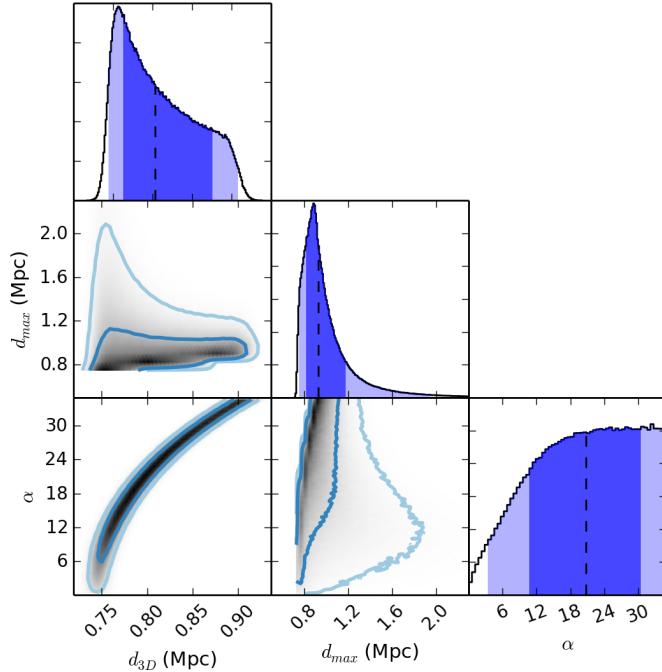
$$P(\beta|M_{ret}) = P(\beta|M_{out}) = \begin{cases} \text{const if } 0.7 \leq \beta \leq 1.3 \\ 0 \text{ otherwise.} \end{cases} \quad (\text{C4})$$

This conservative prior is chosen to reflect our belief that the NW shockfront should propagate at approximately  $v_{3D,NW}(t_{col})$  in the center-of-mass frame, and likewise for the SE shockfront. As more simulations of merger shocks become available, one can update equation C4 accordingly.

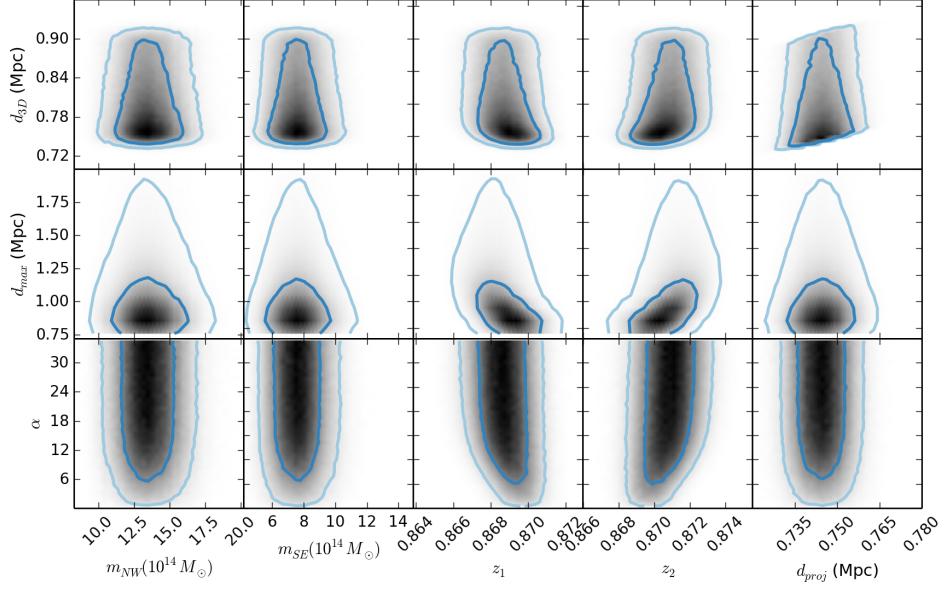
This paper has been typeset from a TeX/ L<sup>A</sup>T<sub>E</sub>X file prepared by the author.



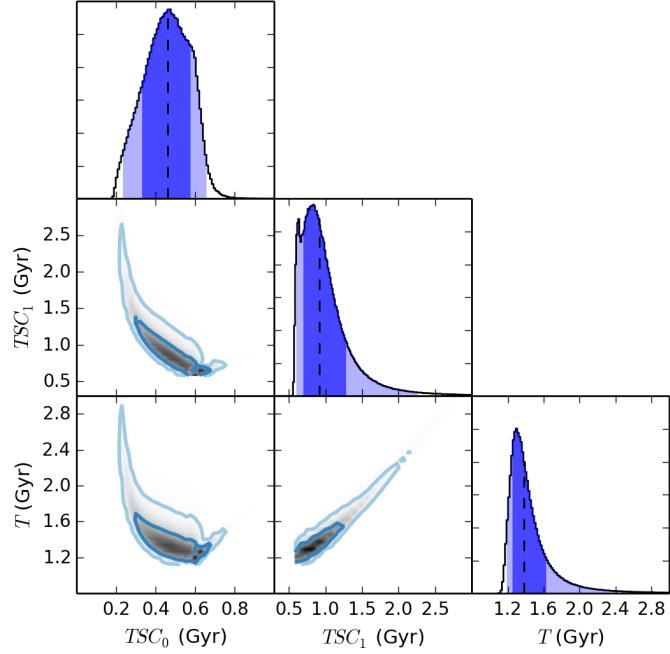
**Figure B2.** Marginalized PDFs of original inputs (vertical axis) and the inputs after applying polarization prior and default priors (horizontal axis). The inner and outer contour denote the central 68% and 95% credible regions respectively. The circular contours show that the application of priors did not introduce uneven sampling of inputs.



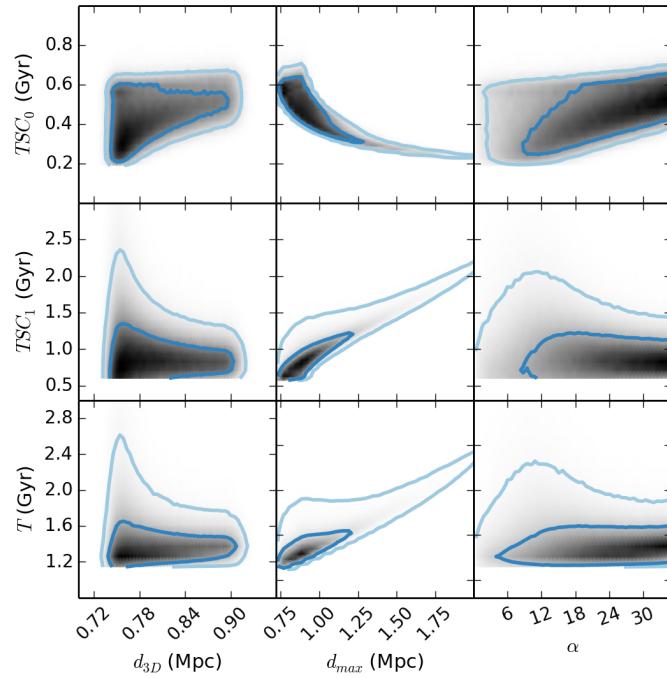
**Figure B3.** One-dimensional marginalized PDFs (panels on the diagonal) and two-dimensional marginalized PDFs of variables denoting characteristic distances and projection angle of the mergers.



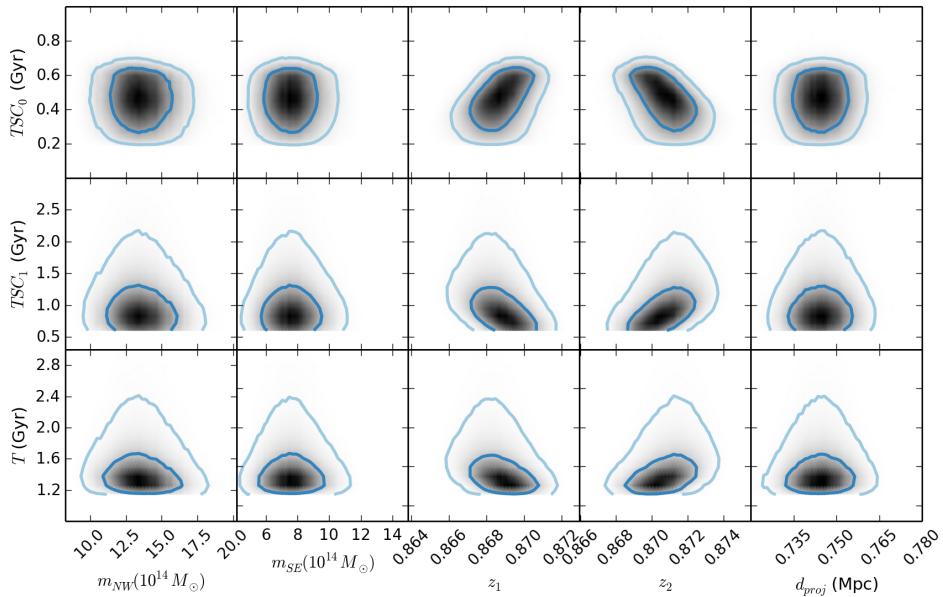
**Figure B4.** Marginalized PDFs of characteristic distances and projection angle of the merger and the inputs of the simulation.



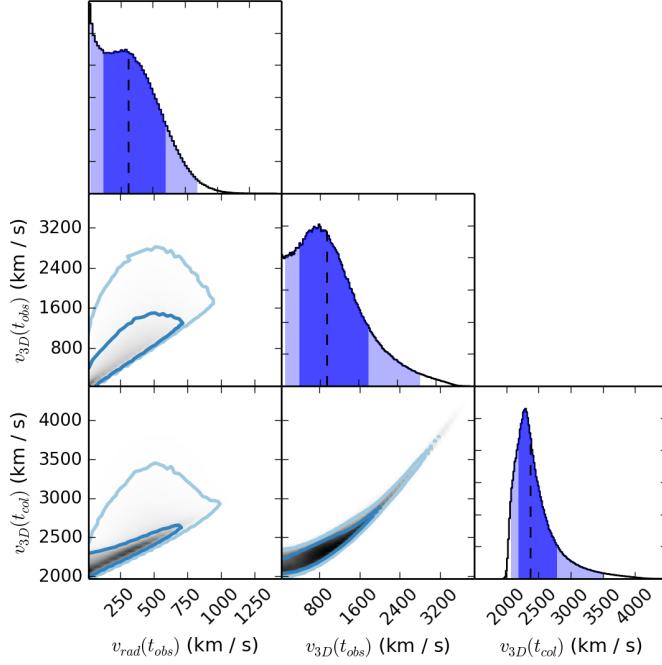
**Figure B5.** One-dimensional PDFs of characteristic timescales of the simulation (panels on the diagonal) and the marginalized PDFs of different timescales.



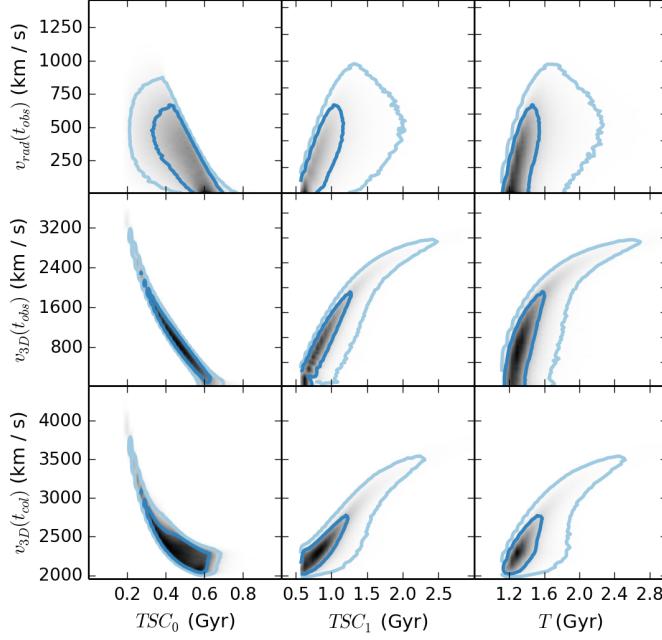
**Figure B6.** Marginalized PDFs of characteristic timescales of the simulation and the characteristic distances and the projection angle of the merger.



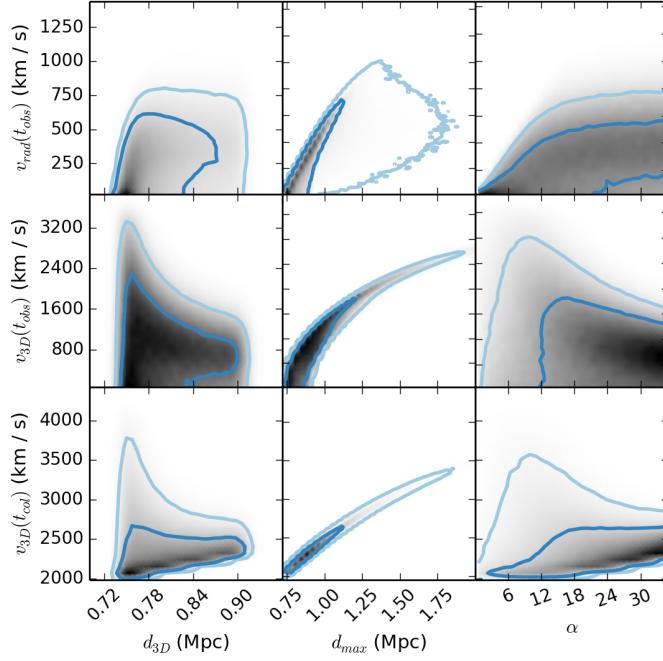
**Figure B7.** Marginalized PDFs of characteristic timescales of the simulation and the inputs.



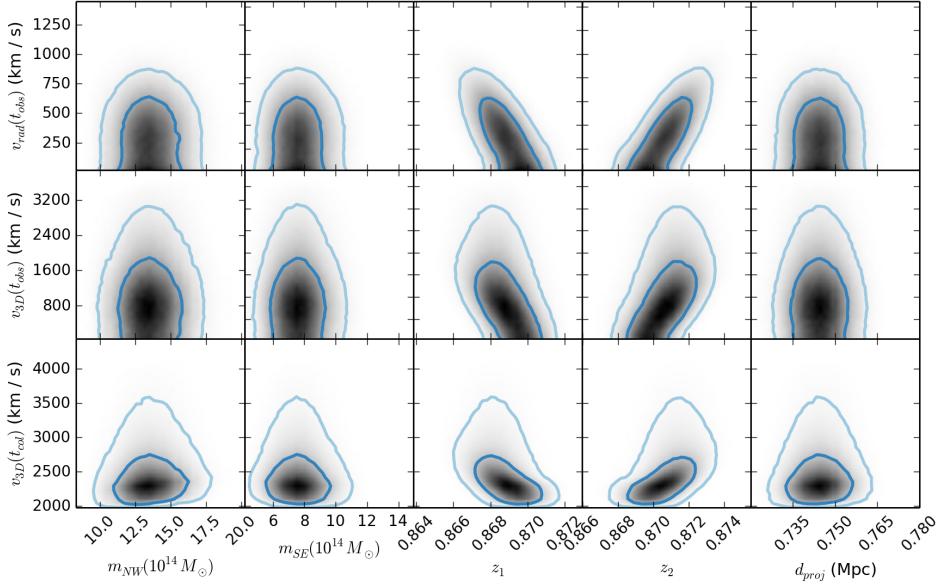
**Figure B8.** One-dimensional marginalized PDFs of velocities at characteristic times (panels on the diagonal) and marginalized PDFs of different velocities.



**Figure B9.** Marginalized PDFs velocities and the characteristic timescales of the simulation against the inputs.



**Figure B10.** Marginalized PDFs of the velocities at characteristic timescales and the characteristic distances and the projection angle of the merger.



**Figure B11.** Marginalized PDFs of relative velocities characteristic timescales of the simulation and the inputs.