



**HAL**  
open science

## **A discrete chemo-dynamical model of M87's globular clusters: Kinematics extending to $\sim 400$ kpc**

Chao Li, Ling Zhu, R. J. Long, Shude Mao, Eric W. Peng, Marc Sarzi, Glenn van de Ven, Hongxin Zhang, Rui Guo, Xiangxiang Xue, et al.

► **To cite this version:**

Chao Li, Ling Zhu, R. J. Long, Shude Mao, Eric W. Peng, et al.. A discrete chemo-dynamical model of M87's globular clusters: Kinematics extending to  $\sim 400$  kpc. Monthly Notices of the Royal Astronomical Society, 2020, 492, pp.2775-2795. 10.1093/mnras/staa027 . insu-03667131

**HAL Id: insu-03667131**

**<https://insu.hal.science/insu-03667131>**

Submitted on 21 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A discrete chemo-dynamical model of M87's globular clusters: Kinematics extending to $\sim 400$ kpc

Chao Li,<sup>1,2★</sup> Ling Zhu<sup>1b, 3★</sup>, R. J. Long,<sup>1,4,5</sup> Shude Mao,<sup>1,4</sup> Eric W. Peng,<sup>6,7</sup> Marc Sarzi,<sup>8,9</sup> Glenn van de Ven<sup>1b, 10</sup>, Hongxin Zhang,<sup>11</sup> Rui Guo,<sup>1,2</sup> Xiangxiang Xue,<sup>1</sup> Alessia Longobardi,<sup>12</sup> Patrick Côté,<sup>13</sup> Laura Ferrarese,<sup>13</sup> Chengze Liu,<sup>14,15</sup> Stephen Gwyn,<sup>13</sup> Sungsoo Lim<sup>7</sup> and Youkyung Ko<sup>6,7,16</sup>

*Affiliations are listed at the end of the paper*

Accepted 2019 December 23. Received 2019 December 19; in original form 2019 January 30

## ABSTRACT

We study the mass distribution and kinematics of the giant elliptical galaxy M87 (NGC 4486) using discrete chemo-dynamical, axisymmetric Jeans equation modelling. Our catalogue comprises 894 globular clusters (GCs) extending to a projected radius of  $\sim 430$  kpc with line-of-sight velocities and colours, and Multi Unit Spectroscopic Explorer integral field unit data within the central 2.4 kpc of the main galaxy. The gravitational potential for our models is a combination of a luminous matter potential with a varying mass-to-light ratio for the main galaxy, a supermassive black hole and a dark matter (DM) potential with a cusped or cored DM halo. The best-fitting models with either a cusped or a cored DM halo show no significant differences and both are acceptable. We obtain a total mass of  $(2.16 \pm 0.38) \times 10^{13} M_{\odot}$  within  $\sim 400$  kpc. By including the stellar mass-to-light ratio gradient, the DM fraction increases from  $\sim 26$  per cent (with no gradient) to  $\sim 73$  per cent within  $1 R_e^{\text{maj}}$  (major axis of half-light isophote, 14.2 kpc), and from  $\sim 84$  per cent to  $\sim 94$  per cent within  $5 R_e^{\text{maj}}$  (71.2 kpc). Red GCs have moderate rotation with  $V_{\text{max}}/\sigma \sim 0.4$ , and blue GCs have weak rotation with  $V_{\text{max}}/\sigma \sim 0.1$ . Red GCs have tangential velocity dispersion anisotropy, while blue GCs are consistent with being nearly isotropic. Our results suggest that red GCs are more likely to be born *in situ*, while blue GCs are more likely to be accreted.

**Key words:** galaxies: individual: M87 – galaxies: kinematics and dynamics.

## 1 INTRODUCTION

In the outer haloes of galaxies, the observed stellar surface brightness is too low for stellar kinematics to be determined. Bright globular clusters (GCs), considered as discrete point sources, are powerful tracers of dark matter (DM) distributions, and fossil records of the halo's formation history.

M87 is a giant elliptical galaxy, located in the dynamical centre of the Virgo cluster. It has the largest GC system in the local supercluster with a total of 15 000 GCs (Peng et al. 2008), and the number of M87 GCs with line-of-sight (LOS) velocities has been growing rapidly in recent years (Strader et al. 2011; Durrell et al. 2014). M87's GCs have been used previously to estimate its mass profile (e.g. Gebhardt & Thomas 2009; Murphy, Gebhardt & Adams 2011; Strader et al. 2011), but only recently, has its data coverage been significantly increased, reducing the uncertainty of mass profiles

extending to  $\sim 150$  kpc (Zhu et al. 2014). Recent observations of M87's GC LOS velocities from The Next Generation Virgo Cluster Survey (NGVS) have reached the boundary of the galaxy's potential. A population of intracluster GCs, which is gravitationally unbound to M87 but bound to the cluster potential, has been identified with a velocity dispersion of  $\sim 640$  km s<sup>-1</sup> (Longobardi et al. 2018a). GCs bound to M87 have a velocity dispersion of  $\sim 350$  km s<sup>-1</sup>.

M87 has diffuse X-ray emission from hot gas, which has been observed extending to  $\sim 1.2$  Mpc (Urban et al. 2011). A mean temperature of  $kT = 2.3$  keV of the hot gas implies  $M_{200} = 1.4 \times 10^{14} M_{\odot}$  and  $R_{200} = 1.08$  Mpc (Pointecouteau, Arnaud & Pratt 2005; Urban et al. 2011). A detailed mass profile has also been obtained from X-ray emission extending to a few hundred kpc (Nulsen & Bohringer 1995; Das et al. 2010). The virial mass obtained from X-ray emission is generally consistent with the estimates from the kinematics of dwarf galaxies extending to  $\sim 2$  Mpc (McLaughlin 1999). There is evidence that the masses obtained from X-ray gas are larger than those obtained from GC kinematics (e.g. Strader et al. 2011; Zhu et al. 2014) at  $R \gtrsim 100$  kpc,

\* E-mail: lchao@nao.cas.cn (CL), lzhu@shao.ac.cn (LZ)

although these previous mass profiles were obtained from GCs not extending beyond  $\sim 150$  kpc. Due to the fact that Virgo may not be fully relaxed (Binggeli, Tammann & Sandage 1987; Urban et al. 2011), it is possible that neither X-ray gas, dwarf galaxies, nor GCs are fully relaxed at all regions from the inner radii of M87 to the outermost cluster. X-ray gas and dwarf galaxies may be better tracers of the mass profile of the outer regions of the Virgo cluster, while GCs, which are more spatially concentrated, may be a better mass tracer for M87 within its dynamical boundary.

How the outer haloes of giant elliptical galaxies are assembled is still under debate. Cosmological simulations suggest that massive early-type galaxies undergo a two-phase formation process: an early phase of rapid *in situ* star formation driven by gas accretion and dissipative mergers (the so-called wet phase), followed by a dry merger-dominated phase with markedly smaller star formation rates (Oser et al. 2010). These two phases result in stars with different chemical and kinematic properties, which are usually difficult to determine from observations alone.

Velocity dispersion anisotropy is a good indicator of underlying orbit distributions, and is widely used to discriminate between stars from the different galaxy formation phases. Galaxy formation simulations show that accreted stars contribute mainly to isotropic to radial anisotropy, while stars formed *in situ* could result in tangential anisotropy (Röttgers, Naab & Oser 2014; Wu et al. 2014).

The colour distributions within GC systems are typically bimodal, with GCs with different colours being associated with different spatial and kinematic distributions. GC colour bimodality is thought to be related to the two phase formation of galaxies (Côté, Marzke & West 1998; Brodie & Strader 2006). Metal-rich (red) GCs are believed to have been created in the dissipational construction of their parent galaxies, while metal-poor (blue) GCs are considered to have been created in the early universe’s low-mass DM haloes and became part of their current galaxies by accretion.

The relationship, however, between the properties of GCs and their host galaxies is not straightforward to determine: different haloes contain different numbers of GCs, and GCs themselves are potentially subject to disruption by a variety of processes. Considering only GC abundance in different haloes, a purely accreted formation history results in GC systems that are mainly radially biased, with velocity anisotropy that is mildly radial in the centre but is more strongly radial in the outer region (Amorisco 2018). Taking different GC formation schemes and disruption into account, the GC systems become nearly isotropic or mildly tangential in the inner region and more radial in the outer region (Prieto & Gnedin 2008). The efficiency of GC disruption depends on their orbits within the halo. GCs that survive disruption could as a system be much more tangentially anisotropic than the original system, if the disruption occurs as a result of tidal effects (Vesperini et al. 2003; Brockamp et al. 2014). To understand the similarities and differences between the GCs and their host galaxies, as before, robust, observationally based estimates of the velocity anisotropy profiles of each population of GCs and central galaxy stars are important.

The velocity anisotropy of GCs in M87 has been provided by several dynamical models. The velocity dispersion anisotropy of GCs as a single population was assessed as near isotropic in Zhu et al. (2014) using a Syer & Tremaine (1996) made-to-measure model as implemented by Long & Mao (2010). From binning the GC data, Zhang et al. (2015) using a two population spherical Jeans model found that the red GC population is highly radially anisotropic, while the blue GC population is tangentially anisotropic. In comparison, Oldham & Auger (2018) used a discrete

chemo-dynamical model, without binning data and also describing population kinematics by spherical Jeans models. They obtained medium radial anisotropies for both red and blue GCs. The central galaxy has been observed with integral field units (IFUs) on a number of occasions [for example, Spectrographic Areal Unit for Research on Optical Nebulae (de Zeeuw et al. 2002) and Multi Unit Spectroscopic Explorer (MUSE; Emsellem, Krajnović & Sarzi 2014)] and stellar kinematic data are readily available.

Our objectives in this investigation are to

- (1) produce a mass estimate for M87 including both stellar and DM,
- (2) determine the velocity dispersion anisotropy of its two populations of GCs, and
- (3) consider what the anisotropy may reveal about the formation of M87 and its GC system.

The paper is arranged as follows. In Section 2, we introduce the approach we take in our investigations. In Section 3, we describe the main characteristics of M87 and the observational data we use. In Section 4, we provide details on our modelling methods, and then, in Section 5, we present results from our modelling runs. Finally, we discuss what our investigation has revealed in Section 6, and draw our conclusions in Section 7.

## 2 APPROACH

Section 1 was concerned with describing the scientific background and objectives for our investigation. This section defines at a top level the approach we are going to take to meet our objectives.

No new modelling scheme is created for our investigation. The existing axisymmetric modelling scheme utilized with NGC 5846 and its GCs (Zhu et al. 2016) is used as the basis of our models. In essence, the scheme is Jeans anisotropic modelling (JAM; Cappellari 2008) plus additions for discrete data (Watkins et al. 2013), and attributes such as metallicity (Zhu et al. 2016). A Markov Chain Monte Carlo (MCMC) toolkit is used for parameter estimation purposes. Note that within this paper, we use colour as a surrogate for metallicity.

The modelling scheme we have chosen (Zhu et al. 2016) is not the only scheme we considered. Other schemes include using higher order velocity moments (Napolitano et al. 2014), and two population spherical Jeans models (Pota et al. 2015; Zhang et al. 2015; Wasserman et al. 2018). These schemes perform a hard cut on colour to separate the two populations of GCs and use different levels of information from their discrete data, and results are not always consistent with each other. The Gauss–Hermite kurtosis equivalent  $h_4$  is taken as a direct indication of underlying velocity anisotropy in spherical systems. Used with our M87 data, it would suffer from poor sampling of data points and the need to bin across a large radius that would potentially affect any results. In the two population spherical Jeans models, the mass-anisotropy degeneracy is supposed to be partially removed. Without a hard cut on colour, there are a few approaches to chemo-dynamical modelling in the literature that try to model the colour and kinematic distributions of multiple GC populations simultaneously. By adopting Gaussian distributions of colour and velocity for each GC population, Agnello et al. (2014) found evidence of three GC populations (red, yellow, and blue) in M87, rather than two. The resulting velocity dispersion profiles of the red and yellow populations are rather noisy with low number statistics. The data can also be explained by combining the red and yellow GCs into one population with a metallicity gradient from the inner to outer regions. In estimating the mass

profile, Oldham & Auger (2018) described the kinematics of each population by spherical Jeans models, with both populations sharing the same potential. As indicated in Section 1, the two population spherical models of Zhang et al. (2015) result in a red GC system that is highly radially anisotropic, while the blue system has tangential anisotropy. Oldham & Auger (2018) found very different results in that both populations have moderate radial anisotropy. Unfortunately in these models, the discrete points are modelled spherically and some of the information on anisotropy is lost. GC systems extending to medium radii with elongated shapes and weak rotation may be better modelled as axisymmetric.

To obtain good constraints from the kinematic data for each population, we consider two populations of GCs. Each model we run is made up of three components in total – one for the central M87 galaxy and one for each of the two colour groups of GCs. Our best-fitting model comes from maximizing a log likelihood function involving contributions from all three-component models. In terms of our objectives, once we have a best-fitting model, we can estimate the mass of M87, and construct velocity dispersion anisotropy profiles for the two GC populations.

The role of the GCs is to constrain the mass distribution at large radii. In modelling the central galaxy, we use MUSE IFU data (Emsellem et al. 2014) and a radially varying mass-to-light ( $M/L$ ) ratio (Sarzi et al. 2018). This ratio has a negative gradient with increasing radius and is caused by a variation in the stellar initial mass function and the star formation history. The net effect of the varying ratio should be improved mass estimates with a decrease in the stellar matter fraction and an increase in the DM fraction, particularly at radii close to the galactic centre.

We avoid binning GC data during modelling in order to utilize fully the GC information content. This should lead to more robust velocity anisotropy estimates. We do use binning as part of the analysis process after modelling is complete, to assess the reproduction of the GC kinematics.

Our gravitational potential has the usual three components arising from a stellar mass central galaxy, a central black hole (modelled as a point mass), and DM. The stellar potential comes from a Multi-Gaussian expansion (MGE; Emsellem, Monnet & Bacon 1994; Cappellari 2002) of the central galaxy's surface brightness together with the varying  $M/L$  ratio. The DM potential is a generalized NFW profile and is used in both cusped and cored forms.

The MCMC toolkit we use to execute our models and arrive at our best-fitting model is the EMCEE package (Foreman-Mackey et al. 2013). In total, 15 parameters have to be determined – 3 for the gravitational potential and 2 for central galaxy, plus a further 5 for each GC population. Note that rotation and anisotropy parametrization are specific to the individual component models and are not handled globally.

The GC density profiles required for modelling with the Jeans equations come from a red/blue colour separation on the full population of GCs, not just those with LOS velocities. After contamination removal, the resulting number profiles are converted to MGEs for modelling purposes. This separation is only an approximation and may require updating after model analysis, with models then being rerun. GCs with LOS velocities are regarded as a single data set with any split if required into red and blue being performed once modelling is complete, using relative probability values. Typically, our calculations for specific GC populations are weighted calculations using these probability values as weights.

Having described the overall approach in this section, in Section 3 we describe our data in more detail, and expand on specific aspects of our modelling scheme in Section 4.

### 3 DATA

M87 (also known as Virgo A or NGC 4486) is a giant elliptical galaxy and is the second most massive and brightest galaxy of the Virgo Cluster. The distance to M87 is 16.5 Mpc (Mei et al. 2007; Blakeslee et al. 2009), and thus 1 kpc = 12.5 arcsec at this distance. Calculating the major axis of half-light isophote  $R_{\text{maj}}^{\text{maj}}$  (as defined in Cappellari et al. 2013) with the stellar image MGEs shown in Table 1 gives 178 arcsec. The position angle, measured from the image  $y$ -axis (North) over East to the photometric major axis, is  $162^\circ$ .

Throughout the paper,  $(x, y)$  represent the coordinates on the projected plane,  $x$  is along the major axis of galaxy,  $z$  is along the LOS,  $R$  indicates the projected elliptical radius along the  $x$  axis ( $R = \sqrt{x^2 + (y/q)^2}$ , we call it elliptical  $x$ -axis radius for simplicity),  $R'$  indicates the signed value ( $R' = \text{sign}(x) \times R$ ), and  $r$  indicates the intrinsic 3D radius ( $r = \sqrt{x^2 + y^2 + z^2}$ ), and  $q$  is the flattening of surface brightness on the 2D observational plane.

#### 3.1 Photometric data

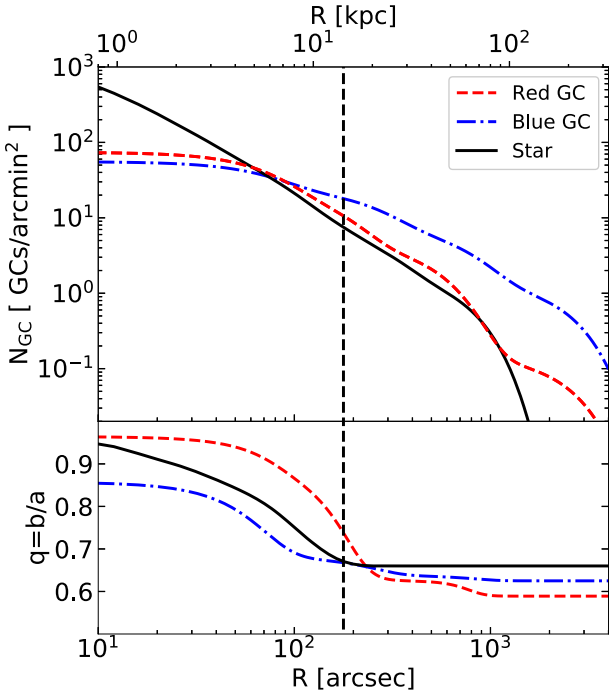
We take the photometric data from NGVS, including a  $g$ -band stellar surface brightness image of the main galaxy (Ferrarese et al. 2016), and GC photometric data. The stellar image covers a region of nearly  $25 \times 25$  arcmin<sup>2</sup>, which means the region  $R < 60$  kpc for M87. We fit the image with a 2D MGE and plot the surface brightness profile along the projected major axis in Fig. 1 (the solid black curve). For comparison purposes, we scale the surface brightness arbitrarily. We use this MGE to construct the luminous matter gravitational potential arising from the main galaxy.

The GC photometric data cover the  $120 \times 120$  arcmin<sup>2</sup> region around M87, and contain all the photometrically selected GC candidates with  $g$ -band magnitude smaller than 25. Following Durrell et al. (2014), GCs are classified as red if meeting the conditions  $0.85 < g - i < 1.15$ ,  $20 < g < 24$  and  $R < 60$  arcmin, and are classified as blue with  $0.55 < g - i < 0.85$ ,  $20 < g < 24$  and  $R < 60$  arcmin. To construct the individual projected surface number density maps for red and blue GCs, we take the following three steps: (1) we construct a smoothed 2D surface density map, replacing the position of each source by a circular Gaussian with  $\sigma = r_5$ , where  $r_5$  is the radial distance from each point to its fifth nearest neighbour (Durrell et al. 2014). In the inner region,  $r_5$  is very small, so we use  $\sigma = 30$  arcsec instead; (2) background and foreground contaminations are determined by calculating the average surface number density for the fields with  $80 < R < 100$  arcmin and where there is no obvious contamination from other galaxies. We subtract this constant value of contamination from the smoothed 2D map to form a cleaned 2D map; and (3) we fit a 2D MGE to the cleaned surface density map.

The best-fitting MGEs to the cleaned 2D surface density maps of the blue and red GCs are shown in Fig. B1 in Appendix B. In Fig. B2, we show how well the MGE model matches the surface density profile along the major and minor axes. For the blue GCs, the MGE model matches the surface density profile from the centre out to  $R \sim 5000$  arcsec (400 kpc), along both the major and minor axes. For the red GCs, there is an asymmetric structure to the north around  $R = 1000$ – $2000$  arcsec in the 2D surface density map. Our MGE model does not match this asymmetric structure, but does match the south side data. Thus, the MGE constructed surface density for the red GCs may have some uncertainties at  $R > 1000$  arcsec (80 kpc). In the kinematic data, the red GCs with velocity measurements are mostly located within 1000 arcsec. Note that the surface density profile

**Table 1.** The parameters of the MGEs for the *g*-band stellar image, red GC surface number density, and blue GC surface number density.  $\Sigma$  is the peak surface brightness or surface number density,  $\sigma$  is the dispersion, and  $q$  is the flattening of each Gaussian.

	g-Band stellar image			Red GCs			Blue GCs		
	$\Sigma$ ( $L_{\odot} \text{pc}^{-2}$ )	$\sigma$ (arcsec)	$q$	$\Sigma$ ( $\text{N arcmin}^{-2}$ )	$\sigma$ (arcsec)	$q$	$\Sigma$ ( $\text{N arcmin}^{-2}$ )	$\sigma$ (arcsec)	$q$
1	1.4597e+04	0.54	0.972	48.6816	51.2923	0.999	30.7399	53.4130	0.999
2	1.6965e+03	4.89	0.945	20.6292	115.4347	0.959	16.7829	179.9053	0.702
3	1.1315e+03	9.57	0.999	4.5376	382.9437	0.628	6.5417	512.8271	0.641
4	4.8415e+02	17.77	0.945	0.1514	1729.8027	0.589	1.4576	1721.5215	0.625
5	1.8451e+02	28.22	0.944	–	–	–	–	–	–
6	1.3272e+02	51.23	0.895	–	–	–	–	–	–
7	4.2822e+01	87.47	0.863	–	–	–	–	–	–
8	1.7848e+01	186.75	0.660	–	–	–	–	–	–
9	5.7816e+00	515.90	0.660	–	–	–	–	–	–

**Figure 1.** Surface density profiles of different tracers for M87. **Top panel:** The GC number density profiles and the stellar surface brightness along the major axis. The red-dashed and dash-dotted blue curves represent the number density profiles of the red and blue GCs, respectively, and they come from the 2D MGE fits to the number density of GCs. The flatness of the GC surface number density profiles in the inner 30 arcsec could partly be caused by catalogue incompleteness and smoothing. The solid black curve is from the MGE fit to the stellar surface brightness of NGVS *g*-band image and it has been arbitrarily scaled for comparison with GCs. **Bottom panel:** The flattening profiles. The red-dashed and dash-dotted blue curves represent the red and blue GCs respectively and the solid black curve is for the stars. The black vertical dashed line indicates the position of  $1 R_e^{\text{maj}}$ .

beyond the coverage of the kinematic data points does not have any effect on our dynamical model. The number density profiles of red and blue GCs along the major axis and the flattening profiles from the MGE models are shown in Fig. 1 (the red-dashed and dash-dotted blue curves). These two MGEs are used as density functions in our Jeans modelling.

The parameters of the MGEs for the *g*-band stellar image, the red GC number density, and the blue GC number density are shown in Table 1. The MGE of the stellar image is composed of nine Gaussians and the MGEs of both the red and blue GCs are composed

of four Gaussians. For each Gaussian of the MGEs,  $\Sigma$  is the peak surface brightness or surface number density,  $\sigma$  is the dispersion, and  $q$  is the flattening.

Comparing the density profiles in the top panel of Fig. 1, the stellar surface brightness is steeper than the GC surface number density profiles in the inner  $\sim 50$  arcsec. The flatness of the GC surface number density profiles in the inner 30 arcsec could partly be caused by catalogue incompleteness and smoothing. At larger radii, the slope of the red curve follows the stellar light, while the slope of the blue curve is shallower.

We also show the projected flattening profiles along the major axis in the bottom panel of Fig. 1. The solid black curve represents the surface brightness flattening profile, which varies from  $q \sim 0.96$  to  $q \sim 0.65$  from the centre to larger radii. The red-dashed and dash-dotted blue curves represent red and blue GCs. The red-dashed curve decreases from  $q \sim 0.95$  to  $q \sim 0.59$ , while the blue from  $q \sim 0.85$  to  $q \sim 0.63$  from the inner  $\sim 100$  arcsec to larger radii, after which they remain constant.

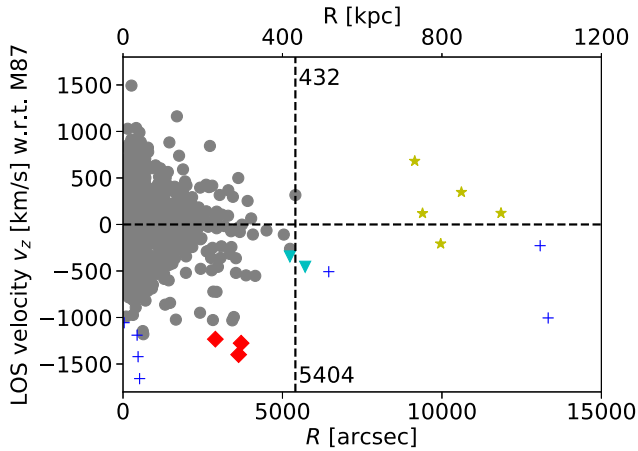
### 3.2 Kinematic data

We have two sets of kinematic data, MUSE IFU data for the central galaxy and LOS velocity data for the GCs. We describe the MUSE data first and then the GC data.

The MUSE data from Emsellem et al. (2014) cover the central region of nearly  $1 \times 1 \text{ arcmin}^2$  of M87 and comprise 523 Voronoi data bins. These bins correspond to stellar-kinematical measurements constructed from spectra binned to reach a minimum  $S/N = 300$ . In our models, we use the MUSE LOS velocity and velocity dispersion values together with their error estimates.

We have 911 M87 GCs with LOS velocities from three different sources, a systematic spectroscopic survey using 2dF/AAOmega multifibre spectrograph on the 3.9 m Anglo–Australian Telescope, an extensive spectroscopic survey using Hectospec multifibre spectrograph on the 6.5 m Multiple Mirror Telescope (Zhang et al. 2015), and the radial velocity catalogues of Virgo GCs compiled by Strader et al. (2011). Our GC catalogue is exactly the same as Zhang et al. (2015) but is slightly different from Zhu et al. (2014; the catalogues are from the same source but there are a few more data points in the Zhu catalogue). We intend to publish this data set with the final catalogue paper in the near future (Youkyung Ko et al. in preparation).

Although most GCs are located within 500 kpc, a small number of GCs are positioned as far out as 1000 kpc, and may be the intracluster GCs (Longobardi et al. 2018a). We expect to be able to discover the intracluster GCs hidden in our catalogue and



**Figure 2.** Groups in the GCs. There is a big group marked with the grey points and several other small groups are marked with the red diamonds, the yellow asterisks, and the cyan triangles. The GCs that are isolated and have no neighborhood are marked with the blue pluses. The vertical dashed line marks the maximum radius of the grey group.

remove them. To do so, we adopt the friends-of-friends algorithm (Huchra & Geller 1982; Geller & Huchra 1983; Eke et al. 2004) and a dimensionless distance between two GCs as the linking length defined by

$$l^2 = w_\phi \phi_{ij}^2 + w_v (v_i - v_j)^2 + w_d (d_i - d_j)^2, \quad (1)$$

where  $\phi_{ij} = \text{atan2}(x_j - x_i, y_j - y_i)$  is the angular separation of two GCs relative to the centre of galaxy,  $v$  is the LOS velocity, and  $d = \sqrt{x^2 + y^2}/q$  is the elliptical  $x$ -axis radius. ( $w_\phi$ ,  $w_v$ ,  $w_d$ ) are three weighting factors used to normalize the quantities (the largest possible angular separation is  $\pi$ , distance 1064 kpc, and velocity  $3150 \text{ km s}^{-1}$ ) and take into account uncertainties. The definition of the linking length  $l$  in this paper is similar to the 4-distance in Starkeburg et al. (2009), but we only possess information on three of the six components in phase space: the azimuth angle  $\phi$  ( $\phi = \text{atan2}(y/x)$ ), the elliptical  $x$ -axis radius  $d$ , and the LOS velocity  $v$ :

$$\begin{aligned} w_\phi &= \left(\frac{1}{\pi}\right)^2, \\ w_d &= \left(\frac{1}{1064}\right)^2 \frac{\left(\frac{d_{\text{err}}(i)}{d(i)}\right)^2 + \left(\frac{d_{\text{err}}(j)}{d(j)}\right)^2}{2 \langle \frac{d_{\text{err}}}{d} \rangle^2}, \\ w_v &= \left(\frac{1}{3150}\right)^2 \frac{v_{\text{err}}^2(i) + v_{\text{err}}^2(j)}{2 \langle v_{\text{err}} \rangle^2}, \end{aligned} \quad (2)$$

where  $d_{\text{err}}$  and  $v_{\text{err}}$  are the uncertainties on the elliptical  $x$ -axis radius and LOS velocity, respectively, and the relative uncertainty  $d_{\text{err}}/d$  is used as distance uncertainties scale with the distance itself. Quantities within  $\langle \rangle$  denote the average uncertainties over the whole sample.

We set the linking length of the friends-of-friends algorithm as  $l = 0.14$  and plot the groups in Fig. 2. There is a big group marked with the grey points and several other small groups are marked with the red diamonds, the yellow asterisks, and the cyan triangles. The GCs that are isolated and have no neighbours are marked with the blue pluses. Most of the GCs are in the grey group. We treat GCs belonging to the non-grey groups as the intracluster GCs (17 in total) and remove them, leaving 894 GCs in our catalogue. These GCs are considered bound by the potential of M87 and trace the dynamical history of the galaxy. The 17

GCs excluded have a velocity dispersion of  $620 \text{ km s}^{-1}$ , which is consistent with the intracluster GCs identified in Longobardi et al. (2018a). The remaining 894 GCs still show some asymmetry in velocity distribution in the region with  $R > 2000 \text{ arcsec}$ . Having checked the points with  $R > 2000 \text{ arcsec}$  and  $V_z < -700 \text{ km s}^{-1}$ , we find that their photometric properties ( $u - i$  versus  $i - K_s$  diagram) are consistent with GCs, and are unlikely to be Milky Way stars (Muñoz et al. 2014). We presume this asymmetry is caused by low number statistics and we keep all the grey points in Fig. 2 as model constraints.

Regarding the non-symmetric structure in the north according to the surface number density maps, we do not mask out any red GCs that might be associated with the structure. At this radius, the number of red GCs is small compared to the number of blue GCs, and any impact on the modelling will be small.

The kinematic data in the projected sky plane are shown in Fig. 3. Each black plus symbol represents one GC, while the yellow square in the centre indicates the area covered by MUSE IFU data. The five dashed green ellipses have  $x$ -axis radii of (1/3, 1, 5, 15, 30)  $R_e^{\text{maj}}$ .

## 4 METHOD

Much of the material in this section just describes an application of the method in Zhu et al. (2016) and so we present it in summarized form.

### 4.1 Gravitational potential

The total potential is a combination of the contributions from stars in the central galaxy, the central supermassive black hole and the DM halo.

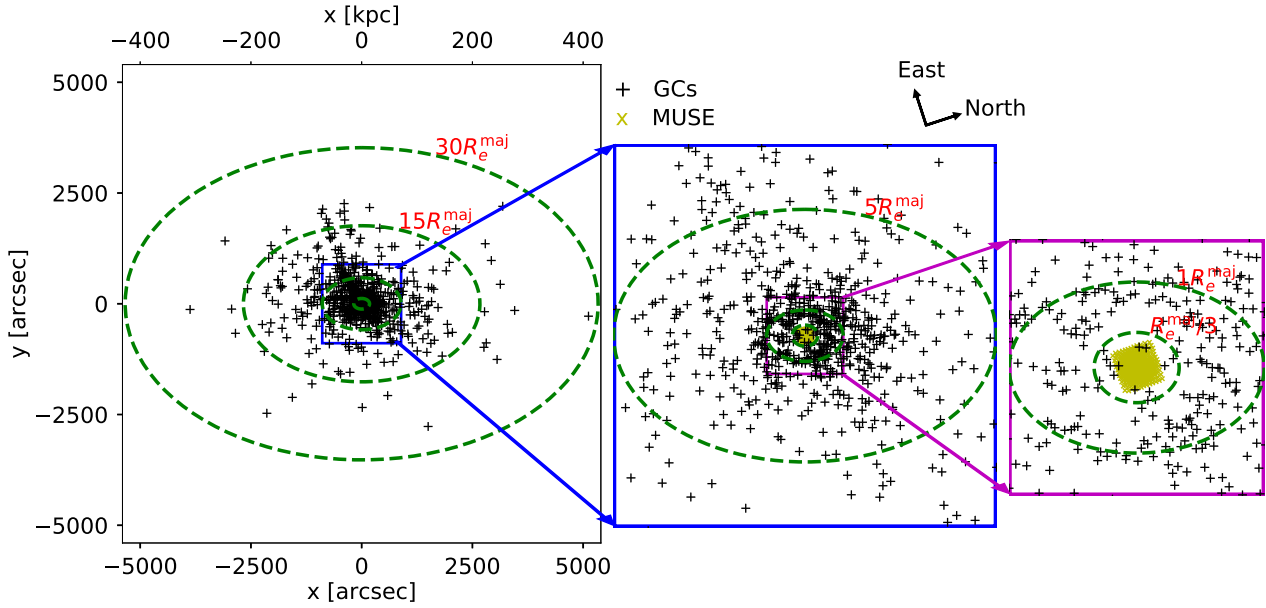
#### 4.1.1 Central galaxy

To create the potential from the stars in the central galaxy, we construct a 2D MGE from the galaxy's NGVS  $g$ -band surface brightness and apply the MGE formalism to the 2D MGE to arrive at the potential. In doing so, we need to accommodate the varying  $M/L$  ratio identified in Section 2 and the inclination of the galaxy to the LOS. We proceed as follows.

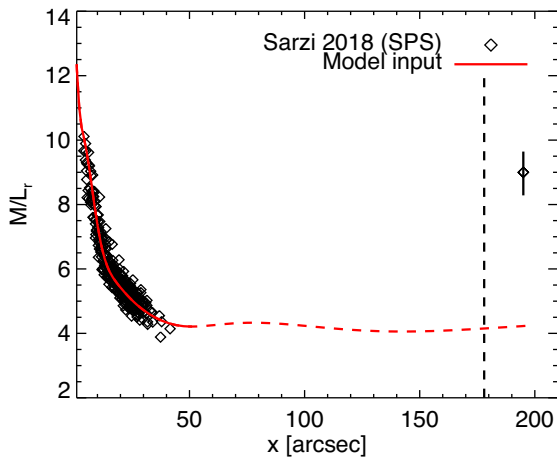
Recent studies show that the stellar  $M/L$  ratio for M87 has a gradient from the centre outwards (Oldham & Auger 2018; Sarzi et al. 2018). The black diamonds in Fig. 4 show the  $M/L$  ratio gradient from stellar population synthesis (Sarzi et al. 2018). Using MUSE observations, Sarzi et al. (2018) found a strong negative stellar  $M/L$  ratio gradient in M87. To impose the stellar  $M/L$  ratio variation, we assign each Gaussian component of the 2D MGE its own  $M/L$  ratio  $\gamma_i$  and define

$$\gamma^*(x) = \frac{\sum_i^N \gamma_i l_i \exp\left(-\frac{x^2}{2\sigma_i^2}\right)}{\sum_i^N l_i \exp\left(-\frac{x^2}{2\sigma_i^2}\right)}, \quad (3)$$

where  $x$  is along the major axis,  $N$  is the number of the MGE components, having peak surface brightness  $l_i$  and dispersion  $\sigma_i$  along the the major axis. We perform a minimum  $\chi^2$  fit to the  $M/L$  ratio gradient points constructed from equation (3), and obtain a solution for the  $\gamma_i$ . As shown in Fig. 4, the resulting best-fitting  $\gamma^*(x)$  profile (the solid red curve) matches the data points well. The data points from Sarzi et al. (2018) only extend from  $\sim 4$  to  $\sim 40 \text{ arcsec}$ . The interpolation inwards is consistent with the  $M/L$  profile used in Oldham & Auger (2018). There is evidence that



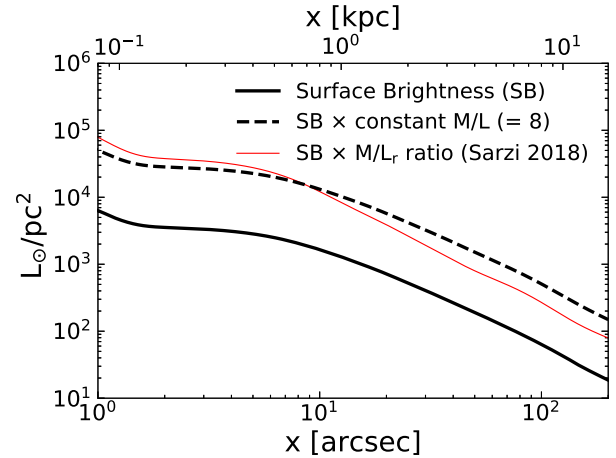
**Figure 3.** The kinematic data in the observational plane.  $x$  and  $y$  indicate the major and minor axes of the galaxy and the figure orientation is shown in the top right-hand corner of the middle panel. The black plus symbols represent the positions of GCs and the yellow multiplication signs represent the integrated stellar light from MUSE. The five green ellipses have  $x$ -axis radii of  $(1/3, 1, 5, 15, 30) R_e^{\text{maj}}$ . The middle panel is a zoom-in of the central  $5 R_e^{\text{maj}}$  region of the left-hand panel and the right-hand panel is a zoom-in of the central  $1 R_e^{\text{maj}}$  region of the middle panel.



**Figure 4.** Stellar  $M/L_r$  ratio gradient of M87.  $x$  is along the major axis. The black diamonds are the  $M/L_r$  ratio gradient from Sarzi et al. (2018). The diamond with the error bar in the middle right presents the mean of errors. The solid red curve is our best-fitting  $M/L_r$  ratio gradient with equation (3). The red-dashed curve is the extension of our best-fitting  $M/L_r$  ratio to large radii. The vertical dashed line indicates a radius of  $1 R_e^{\text{maj}}$ .

M87 has constant age and metallicity outside  $\sim 60$  arcsec (Montes et al. 2014), and so we extrapolate the  $M/L$  ratio gradient from Sarzi et al. (2018) to large radii by assuming there is no strong variation in the outer region (the red-dashed curve).

Fig. 5 shows the comparison between the stellar surface brightness profile and  $M/L$  ratio gradient revised surface mass density profile along the major axis. The solid black curve is the stellar surface brightness profile. The dashed black curve is the stellar surface mass density profile using constant  $M/L = 8$ , and the solid red thin curve is the surface mass density profile constructed with the  $M/L$  profile as shown in Fig. 4.



**Figure 5.** Stellar surface brightness profiles of M87.  $x$  is along the major axis. The solid black, dashed black and thin red curves represent the stellar surface brightness profile, the stellar surface mass density profile using a constant  $M/L$  ratio = 8, and the surface mass density profile using the  $M/L_r$  profile from Fig. 4.

Note that the  $M/L$  radial profile from Sarzi et al. (2018) is in the SDSS  $r$ -band whereas the projected stellar surface brightness used in our models is from the NGVS  $g$ -band photometric image, thus we have used  $L_g * M/L_r$  to impose the  $M/L$  gradient in constructing the stellar mass model. M87 has a rather constant ratio of  $L_r/L_g$  with increasing radius (Montes et al. 2014), thus the above assumption works for including the gradient itself, but has a mismatch in the absolute normalization of stellar mass. Kinematics can be independent, and potentially stronger, constraints on the scale of stellar  $M/L$  ratio (Lyubenova et al. 2016). Although Sarzi et al. (2018) has tried to pin down the effect of a varying IMF, we still allow a constant scale factor  $\Upsilon$  as a free parameter in our models.

In order to deproject the 2D mass density to give a 3D density and then the potential, we need to know the inclination of M87 to the LOS. Inclination angle is poorly constrained by kinematic data for a slow-rotating galaxy (van den Bosch & van de Ven 2009). We tried an initial model which showed that an inclination angle of  $i = 90^\circ$  is preferred but it has large uncertainties (shown in Appendix C). Taking this into account, we choose to fix the inclination angle  $i = 90^\circ$  (edge-on) for the rest of the models.

#### 4.1.2 Black hole

M87 is believed to have a supermassive black hole. We adopt a black hole with a mass of  $(6.6 \pm 0.4) \times 10^9 M_\odot$  (Gebhardt et al. 2011) and model it as a point mass. The sphere of influence radius of the black hole is about 4 arcsec at the M87 distance of 16.5 Mpc. Note that stars anywhere in the galaxy, depending on their orbits, may have been affected by the black hole.

#### 4.1.3 Dark matter halo

We adopt a generalized Navarro, Frenk & White (1997) (NFW) DM density distribution

$$\rho(r) = \frac{\rho_s}{(r/r_s)^\gamma (1 + r/r_s)^{3-\gamma}}, \quad (4)$$

with  $r^2 = x^2 + y^2 + z^2/p_z^2$  for an axisymmetric model. The scale density  $\rho_s$ , the scale radius  $r_s$  and the density slope  $\gamma$  are the three DM halo parameters. Cosmological simulations show that the shape of a DM halo may be correlated with, but is usually rounder than, the luminous tracers (Wu et al. 2014). In the case of M87, it is unlikely to have a very flattened DM halo. The flattening  $p_z$  of a DM halo is highly degenerate with its radial profile, and it is hard for it to be constrained by the kinematics of an external galaxy. To simplify matters, we just adopt a spherical DM halo ( $p_z = 1$ ). We construct two models with different radial profiles in this paper, one with a classic cusped DM density profile ( $\gamma = 1$ ), and the other with a cored DM density profile ( $\gamma = 0$ ). In each case,  $\rho_s$  and  $r_s$  are the two free parameters for the DM halo.

## 4.2 The likelihood function

As indicated earlier, we follow the approach in Zhu et al. (2016). We include, for completeness, a short description of the likelihood calculations.

#### 4.2.1 Discrete GCs

The likelihood function for the GCs can be constructed by considering three probabilities involving chemistry, spatial position, and kinematics. We describe these three probabilities individually and then combine them together to form the GCs' likelihood function.

We have a catalogue of  $N_{GC}$  discrete GCs where the  $i$ th GC has sky coordinates  $(x_i, y_i)$ , LOS velocity  $v_{z,i} \pm \delta v_{z,i}$  and measured metallicity  $Z_i \pm \delta Z_i$ . As stated earlier, we adopt the colour  $g - i$  as a proxy for metallicity. For this paper, we consider two GC populations, red and blue, corresponding to the metal rich and metal poor.

For each population  $k$  ( $k = \text{red or blue}$ ), we adopt a Gaussian metallicity distribution with mean metallicity  $Z_0^k$  and dispersion  $\sigma_z^k$

as the free parameters. The metallicity probability of the  $i$ th GC in population  $k$  is

$$P_{\text{chm},i}^k = \frac{1}{\sqrt{2\pi[(\sigma_z^k)^2 + \delta Z_i^2]}} \exp\left[-\frac{1}{2} \frac{(Z_i - Z_0^k)^2}{(\sigma_z^k)^2 + \delta Z_i^2}\right]. \quad (5)$$

Taking the observed projected number density for population  $k$  as  $\Sigma^k(x, y)$ , the spatial probability of the  $i$ th GC in population  $k$  is

$$P_{\text{spa},i}^k = \frac{\Sigma^k(x_i, y_i)}{\Sigma_{\text{obj}}(x_i, y_i)}, \quad (6)$$

where  $\Sigma_{\text{obj}} = \Sigma^{\text{red}} + \Sigma^{\text{blue}}$  is the combined density in our two population model.

Assuming a Gaussian velocity distribution with mean value and dispersion predicted by an axisymmetric JAM model for population  $k$ , the kinematic probability of the  $i$ th GC in population  $k$  is

$$P_{\text{dyn},i}^k = \frac{1}{\sqrt{(\sigma_i^k)^2 + (\delta v_{z,i})^2}} \exp\left[-\frac{1}{2} \frac{(v_{z,i} - \mu_i^k)^2}{(\sigma_i^k)^2 + (\delta v_{z,i})^2}\right], \quad (7)$$

where  $\mu_i^k$  and  $\sigma_i^k$  are the LOS mean velocity and the velocity dispersion predicted by the model at the position  $(x_i, y_i)$ . Axisymmetric JAM models assume (i) the velocity ellipsoid is aligned with the cylindrical coordinates and (ii) the velocity anisotropy in the meridional plane  $\beta_z^k = 1 - \overline{v_z^2}/\overline{v_R^2}$  is constant (Cappellari 2008).

In JAM modelling, the number density of population  $k$  is expressed in terms of an MGE. In principle, for each Gaussian  $j$  in the MGE (see Table 1), the anisotropy parameters  $\beta_{z,j}^k$  and rotation parameters  $\kappa_j^k$  can take different values, and thus it is possible to model the velocity anisotropy in the meridional plane and intrinsic rotation at different radii. To investigate the rotation properties at inner and outer radii, we assume two different constant values for the first three and the last Gaussians in the MGE, and so we include two free parameters  $\kappa_{\text{inner}}^k$  and  $\kappa_{\text{outer}}^k$  for population  $k$ . Under the assumptions of axisymmetric JAM modelling, the velocity anisotropy in the meridional plane  $\beta_z^k$  is constant and we include  $\beta_z^k$  of population  $k$  as the another free parameter so we can investigate anisotropy.

Having defined the three probabilities individually, we now combine them to give a likelihood value. Given the chemical, spatial, and kinematical probabilities, the likelihood for the  $i$ th GC in the model is

$$L_i = \sum_k P_i^k = \sum_k P_{\text{chm},i}^k P_{\text{spa},i}^k P_{\text{dyn},i}^k, \quad (8)$$

where  $k = \text{red or blue}$ . The summation is over the red and blue populations in our model.  $P_i^k = P_i^k/L_i$  is the relative probability of the  $i$ th GC belonging to population  $k$ .

The total log likelihood of all the GCs is

$$\mathcal{L} = \sum_i^{N_{GC}} \log L_i. \quad (9)$$

#### 4.2.2 Stellar kinematics

The MUSE IFU data provide the LOS mean velocities and velocity dispersions ( $\mu_j^s \pm \delta\mu_j^s, \sigma_j^s \pm \delta\sigma_j^s$ ) of stars in Voronoi bins with centroids  $(x_j, y_j)$ . We use another axisymmetric JAM model with free velocity anisotropy and rotation parameters to describe the stellar kinematics. We calculate  $\chi_{\text{star}}^2$  between the observed stellar mean velocities and velocity dispersions and the model predictions



as

$$\chi_{\text{star}}^2 = \sum_{j=1}^{N_s} \left( \frac{\mu_j^s - \mu_j^{\text{star}}}{\delta\mu_j^s} \right)^2 + \sum_{j=1}^{N_s} \left( \frac{\sigma_j^s - \sigma_j^{\text{star}}}{\delta\sigma_j^s} \right)^2, \quad (10)$$

where  $N_s$  is the number of Voronoi bins and  $(\mu_j^{\text{star}}, \sigma_j^{\text{star}})$  are the model predicted mean velocities and the velocity dispersions.

#### 4.2.3 Total likelihood

We obtain the best-fitting parameters by maximizing the log likelihood of the GCs and minimizing  $\chi_{\text{star}}^2$ . This can be achieved by maximizing the combined likelihood as defined by

$$\mathcal{L}_{\text{tot}} = \mathcal{L} - \frac{1}{2} \alpha_s \chi_{\text{star}}^2, \quad (11)$$

where  $\alpha_s$  is a weight parameter that is used to balance the relative importance of the GCs and the stellar kinematics and also to balance the two terms numerically. The value of  $\alpha_s$  is determined by experimentation. We set  $\alpha_s = 0.004$  in our model. With this choice of  $\alpha_s$ , the kinematics of GCs and stars are both matched well by our model.

### 4.3 Model parameters

We summarize the 15 free parameters in our model. Three of them are related to the gravitational potential as described in Section 4.1:

- (1)  $\rho_s$ , DM scale density;
- (2)  $d_s \equiv \log(\rho_s^2 r_s^3)$ , instead of DM scale radius  $r_s$  for the purpose of reducing the strong degeneracy between  $\rho_s$  and  $r_s$ ;
- (3)  $\Upsilon$ , a constant scale factor of stellar mass.

For the stars in the central galaxy, the projected stellar density profile is set by the surface brightness profile as shown in Fig. 1 and the dynamical distribution is described by an axisymmetric JAM model, leaving two dynamical free parameters for the stars:

- (4)  $\lambda^{\text{star}} \equiv -\ln(1 - \beta_z^{\text{star}})$ , a symmetric re-casting of the constant velocity anisotropy parameter of the stars;
- (5)  $\kappa^{\text{star}}$ , the rotation parameter of the stars.

We consider two GC populations, red and blue, each with a Gaussian metallicity (colour) distribution, a fixed spatial density profile as shown in Fig. 1 and a dynamical distribution described by an axisymmetric JAM model, leaving two chemical and three dynamical free parameters for each population, which have been introduced in detail in Section 4.2. So, for each population, there are five free parameters:

- (6)  $Z_0^k$ , mean of the Gaussian metallicity (colour) distribution for population k (red or blue);
- (7)  $\sigma_{Z^k}$ , sigma of the Gaussian metallicity (colour) distribution;
- (8)  $\lambda^k \equiv -\ln(1 - \beta_z^k)$ , a symmetric re-casting of the constant velocity anisotropy in the meridional plane;
- (9)  $\kappa_{\text{inner}}^k$ , the inner rotation parameter; and
- (10)  $\kappa_{\text{outer}}^k$ , the outer rotation parameter.

## 5 RESULTS

Both the cusped and cored models described in Section 4.1.3 are acceptable. For simplicity, we show the results from the cored model in this section and the corresponding results from the cusped model in the appendix A. The results from the cusped model hardly change relative to the cored model.

### 5.1 Best-fitting parameters

We create two sets of models, one with a cusped DM halo and the other with a cored DM halo. As indicated in Section 2, we use the EMCEE package to run our models. For each model, we use 100 walkers and run for 600 steps to ensure that the parameter space is sufficiently explored and the parameter values have converged. The best-fitting parameters for the two models are shown in Table 2. Note that we have converted  $d_s$  back to  $r_s$ , and  $\lambda^{\text{star}}$ ,  $\lambda^{\text{red}}$ , and  $\lambda^{\text{blue}}$  back to  $\beta_z^{\text{star}}$ ,  $\beta_z^{\text{red}}$ , and  $\beta_z^{\text{blue}}$ .

Fig. 6 shows the posterior marginalized distributions of the last 50 steps for 11 of the 15 parameters, which are related directly to the dynamical properties. The distributions are from the model with a cored DM halo. The points in each scatter plot show the 2D posterior marginalized distribution of the two corresponding parameters and the colour of each point indicates the relative probabilities from low (blue) to high (red). The three ellipses in each scatter plot show the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  regions of the covariance ellipses of the distribution. When the covariance ellipse is tilted, the ellipticity indicates the degree of correlation of the two parameters such that the larger the ellipticity, the greater the correlation, indicating a stronger degeneracy. The histogram in each row shows the 1D posterior marginalized distribution of the parameter and the curve is a Gaussian curve plotted with the mean and standard deviation of the distribution.

The cored and cusped models lead to similar maximum likelihoods, and so we have little ability to constrain the DM density slope. For the cored/cusped models, there are still some degeneracies between the DM scale density  $\rho_s$ , the DM scale radius  $d_s$ , and the stellar  $M/L$  ratio scale factor  $\Upsilon$  (see Fig. 6) even for the fixed DM density slope  $\gamma$ . There are no noticeable degeneracy concerns between the anisotropy and rotation parameters.

The  $\rho_s$  and  $r_s$  of the best-fitting cored halo correspond to virial values of  $M_{200} = (2.2 \pm 0.4) \times 10^{13} M_{\odot}$  and  $r_{200} = 573 \pm 30$  kpc, while those for the cusped halo correspond to  $M_{200} = (3.1 \pm 0.4) \times 10^{13} M_{\odot}$  and  $r_{200} = 647 \pm 30$  kpc.  $M_{200}$  and  $r_{200}$  of the cored and cusped models are roughly consistent with each other. Note that for the cusped model, its concentration is  $C = 13 \pm 3$ . The concentration uncertainty is relatively high, but consistent within  $2\sigma$  scatter, compared to the prediction of the  $M_{200} \sim r_{200}$  relation from cold dark model (Dutton & Macciò 2014).

In using the stellar  $M/L_r$  gradient from Stellar Population Analysis (SPS, Sarzi et al. 2018) and  $g$ -band surface brightness, we allow a scale parameter  $\Upsilon$  to be free. We obtain  $\Upsilon = 0.85 \pm 0.03$  for the cusped model, and  $\Upsilon = 0.93 \pm 0.02$  for the cored model. The ratio of stellar mass obtained by our dynamical model to that obtained by SPS in Sarzi et al. (2018) is  $M^{\text{dyn}}/M^{\text{SPS}} = L_g * \Upsilon / L_r$ . We have  $g - r \sim 0.8$  for M87 (Chen et al. 2010) and  $g - r = 0.46$  for the Sun, thus  $M^{\text{dyn}}/M^{\text{SPS}} \sim 0.73\Upsilon$  that is  $\sim 0.62$  for the cusped model and  $\sim 0.68$  for the cored model. The total stellar mass from our dynamical model is  $\sim 4.34 \times 10^{11} M_{\odot}$ , which leads to  $M_{\text{star}}/M_{\text{tot}} \sim 0.017$ . This value is a little higher than might be expected from the stellar-to-halo mass relation (e.g. Behroozi, Conroy & Wechsler 2010).

M87 is at the centre of the Virgo cluster, which is clearly not relaxed according to the non-continuous variation in velocity dispersions between the GCs within the dynamical boundary and the intracluster members. We would like to emphasize that the virial mass we have estimated here does not represent the virial mass of the Virgo cluster.

**Table 2.** The best-fitting parameters for the cusped and cored DM halo models. The parameters are presented in two rows for each model. The first row from left to right: DM scale density  $\rho_s$  [ $10^{-3} M_\odot \text{pc}^{-3}$ ], DM scale radius  $r_s$  [kpc], the  $M/L$  ratio scale factor  $\Upsilon$ , the velocity anisotropy, and the rotation parameters of the stars  $\beta_z^{\text{star}}$ ,  $\kappa^{\text{star}}$ . The second row from left to right: the mean and dispersion of the colour  $g - i$  Gaussian distributions of the red GCs  $Z_0^{\text{red}}$  and  $\sigma_Z^{\text{red}}$ , the velocity anisotropy parameter for the red GCs  $\beta_z^{\text{red}}$ , the inner and outer rotation parameters for the red GCs  $\kappa_{\text{inner}}^{\text{red}}$ ,  $\kappa_{\text{outer}}^{\text{red}}$ , and the maximum likelihood  $\mathcal{L}_{\text{max}}$ . The third row from left to right: the mean and dispersion of the colour  $g - i$  Gaussian distributions of the blue GCs  $Z_0^{\text{blue}}$  and  $\sigma_Z^{\text{blue}}$ , the velocity anisotropy parameter for the blue GCs  $\beta_z^{\text{blue}}$ , the inner and outer rotation parameters for the blue GCs  $\kappa_{\text{inner}}^{\text{blue}}$ ,  $\kappa_{\text{outer}}^{\text{blue}}$ , and the maximum likelihood  $1 \sigma$  confidence level  $\delta\mathcal{L}$ . Note that we have converted  $d_s$  back to  $r_s$ , and  $\lambda^{\text{star}}$ ,  $\lambda^{\text{red}}$ , and  $\lambda^{\text{blue}}$  back to  $\beta_z^{\text{star}}$ ,  $\beta_z^{\text{red}}$ , and  $\beta_z^{\text{blue}}$ .

DM	$\rho_s$ $Z_0^{\text{red}}$ $Z_0^{\text{blue}}$	$r_s$ $\sigma_Z^{\text{red}}$ $\sigma_Z^{\text{blue}}$	$\Upsilon$ $\beta_z^{\text{red}}$ $\beta_z^{\text{blue}}$	$\beta_z^{\text{star}}$ $\kappa_{\text{inner}}^{\text{red}}$ $\kappa_{\text{inner}}^{\text{blue}}$	$\kappa^{\text{star}}$ $\kappa_{\text{outer}}^{\text{red}}$ $\kappa_{\text{outer}}^{\text{blue}}$	$\mathcal{L}_{\text{max}}$ $\delta\mathcal{L}(1 \sigma)$
Cusped	$10.8^{+2.8}_{-2.2}$	$51^{+11}_{-9}$	$0.85 \pm 0.03$	$0.013^{+0.022}_{-0.023}$	$0.005 \pm 0.032$	
	$0.946 \pm 0.009$	$0.111 \pm 0.008$	$-0.06^{+0.13}_{-0.15}$	$-0.53 \pm 0.16$	$0.22 \pm 0.32$	$-19806.5$
	$0.751 \pm 0.004$	$0.072 \pm 0.004$	$0.36^{+0.05}_{-0.05}$	$-0.35 \pm 0.21$	$0.10 \pm 0.18$	$6.9$
Cored	$208^{+36}_{-30}$	$15.7^{+2.3}_{-2.0}$	$0.93 \pm 0.02$	$0.007^{+0.023}_{-0.023}$	$0.01 \pm 0.03$	
	$0.944 \pm 0.008$	$0.112 \pm 0.008$	$-0.09^{+0.19}_{-0.23}$	$-0.47 \pm 0.17$	$0.04 \pm 0.39$	$-19806.3$
	$0.751 \pm 0.004$	$0.071 \pm 0.003$	$0.32^{+0.05}_{-0.05}$	$-0.40 \pm 0.20$	$-0.08 \pm 0.15$	$7.4$

## 5.2 Models

### 5.2.1 GCs separation into two populations

The relative probability of the  $i$ th GC being red ( $P_i^{\text{red}}$ ) or blue ( $P_i^{\text{blue}}$ ) is calculated from the best-fitting model. With the posterior likelihood, the spatial, colour, and velocity distribution of population  $k$  can be calculated from the overall observed GC data values by weighting each by  $P_i^k$ . For convenience of analysis of the two populations after modelling, we separate the GCs based on the relative probability of the  $i$ th GC being red or blue in Section 5.3 (Figs 11 and 12) and Section 5.4 (Figs 13 and 14). The criteria used are that the  $i$ th GC is identified as red if  $P_i^{\text{red}} > 0.5$ , or blue if  $P_i^{\text{blue}} > 0.5$ .

Fig. 7 shows the spatial, colour, and kinematic distributions of the red and blue populations from our best-fitting model with a cored DM halo. The top and bottom scatter panels show the  $g - i$  and LOS velocity distributions along  $R'$ . The points are coloured by the relative probabilities of belonging to the red GCs ( $P_i^{\text{red}}$ ) from low (blue) to high (red), which means the red points have higher probability of belonging to the red GCs and the blue points have higher probability of belonging to the blue GCs. In the top and right-hand panels, we show the overall spatial, colour, and velocity distributions for the red and blue populations. The histograms of the red and blue populations are calculated from all the kinematic GC data points weighted by  $P_i^{\text{red}}$ , and  $P_i^{\text{blue}}$ . The curves for the density distributions are from the best-fitting MGEs for the photometric red (dashed) and blue (solid) GCs. The curves for the colour and velocity distributions of red/blue GCs are from the best-fitting chemo-dynamical models. The histograms and the corresponding curves of density and colour distribution constructed by the kinematic data are generally consistent with the photometric data. The density curves and histograms in the inner region are depressed due to incompleteness in the photometric and kinematic data. The histograms of the  $g - i$  distributions for the red and blue GCs show an overlap centred at  $g - i \sim 0.85$ , which is consistent with the hard cut colour value used in Section 3.1.

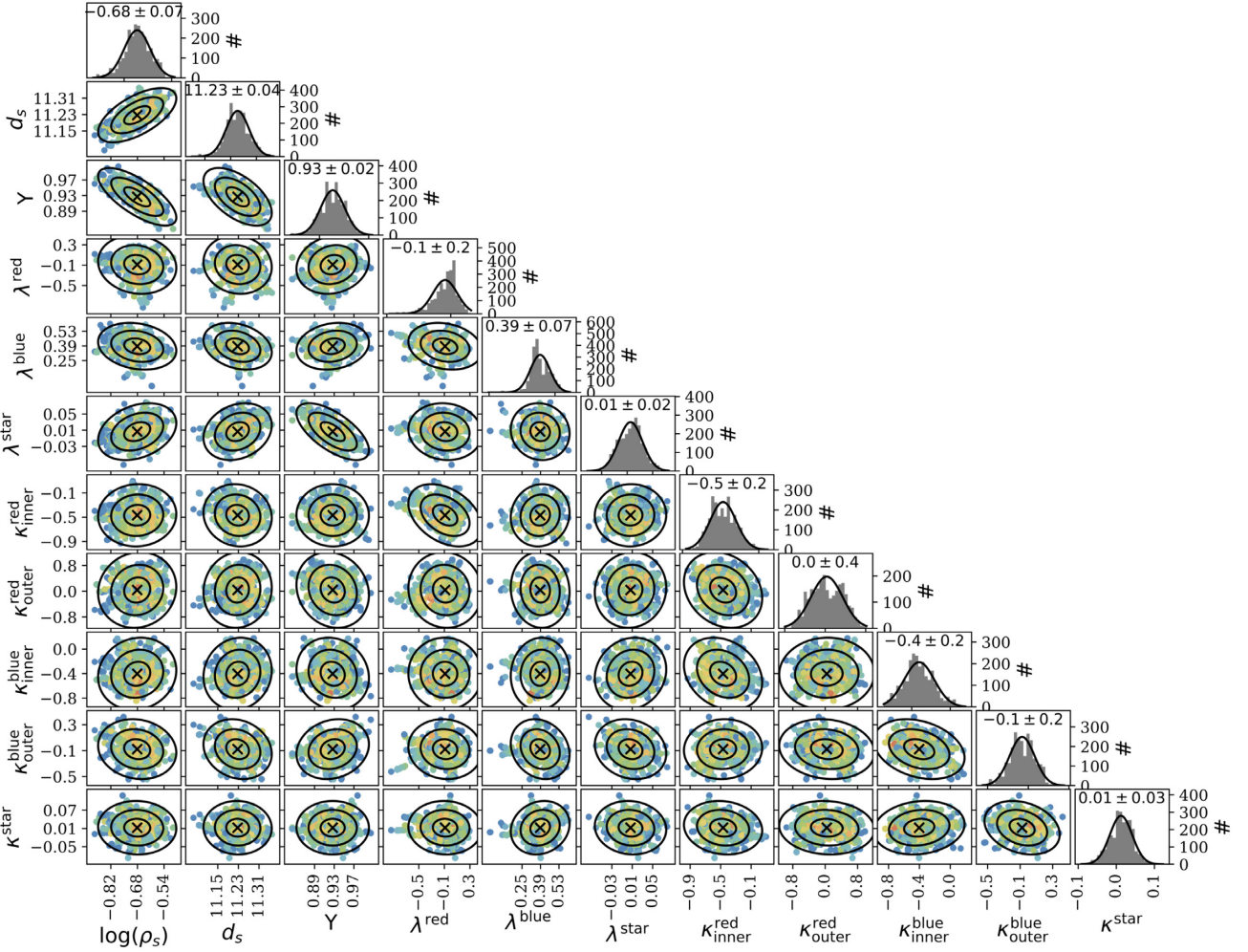
Zhang et al. (2015) used a double-Gaussian fitting to the NGVS  $g - i$  bimodal colour distribution of the photometrically selected GCs in M87. This Gaussian fitting indicates that the blue and red components overlap at  $g - i = 0.89$  mag. We use this value of  $g -$

$i = 0.89$  to separate initially the red and blue photometric GCs so that we are able to construct MGEs for each component. Having run our model, which utilizes the kinematical GC data, using these MGEs, we find that the overlap is now centred at  $g - i \sim 0.85$  mag. We reconstruct our MGEs using this revised overlap value, and rerun our model. This second modelling run returns the consistent results we have shown. The photometrically selected GCs may contain other objects, such as stars, which may cause a bias in the  $g - i$  separation from the double-Gaussian fitting. Our approach ensures we achieve consistency between the photometric and kinematical GCs, which may eliminate this bias.

### 5.2.2 Kinematic data reproduction

In this section, we aim to quantify how well our model can recover the kinematics for both the discrete GCs and the stars in the central galaxy. In Fig. 8, for the best-fitting cored model, we show the mean velocity and velocity dispersion maps for the red and blue GCs separately. The red GCs have moderate rotation, while the rotation of the blue GCs is much weaker. On the velocity dispersion maps, the red GCs have larger dispersions along the major axis than those along the minor axis, while conversely, the blue GCs have larger dispersions along the minor axis, which are typical features for near edge-on systems with negative  $\beta_z^{\text{red}} (\sim -0.08)$  and positive  $\beta_z^{\text{blue}} (\sim 0.35)$  (Cappellari 2008).

We maximize the likelihood of the discrete data points within a model. For the discrete data points with their single LOS velocity  $V_z$  at a given position, we must bin the data points on the 2D plane to obtain the mean velocity and velocity dispersion values for direct comparison with the model predictions. In order to give a visual impression of how well our model matches the observed kinematic data, we bin the model predictions and observed data with the same binning scheme. The binning scheme is shown in the top left-hand panel of Fig. 8. First, we divide the projected plane into eight sectors equally and the regions along the major axis, minor axis, and the  $45^\circ$  and  $135^\circ$  diagonals are filled with yellow, grey, blue, and orange. Secondly, we bin the observed data and model predictions located in each coloured region along the radial radius from the left to right, or bottom to top. We use axisymmetry and point symmetry for both the data and model predictions to increase the number of points



**Figure 6.** The posterior marginalized distributions for the model with a cored DM halo. The points in each scatter plot show the 2D posterior marginalized distribution of the two corresponding parameters and the colour of each point indicates the relative probabilities from the low (blue) to high (red). The three ellipses in each scatter plot show the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  regions of the covariance ellipses of the distribution. The histogram in each row shows the 1D posterior marginalized distribution of the parameter and the curve is a Gaussian curve plotted with the mean and standard deviation of the distribution; the values are shown in the top. The parameters we show here, from left to right, are logarithm of the DM scale density  $\log \rho_s$ ;  $d_s = \log(\rho_s^2 r_s^3)$ , where  $r_s$  is the DM scale radius; the stellar  $M/L$  ratio scale factor  $Y$ ; the velocity anisotropy parameters for the red GCs  $\lambda^{\text{red}}$ , blue GCs  $\lambda^{\text{blue}}$  and the central stellar light  $\lambda^{\text{star}}$ ; the inner and outer region rotation parameters for the red GCs  $\kappa^{\text{red}}_{\text{inner}}$ ,  $\kappa^{\text{red}}_{\text{outer}}$ , blue GCs  $\kappa^{\text{blue}}_{\text{inner}}$ ,  $\kappa^{\text{blue}}_{\text{outer}}$ ; and the rotation parameter for the central stellar light,  $\kappa^{\text{star}}$ .

by a factor of 4 before binning. After the symmetrization, the data and model predictions in the blue and orange regions along the diagonals are the same.

For the discrete GCs, we take equal number of points in each bin and the bin size  $n = 102$ , but the size of the last bin is bigger because it contains all the remaining points. The observed mean velocity and dispersion values in each bin for the red and blue populations are calculated as follows:

$$\mu = \frac{\sum_i^n P_i^k v_i}{\sum_i^n P_i^k}, \quad (12)$$

$$\sigma^2 = \frac{\sum_i^n P_i^k (v_i - \mu)^2}{\sum_i^n P_i^k}, \quad (13)$$

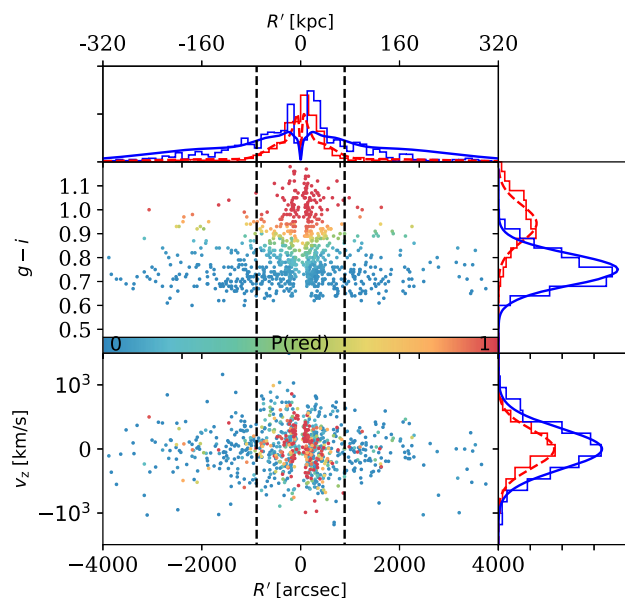
where  $n$  is the bin size and  $k$  represents the red or blue population. For the model, we just calculate the mean values of velocity and dispersion predicted at the different positions in that bin. Then, we can calculate the mean  $\chi^2$  per data point of the best-fitting model

by

$$\overline{\chi^2} = \frac{\sum_{j=1}^N \left( \frac{\mu'_j - \mu_j}{\delta \mu'_j} \right)^2 + \sum_{j=1}^N \left( \frac{\sigma'_j - \sigma_j}{\delta \sigma'_j} \right)^2}{2N}, \quad (14)$$

where  $N$  is the number of bins,  $(\mu'_j, \sigma'_j)$  represent the mean velocity and dispersion of the observed GC data in the  $j$ th bin and  $(\mu_j, \sigma_j)$  represent that of the model predicted data. Our model matches the GC data well with a mean  $\chi^2$  value of 0.94 for the red population and 0.49 for the blue population. Fig. 9 shows the resulting comparison of mean velocity and velocity dispersion profiles from the observed GC data and the model predictions along the major axis, the minor axis, and the diagonals. In each panel, points with the error bars are the data, the thick and thin solid curves are the mean and the  $1\sigma$  uncertainties of the best-fitting models.

For the stars,  $\chi^2$  comes directly from the model as shown in equation (10). The mean stellar  $\chi^2$  value from the best-fitting cored model is 0.31. The visual comparison is shown in Fig. 10. We also bin the stellar data and model predictions as in Fig. 8 and the

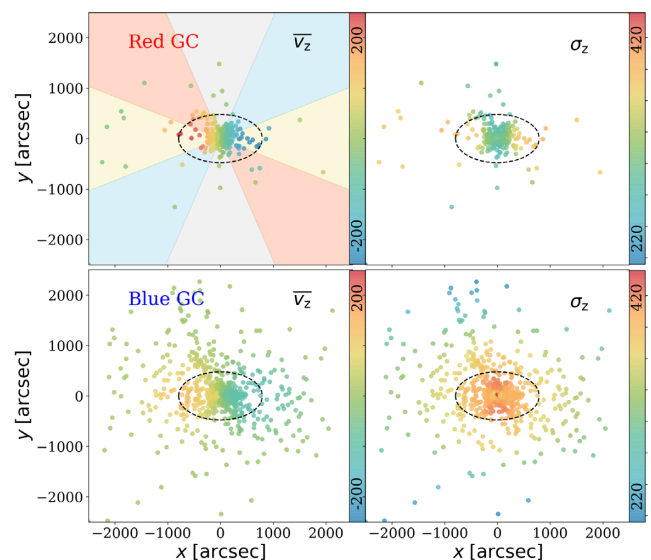


**Figure 7.** Chemo-dynamical modelling results for the spatial, colour, and kinematic distributions of GCs. **Top scatter panel:**  $R'$  versus the colour  $g - i$ . **Bottom scatter panel:**  $R'$  versus LOS velocity  $v_z$ . GCs are plotted with the points coloured by  $P_i^{\text{red}}$ , the relative probabilities of belonging to the red population, from blue (low) to red (high). In the top and right-hand panels, we show the overall spatial, colour, and velocity distributions for the red and blue populations. The histograms of the red and blue populations are calculated from all kinematic GC data points weighted by  $P_i^{\text{red}}$ , and  $P_i^{\text{blue}}$ . The curves for the density distributions are from the best-fitting MGEs for the photometric red (dashed) and blue (solid) GCs. The curves for the colour and velocity distributions of red/blue GCs are from the best-fitting chemo-dynamical models. The histograms and the corresponding curves of density and colour distribution constructed from the kinematic data are generally consistent with the photometric data. The density curves and histograms are depressed in the inner region due to incompleteness in the photometric and kinematic data. The two vertical dashed lines indicate the position with elliptical  $x$ -axis radii of  $-5 R_e^{\text{maj}}$  and  $5 R_e^{\text{maj}}$ .

number of points per bin is 20. Note that we have symmetrized the data points to reduce noise, and so the data bins in Figs 9 and 10 are not independent.  $\chi^2$  will increase by a factor of 2 if we model without data symmetrization, and are then using independent bins.

For the velocity profiles, the model matches the data to within  $1\sigma$  uncertainties for both the GCs and the stars in the central galaxy. As shown in the first and third left-hand panels of Fig. 9, the red GCs have moderate rotation around the minor axis with  $V_{\text{max}}/\sigma \sim 0.4$  at  $R \sim 40$  kpc, and the blue GCs have a weak rotation around the minor axis with  $V_{\text{max}}/\sigma \sim 0.1$ , and our model matches the rotation for each of the GC populations. The observed integrated starlight has no significant rotation around the minor axis, and our model prediction agrees with this.

For the velocity dispersion profiles, the model generally matches the radial profiles of both the GCs and the stars in the central galaxy, except for the red GCs along the diagonals. This explains why we obtained a higher mean  $\chi^2$  for the red GCs. The dispersions along the major and minor axes are different for the two populations, with the red GCs having a larger dispersion along the major axis than that along the minor axis and the blue GCs behaving in just the opposite way. Our model matches the trend for both GC populations.



**Figure 8.** Mean velocity ( $\bar{v}_z$ ) and velocity dispersion ( $\sigma_z$ ) maps of red (upper panels) and blue (lower panels) GCs from the best-fitting model with a cored DM halo. Each point represents the position of a GC in the data sample, colours indicate the value of  $\bar{v}_z$  and  $\sigma_z$  as scaled by the colour bars. On the top left-hand panel, we illustrate the binning scheme we use for a direct comparison of the data and model predictions. The projected plane is divided into eight sectors and the regions along the major axis, minor axis and the  $45^\circ$  and  $135^\circ$  diagonals are filled with yellow, grey, blue, and orange. The data located in each coloured region are binned along the radial radius from the left to right, or bottom to top. The  $x$ -axis is along the major axis and  $y$  the minor axis. The dashed ellipse in each panel has an  $x$ -axis radius of  $5 R_e^{\text{maj}}$ .

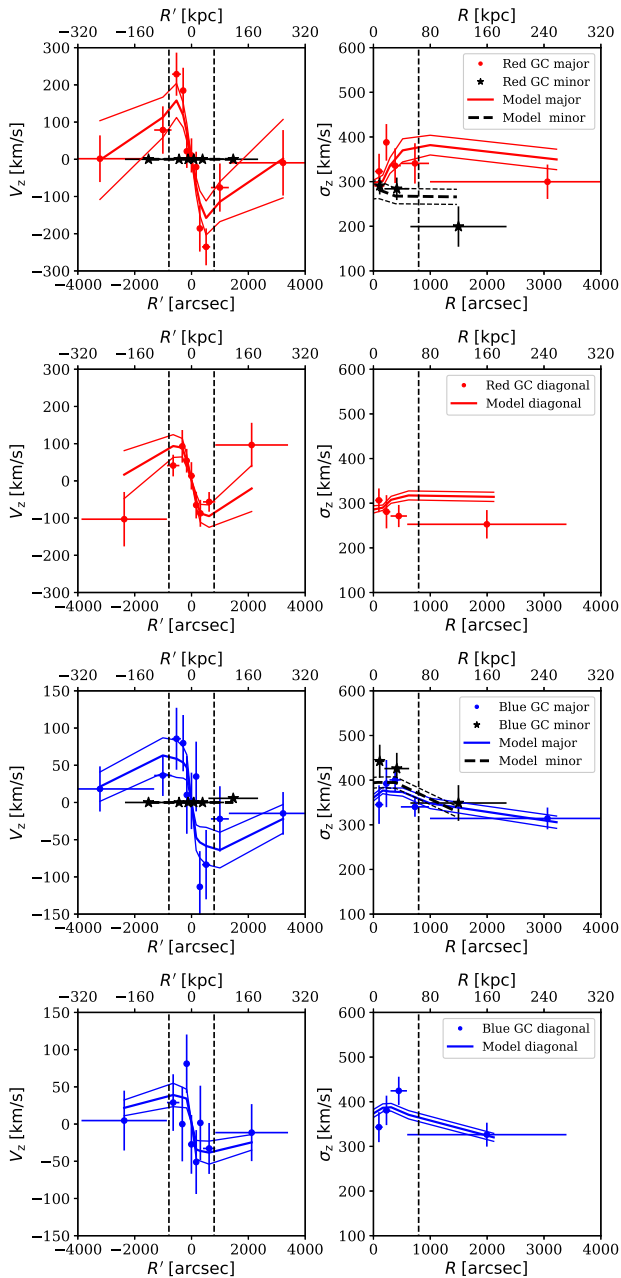
### 5.3 Internal rotation

As shown in Figs 9 and 10, only the red GCs show an obvious rotation around the minor axis, with  $V_{\text{max}}/\sigma \sim 0.4$ , which is matched by our axisymmetric models. Here, we check whether the rotation axis of red GCs is really aligned with the minor axis.

Fig. 11 shows the positions of the red GCs coloured by their LOS velocities. The plus symbols represent the red GCs coloured by the measured LOS velocity, while the size of each symbol is scaled with the absolute value of its LOS velocity. The vertical black line is the photometric minor axis of the galaxy. From a visual inspection of Fig. 11, it should be apparent that the rotation direction of the red GCs is roughly around the photometric minor axis.

The variation of the LOS velocities of the red GCs along the azimuthal angle  $\phi$  measured from the major axis is shown in Fig. 12. We fit the velocities of the red GCs within  $5 R_e^{\text{maj}}$  with a sinusoid  $v = v_{\text{max}} \sin(\phi + \phi_0)$  and obtain  $\phi_0 = -91^\circ \pm 3^\circ$  and  $v_{\text{max}} = (114 \pm 6) \text{ km s}^{-1}$ .  $\phi_0 = \pm 90^\circ$  means rotation exactly around the minor axis. The values we obtained are consistent with minor axis rotation to within  $1\sigma$  error. Consequently, an axisymmetric assumption is reasonable, and our model fits the rotation of the red GCs well. Red GCs may have a velocity twist outside  $5 R_e^{\text{maj}}$  as shown in the velocity map of Zhang et al. (2015), however, there are only a few red GCs outside  $5 R_e^{\text{maj}}$  as can be seen in Fig. 11.

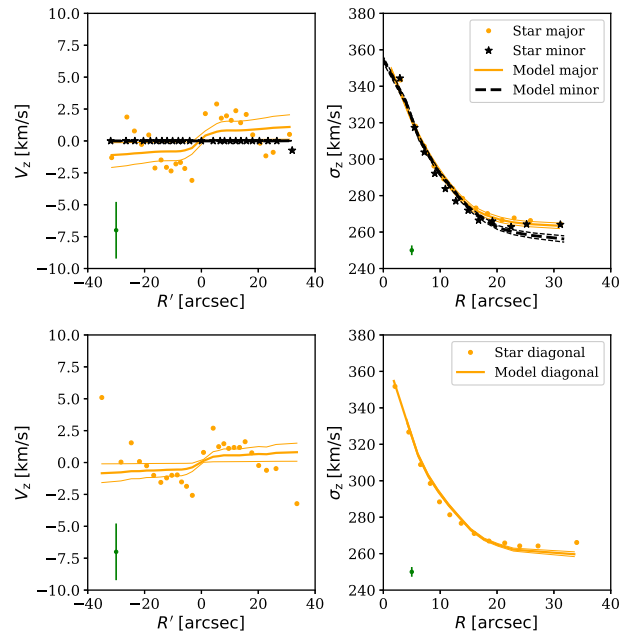
Our red GCs have the same rotation axis and rotation direction as M87's planetary nebulae (PNe; Longobardi et al. 2018a). PNe have their strongest rotation in the region of  $R \sim 500\text{--}1000$  arcsec (Longobardi et al. 2018a), but with  $V_{\text{max}}/\sigma \sim 0.1$  that is smaller than we found for red GCs.



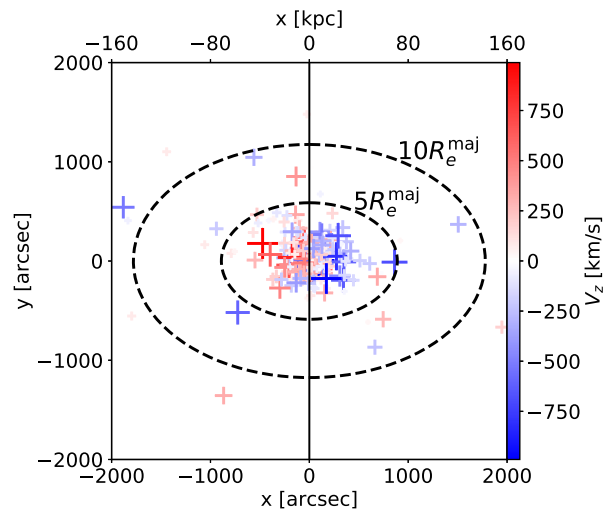
**Figure 9.** Comparison of the observed GC data with the best-fitting predictions from the cored DM halo model. In each panel, points with error bars are the data, thick and thin solid curves are the mean and  $1\sigma$  uncertainties of best-fitting models.  $R'$  is increasing from the left to right, or bottom to top for each binning region as shown in Fig. 8. The vertical dashed lines in the panels mark the elliptical  $x$ -axis radius of  $5R_e^{\text{maj}}$ .

#### 5.4 Velocity dispersion anisotropy

The velocity dispersion anisotropy parameter  $\beta_r = 1 - (\sigma_\phi^2 + \sigma_\theta^2)/2\sigma_r^2$  is widely used as an indication of underlying orbit distribution:  $\beta_r > 0$  indicates radially biased anisotropy,  $\beta_r < 0$  tangentially biased anisotropy, and  $\beta_r = 0$  an isotropic distribution (p.294; Binney & Tremaine 2008). In axisymmetric JAM models,  $\beta_r$  is not used as an explicit free parameter, but we can calculate it as a function of radius from the intrinsic first and second velocity moments predicted by the model. Fig. 13 shows the velocity anisotropy profiles obtained from our best-fitting models. The red

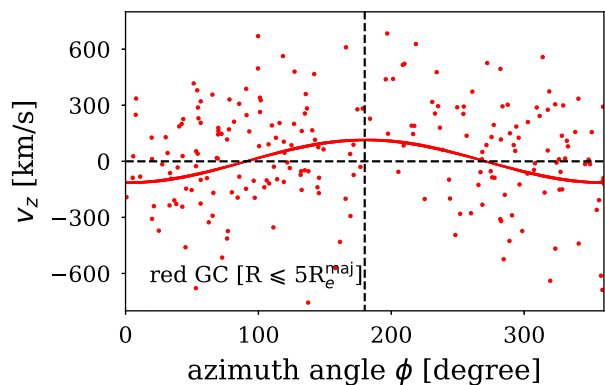


**Figure 10.** Comparison of the stellar data in the central galaxy with the best-fitting predictions from the cored DM halo model. In each panel, points are the data, the thick and thin solid curves are the mean and  $1\sigma$  uncertainties from the best-fitting model. The green points with the vertical error bars in the bottom left-hand corners of the panels indicate the mean  $1\sigma$  uncertainties of the binned observed data.

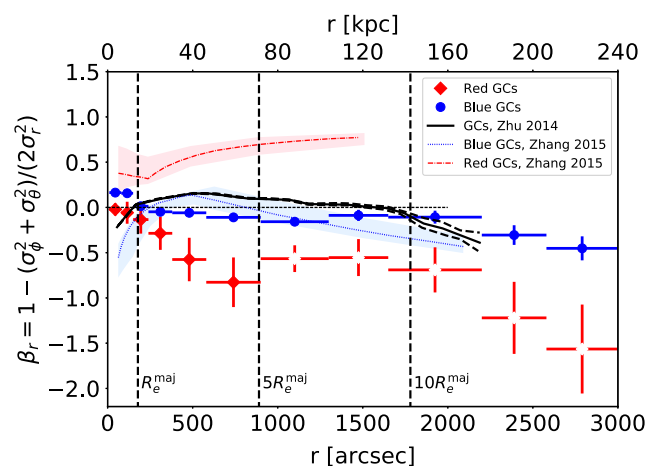


**Figure 11.** The positions of the red GCs identified by the cored model.  $x$  is along the major axis and  $y$  the minor axis. The plus symbols represent the red GCs coloured by the measured LOS velocity, while the size of each symbol is scaled by the absolute value of its LOS velocity. The two dashed ellipses have  $x$ -axis radii of  $5R_e^{\text{maj}}$ ,  $10R_e^{\text{maj}}$ . The vertical black line is the photometric minor axis of the galaxy.

diamonds and the blue circles with the error bars represent the velocity anisotropy for our GCs from our model with a cored DM halo. There are few red GCs at  $R > 1000$  arcsec, and our model has difficulty matching the dispersion along diagonals for red GCs in this region as shown in Fig. 9. Thus, the anisotropy for red GCs at  $r > 1000$  arcsec may be larger than the errors we obtain from the model scatter. We mark these points with the open symbols in the



**Figure 12.** The LOS velocity as a function of azimuthal angle  $\phi$  from the major axis. The points are the red GCs and the overlay curve is the corresponding best-fitting sine curve  $v = (114 \pm 6) \times \sin(\phi - (91^\circ \pm 3^\circ))$  [km s $^{-1}$ ]. The vertical dashed line marks the azimuth angle of  $180^\circ$ .



**Figure 13.** The velocity dispersion anisotropy  $\beta_r$  profiles. The red diamonds and the blue circles with the error bars are the profiles of the red and blue GCs from the best-fitting models with a cored DM halo.  $\beta_r$  for red GCs at  $r > 1000$  arcsec should have large uncertainties and we mark these points with the open symbols. The error bars indicate the scatter of  $\beta_r$  in the models within a  $1\sigma$  confidence level. The dash-dotted red and dotted-blue curves are the results obtained by the spherical Jeans models from Zhang et al. (2015) and the red and blue regions mark the  $1\sigma$  confidence intervals. The solid black curve is the profile obtained by the made-to-measure method from Zhu et al. (2014) and the dashed curves are the corresponding uncertainties. Note that the profile from Zhu et al. (2014) is for all GCs. The three vertical dashed lines mark radii of (1, 5, 10)  $R_e^{maj}$ .

figure. The cusped model also predicts an almost identical velocity anisotropy profile, and we show it in the Appendix A.

According to our model, the red GC system is nearly isotropic in the very inner region, becoming significantly tangential in the outer region. The blue GC system has slightly radial velocity anisotropy in the inner region, becoming slightly tangential in the outer region. The difference between the red and blue GCs velocity anisotropy could be indicated by the velocity dispersion difference along the major and minor axes. The red GCs have a larger dispersion along the major axis, while it is the opposite for the blue GCs. This results in the different velocity anisotropy we obtained for the two populations.

The dash-dotted red and dotted blue curves in Fig. 13 are the results obtained by the spherical Jeans model from Zhang et al.

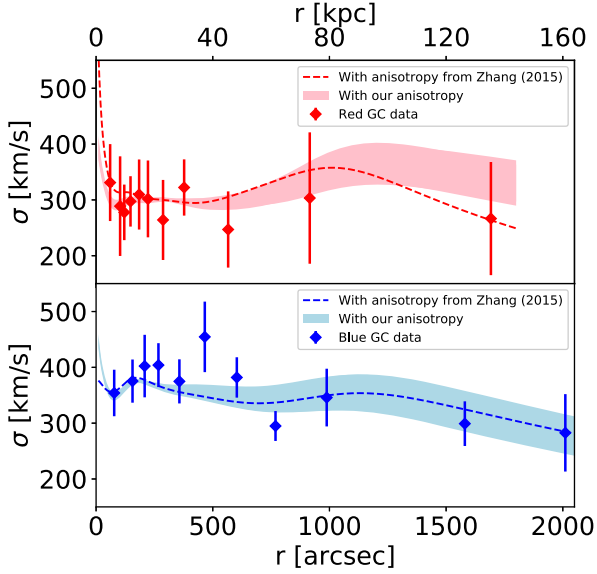
(2015) with the red and blue regions marking the  $1\sigma$  confidence intervals. Zhang et al. (2015) divided GCs into red and blue population in advance by colour  $g - i = 0.89$  and constructed two spherical Jeans models for the red and blue GCs. The solid black curve is the profile obtained by the made-to-measure method from Zhu et al. (2014) and the dashed curves are the corresponding uncertainties. Zhu et al. (2014) treated GCs as one population in the model.

The spherical Jeans model from Zhang et al. (2015) predicted strong radial anisotropy for the red GCs overall and slightly tangential anisotropy for the blue GCs. Oldham & Auger (2018) obtained  $\beta_r \sim 0.32$  for red GCs and  $\beta_r \sim 0.26$  for blue GCs that we do not show in the figure. The velocity anisotropy profiles for the blue GCs in the literature are roughly consistent with our results within  $1\sigma$  confidence in the outer regions beyond  $1R_e^{maj}$ . For the red GCs, the radial anisotropy obtained by spherical Jeans models in the literature are not consistent with each other, and not consistent with our results.

Given the apparent inconsistency of the velocity anisotropy of the red GCs we obtained, especially when compared with that of Zhang et al. (2015), we perform a simple test to check whether our velocity anisotropies are consistent with the data if we also make a spherical symmetry assumption.

We construct spherical Jeans models for the red GCs and blue GCs of M87, using the total mass profiles from our axisymmetric Jeans models, the velocity dispersion anisotropy  $\beta_r^{\text{red}}$ ,  $\beta_r^{\text{blue}}$  from our models or Zhang et al. (2015; shown in Fig. 13), and the surface number density profiles of red and blue GCs (shown in Fig. 1) as input. We still use MGEs to describe the surface number density of each population. We impose a radial variation of  $\beta_r$  in our spherical Jeans models by setting  $\beta_k = \beta_r(r = \sigma_k)$  for each Gaussian  $k$  with its dispersion  $\sigma_k$  (Cappellari 2008). We then compare the model predictions from the spherical Jeans models with the observed velocity dispersions as shown in Fig. 14. The pink and light blue regions are the velocity dispersion predictions from the spherical Jeans models after imposing our velocity anisotropy profiles for the red and blue GCs, respectively. The red- and blue-dashed curves are the predictions from the spherical Jeans models using velocity anisotropy profiles from Zhang et al. (2015). Although the model predictions using our velocity anisotropy do not match the data for red GCs in the outmost region as well as those models with the velocity anisotropy taken from Zhang et al. (2015), our model still generally matches the data points within  $1\sigma$  uncertainties. An exception is that the observed dispersion of the blue GCs at  $\sim 400$  arcsec (32 kpc) is about 20 per cent higher than the model prediction, and is likely due to a substructure (Romanowsky et al. 2012) in the blue GCs in this region.

In a spherical Jeans model, the tangential velocity dispersion  $\sigma_t$  and the radial velocity dispersion  $\sigma_r$  are sensitive to  $\beta_r$ . For a system with a fixed mass profile  $M(r)$ , when  $\beta_r$  increases,  $\sigma_r$  increases and  $\sigma_t$  decreases (see equations 4–55 of Binney & Tremaine 1987). Both  $\sigma_r$  and  $\sigma_t$  contribute to the velocity dispersion along the LOS. The LOS velocity dispersion does not depend strongly on  $\beta_r$ , as the variation of  $\sigma_r$  and  $\sigma_t$  partly cancel each other (see equations 4–60 of Binney & Tremaine 1987). The dependence of the LOS velocity dispersion on  $\beta_r$  is affected by the tracer density profile. With the tangential anisotropy we obtained for red GCs, very different from the radial anisotropy obtained by Zhang et al. (2015), our tests show that the dispersion obtained from the spherical Jeans models could still be consistent with the LOS velocity dispersion from the data. This indicates that the red GCs velocity anisotropy profiles from spherical models could contain large uncertainties,



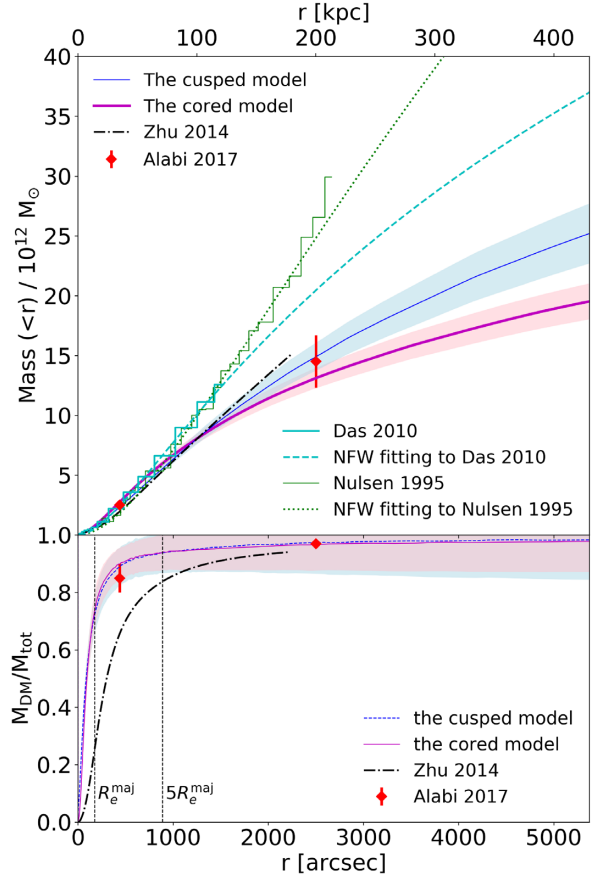
**Figure 14.** Comparison of the observed velocity dispersions with the predictions by the spherical Jeans models. The red and blue GCs are classified by the best-fitting chemo-dynamical model as mentioned above. The points with the error bars are the observed velocity dispersion profiles of the red (top) and blue (bottom) GCs, and they are binned along the spherical radius. The pink and light blue regions are the predictions for the velocity dispersions with  $1\sigma$  uncertainties by the spherical Jeans models after imposing our velocity anisotropy profiles for the red and blue GCs, respectively. The red- and blue-dashed curves are the predictions by the spherical Jeans models with the velocity anisotropy profiles from Zhang et al. (2015).

and perhaps the error bars in Zhang et al. (2015) have been underestimated. Axisymmetric Jeans models are most useful for stronger anisotropy constraints, which are provided mainly via the azimuthal information in the kinematics. We have discussed this by showing the difference in velocity dispersions along the major and minor axes in the second paragraph of this section.

### 5.5 Mass profile

We show in Fig. 15 the total mass profiles and the DM fractions obtained by our best-fitting models, with blue and magenta representing the cusped and the cored models. The coloured regions around the solid curves indicate the corresponding  $1\sigma$  uncertainties of the total mass profile. The bottom panel shows the DM fraction. The light blue and pink regions around the solid curves represent the  $1\sigma$  uncertainties of the DM fraction from the cusped and cored models.

The total mass profiles from our two models are consistent with each other to within  $1\sigma$  uncertainties up to  $\sim 220$  kpc. The difference in the mass profiles becomes obvious beyond 220 kpc. By comparing the observed kinematics with model predictions from the cored (Fig. 9) and cusped (Fig. A1) models, we see that the cored model matches the data slightly better than the cusped model beyond 150 kpc for the red GCs and the corresponding mean  $\chi^2$  of the cored model is slightly lower. However, the preference of the cored model is not statistically significant due to the small number of data points at large radii. As the scale factors  $\Upsilon$  obtained by the two sets of models are close, the stellar mass profiles are almost the same and the DM fraction profiles from our two models are also consistent. The DM halo contributes  $(73 \pm 8)$  per cent of the total



**Figure 15.** Total mass profile and DM fraction. **Top panel:** The total mass profiles. The solid blue, red, and dash-dotted black curves represent the total mass profiles obtained by our models with a cusped DM halo, with a cored DM halo and by the made-to-measure method from Zhu et al. (2014). The coloured regions around the solid curves represent the  $1\sigma$  uncertainties of the total mass. The red diamonds mark the total mass obtained from Alabi et al. (2017). The cyan step curve represents the mass profile obtained from hot gas emission based on *Chandra* observations (Das et al. 2010), while the green is based on *ROSAT* observations (Nulsen & Bohringer 1995). The step curves represent the middle values of their measurements within the data coverage, and the corresponding dotted and dashed curves are NFW profiles fitted to the step curves. **Bottom panel:** The DM fraction. The coloured light blue and pink regions around the solid curves represent the  $1\sigma$  uncertainties of the DM fraction from the cusped and cored models. The red diamonds mark the DM fraction obtained from Alabi et al. (2017). The two vertical dashed lines indicate radii of  $1 R_e^{\text{maj}}$ ,  $5 R_e^{\text{maj}}$ .

mass within  $1 R_e^{\text{maj}}$  (14.2 kpc), and  $(94 \pm 6)$  per cent within  $5 R_e^{\text{maj}}$  (71.2 kpc).

With data extending to  $\sim 400$  kpc, we obtain total mass values of  $M_{\text{tot}} = (1.95 \pm 0.17) \times 10^{13} M_{\odot}$  for the cored model and  $M_{\text{tot}} = (2.32 \pm 0.22) \times 10^{13} M_{\odot}$  for the cusped model. Incorporating the uncertainties in the cored and cusped models, we conclude  $M_{\text{tot}} = (2.16 \pm 0.38) \times 10^{13} M_{\odot}$  within  $\sim 400$  kpc for M87.

The red diamonds at different radii in Fig. 15 mark the total mass and DM fraction obtained by Alabi et al. (2017). They obtained their mass estimates using a tracer mass estimator (Watkins, Evans & An 2010) and a stellar  $M/L$  ratio that accounts for stellar age variations, along with GC kinematic data from the SLUGGS survey (Brodie et al. 2014) and Strader et al. (2011). Their results are consistent with ours as may be seen from Fig. 15. They found that the DM

fractions are  $(85 \pm 5)$  per cent within 35 kpc, and  $(97 \pm 1)$  per cent within 200 kpc.

The black dash-dotted curves in Fig. 15 represent the total mass profile and DM fraction obtained by Zhu et al. (2014). Their mass estimates were determined using the made-to-measure method with GC kinematic data as described in Section 3.2. Comparing with our results, the total mass profile from Zhu et al. (2014) is consistent with the  $1\sigma$  uncertainties inside  $\sim 180$  kpc, but, in the inner region, the stellar mass profile is about twice as high as ours and the DM mass profile is lower. As shown in the bottom panel, the DM fraction from Zhu et al. (2014) is  $\sim 26$  per cent within  $1 R_{\text{e}}^{\text{maj}}$  (14.2 kpc) and  $\sim 84$  per cent within  $5 R_{\text{e}}^{\text{maj}}$  (71.2 kpc) that are significantly smaller than our inner region predictions. This can be explained by the fact that Zhu et al. (2014) assumed a constant  $M/L_r$  ratio, whereas we modelled with a varying  $M/L$  ratio constructed by stellar population synthesis. The DM contribution is less than 17 per cent with the cusped DM and 7 per cent with the cored DM in the inner 25 arcsec ( $\sim 2$  kpc). To match the total mass (mainly stellar mass) in the very inner regions, a model with a constant stellar  $M/L$  ratio will require a higher overall  $M/L$  ratio, which leads to more stellar mass in the outer regions ( $r \sim$  a few hundred arcsec) where both stellar mass and DM are important (see Fig. 5).

The cyan step curve in Fig. 15 represents the mass profile obtained from hot gas X-ray emissions observed by *Chandra* (Das et al. 2010), and the green is based on *ROSAT* observations (Nulsen & Bohringer 1995). M87 has also been observed to large radii by *XMM-Newton* (Urban et al. 2011). Their temperature profile of hot gas is consistent with that from the *ROSAT* observations, and thus similar mass profiles are expected. The step curves represent the middle values of the measurements within the data coverage, and the corresponding dotted and dashed curves are NFW profiles fitted to the step curves. The mass profile from Nulsen & Bohringer (1995) has a large uncertainty with the lower boundary consistent our results. Within the data coverage (from 0 to 200 kpc), the mass obtained from X-ray emitting hot gas is  $\sim 20$ – $100$  per cent larger than our mass profile obtained from GC kinematics. There will be further discussion of X-ray mass profiles in Section 6.

## 6 DISCUSSION

### 6.1 Uncertainties on mass profiles

In the early literature, it was proposed that the DM around M87 is associated with the whole of Virgo (McLaughlin 1999), as a universal NFW halo can fit the kinematics of GCs (within 100 kpc), dwarf galaxies (100 kpc  $< R < 2$  Mpc) and X-ray gas simultaneously. The mass profile presented by McLaughlin (1999) is generally consistent with the mass profile derived from X-ray emission. However, with the increase in the numbers of GC LOS velocity measurements, it has become obvious that GCs (Longobardi et al. 2018b) and PNe (Longobardi et al. 2018a) have much smaller velocity dispersions than dwarf galaxies (McLaughlin 1999; Ko et al. 2017) in the region of  $100 < R < 200$  kpc. A kinematically distinct intracluster population of GCs has been identified (Longobardi et al. 2018b), which has a dispersion of  $\sim 600$  km s $^{-1}$ , consistent with the velocity dispersion of the 17 GCs identified in Section 3.2, as well as the dispersion of the dwarf galaxies. A similar intracluster component has also been identified for PNe (Longobardi et al. 2018a).

The mass obtained from X-ray gas  $M(< 400 \text{ kpc}) = 5.2 \times 10^{13} M_{\odot}$  (Nulsen & Bohringer 1995; Das et al. 2010; Urban et al. 2011) and that of  $M(< 400 \text{ kpc}) = 8.5 \times 10^{13} M_{\odot}$  from dwarf galaxy kinematics (McLaughlin 1999) is larger than our value of

$M(< 400 \text{ kpc}) = (2.16 \pm 0.38) \times 10^{13} M_{\odot}$ . Our mass is about a factor of  $\sim 2.5$  smaller than that from X-ray gas, and a factor of  $\sim 4$  smaller than that from dwarf galaxy kinematics.

Due to the fact that Virgo cluster may not be fully relaxed (Binggeli et al. 1987; Urban et al. 2011), M87 seems to be located at the dynamical centre of one subclump of galaxies. The differences in mass profiles obtained from different tracers may be caused by the unrelaxed state. Dwarf galaxies, affected by the formation of the cluster, have large velocity dispersions that are consistent with the intracluster GCs. The X-ray mass profile is also significantly larger than that indicated by GCs at  $R > 100$  kpc. It is likely that dwarf galaxies and X-ray gas are more relaxed within the gravitational potential of the outer regions of the cluster, but not in the disturbed central regions (Wood et al. 2017). Our study, however, has attempted to carefully remove the intracluster GCs (see Fig. 2), and thus the remaining GCs are supposed to be well relaxed with the gravitational potential of the central galaxy (M87).

We use JAM models to describe the kinematics of stars, red GCs, and blue GCs in M87. The GC systems extending to large radii are oblate and have rotation around the photometric minor axis, and thus suit the oblate assumption of JAM modelling. The stellar component of M87 could actually be triaxial with a kinematic decoupled core (Emsellem et al. 2014). Various tests show that JAM modelling could still work well for estimating the total mass distributions of triaxial and prolate systems with  $\sim 20$  per cent uncertainty (Li et al. 2016). The DM mass dominates in the outer regions and is mainly constrained by the GC kinematics. We presume the uncertainty in the mass profile in the outer regions caused by triaxiality of the stellar component should be minor.

Even with 894 GCs, the data are still poorly sampled in the outer regions (see Fig. 2). We have checked that the asymmetry in velocity distribution in the outer regions is unlikely to be caused by Milky Way stars from their photometry properties. It is hard to tell if they are actually caused by low number statistics, or contamination of foreground GCs in the Virgo cluster, where the latter could potentially cause an overestimate of velocity dispersion, thus affecting the mass profile. More data points at  $R > 2000$  arcsec are needed to investigate this scenario. There is also an indication of substructure (Romanowsky et al. 2012) in the GC system around  $R \sim 500$  arcsec, which ideally should be cleaned before the modelling. We note that our models are not trying to match this ‘bump’ in velocity dispersion, but mostly matched the velocity dispersion profiles in the outer regions. In this sense, the effect of this ‘substructure’ on our results should be small. All these effects on the data may actually cause extra systematic uncertainty rather than the statistical uncertainty to our results. In this sense, we could still underestimate the uncertainty of the masses we reported.

As discussed in Section 3.1, the MGE constructed surface density profile for red GCs may have relatively large uncertainties in the northern region at  $R > 1000$  arcsec. Our MGE model fits the southern side, and so we may have created a too steep surface density slope. This may cause an overestimation of the enclosed mass with fixed kinematics. In our case, there are 23 red GCs with elliptical  $x$ -axis radii  $R > 1000$  arcsec, including only eight in the northern regions, thus the mass profile in the outer regions is dominated by fitting to the blue GCs (as can be seen in Fig. 9, the kinematics of the blue GCs are fitted well in the outer regions, but not for the red GCs). We therefore presume that the red GCs surface density profile uncertainties have only a minor effect on the overall mass profile estimation.



We have few constraints on the DM inner slope, and no significant preference is found for a cored or cusped DM halo. A cored DM is preferred in Oldham & Auger (2018). Note that in their work, all components are modelled with radially anisotropic spherical models (constant  $\beta_r$  for red and blue GCs, and a scaled Osipkov–Merrit profile for stars which is isotropic at  $r = 0$  becoming radially anisotropic with a constant value at infinity). These are strong assumptions and act to narrow down the statistical uncertainty on their mass profiles. Stronger assumptions in models can cause larger systematic uncertainties, which are usually poorly evaluated. We have allowed our velocity anisotropy and rotation parameters to be totally free in our axisymmetric models. This gives more freedom in the determination of mass profiles, and naturally causes relatively large statistical uncertainties on the gravitational potential parameters including the DM density slope  $\gamma$ .

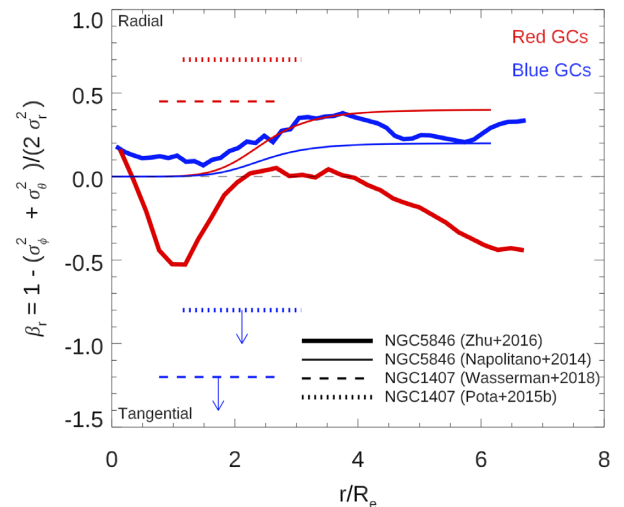
We have included the  $M/L$  gradient from SPS Sarzi et al. (2018). However, the scale parameter we obtained is smaller than 1, with the stellar mass  $M^{\text{dyn}}/M^{\text{SPS}} \sim 0.62 - 0.68$  regardless of the DM model. Note that DM contributes less than  $\sim 7$  per cent with the cored DM in the inner  $\sim 2$  kpc, so a higher stellar mass can not be accommodated in our model. Since the stellar system of M87 is triaxial, the uncertainty in the total mass could be caused by applying a JAM model to a triaxial system. Li et al. (2017) assessed the mass uncertainty in such a case to be  $\sim 20$  per cent, but this would not fully resolve the discrepancy here. Note that our result is consistent with Oldham & Auger (2018), who obtained a similar  $M/L$  gradient independently from dynamical models, but with a systematically lower ratio than Sarzi et al. (2018). Our result indicates that the  $M/L_r$  from SPS (Sarzi et al. 2018) might be overestimated, especially in the very inner regions. This would cause an overestimate of stellar mass and an underestimate of DM in the very inner regions in our model. However, overall, it should have limited impact on the whole DM profile and on the fit to the GC kinematics.

## 6.2 Velocity anisotropies of GCs in other galaxies

The velocity anisotropy  $\beta_r$  values obtained by different methods are not always consistent with each other. As shown in Fig. 13, we obtain tangential anisotropy for the red GCs of M87, while the two-component spherical Jeans model yields a highly radial anisotropy  $\beta_r \sim 0.8$  in Zhang et al. (2015) and medium radial anisotropy  $\beta_r \sim 0.3$  in Oldham & Auger (2018). Also, highly radial anisotropy was obtained for the red GCs of NGC 1407 using a similar two-component spherical Jeans model (Pota et al. 2015; Wasserman et al. 2018).

In Fig. 16, we summarize the velocity anisotropy profiles of red and blue GCs obtained by various two-component dynamical models in the literature: NGC 5846 from Zhu et al. (2016) obtained by discrete chemo-dynamical JAM models and confirmed with the made-to-measure method, NGC 5846 from Napolitano et al. (2014) obtained using kurtosis, NGC 1407 obtained by two-component spherical Jeans model from Wasserman et al. (2018) and Pota et al. (2015).

Highly radial anisotropy of red GCs and highly tangential anisotropy of blue GCs are obtained for NGC 1407 using a similar two-component spherical Jeans model (Pota et al. 2015; Wasserman et al. 2018) as Zhang et al. (2015) used for M87. Unlike an axisymmetric JAM model, with a spherical Jeans model, the information on anisotropy in the data itself across the 2D observational plane is actually washed out. The velocity anisotropies are obtained by the balance of the multiple tracer populations with different number density slopes and velocity dispersions. The red GCs have



**Figure 16.** The velocity anisotropy profiles of red and blue GCs obtained by various two-component dynamical models in the literature: NGC 5846 from Zhu et al. (2016) obtained by discrete chemo-dynamical JAM models and confirmed with the made-to-measure method (the thick solid lines), NGC 5846 from Napolitano et al. (2014) obtained using kurtosis (the thin solid lines), NGC 1407 obtained by two-component spherical Jeans models from Wasserman et al. (2018; the dashed lines) and Pota et al. (2015; the dotted lines). The red and blue lines represent  $\beta_r$  of red and blue GCs, respectively. Fig. 13 shows the anisotropy profiles of M87 for comparison.

smaller velocity dispersions than blue GCs, both in M87 and NGC 1407, which we presume is the direct reason for obtaining radial anisotropy for red GCs and tangential anisotropy for the blue GCs through the two-component spherical Jeans models. The GC systems have rotation and different velocity dispersions along the major and minor axes, both in M87 and NGC 1407. These systems are more axisymmetric and deviate more from the spherical assumption, thus the uncertainty of the velocity anisotropy obtained by this spherical Jeans method should be large. As described in Section 5.4, we suggest that the uncertainty in velocity anisotropy obtained from the spherical Jeans modelling could be heavily underestimated.

The velocity anisotropy profiles we obtained here for M87 by discrete chemo-dynamical JAM models are similar to those of NGC 5846 obtained by Zhu et al. (2016), where blue GCs are near isotropic and red GCs are tangentially anisotropic. They are consistent with the scenario that blue GCs are mostly accreted, and that red GCs, even at very large radii, could either have been born in situ, or have been heavily disrupted on radial orbits. Red GCs have qualitatively similar rotation and velocity dispersion profiles to PNe out to  $\sim 2000$  arcsec (Longobardi et al. 2018a), which supports the *in situ* formation history. The near isotropic blue GCs and tangential red GCs suggest that there could be significant GC disruption even at very large radii. GC disruption (Vesperini & Zepf 2003; Vesperini et al. 2013; Brockamp et al. 2014) is an interesting open question requiring further investigation. It may be better understood by considering the formation and disruption of GCs within cosmological simulations. Ramos-Almendares et al. (2019) using the Illustris simulation (Vogelsberger et al. 2014) provides a good reference on how to tag intracluster GCs in simulations. They found that intracluster GCs are expected to be highly radially biased ( $\beta \geq 0.3$ ) at all radii within the cluster. The slightly tangential anisotropy for blue GCs could be partially caused by localized

substructures because there is a known substructure at  $\sim 50\text{--}100$  kpc (Romanowsky et al. 2012), which we did not remove.

## 7 SUMMARY

In Section 1, the Introduction, we set out our objectives for our investigation. We have been successful in that we have met all of the objectives. We now have

- (1) a new mass estimate for M87 extending to  $\sim 400$  kpc,
- (2) estimates of the velocity dispersion anisotropy of M87's two populations of GCs, and
- (3) an assessment of the formation and capture of the GC populations that is consistent with established formation/capture scenarios.

We find that the mass profiles are well constrained by the central galaxy stellar kinematics and the GC kinematics from the innermost to the outermost region. Both cusped and cored DM haloes are found to be acceptable. The total mass within  $5 R_{\text{e}}^{\text{maj}}$  (71.2 kpc) is  $(5.8 \pm 0.5) \times 10^{12} M_{\odot}$ . The DM halo contributes  $(73 \pm 8)$  per cent of the total mass within  $1 R_{\text{e}}^{\text{maj}}$  (14.2 kpc), and  $(94 \pm 6)$  per cent within  $5 R_{\text{e}}^{\text{maj}}$  (71.2 kpc). Our estimates benefit from both the large number of GCs we have used and the large spatial distribution of the GCs around the central galaxy. Estimates for the central region benefit from the use of a radially varying  $M/L$  ratio taken from Sarzi et al. (2018).

With data extending to  $\sim 400$  kpc, we obtain total mass values of  $M_{\text{tot}} = (1.95 \pm 0.17) \times 10^{13} M_{\odot}$  for the cored model and  $M_{\text{tot}} = (2.32 \pm 0.22) \times 10^{13} M_{\odot}$  for the cusped model. Incorporating the uncertainties in the cored and cusped models, we conclude  $M_{\text{tot}} = (2.16 \pm 0.38) \times 10^{13} M_{\odot}$  within  $\sim 400$  kpc for M87.

The red GCs present no anisotropic dispersion bias in the very inner region, and become tangentially anisotropic as the radius increases. The blue GC system varies from slightly radially to slightly tangentially anisotropic from the inner to outer region. Considering GC formation, these anisotropy values are consistent with the scenario that blue GCs are mostly accreted. The red GCs, even at very large radii, could either have been born *in situ*, or have been heavily tidally disrupted by being on radial orbits (Brockamp et al. 2014; Wu et al. 2014).

For the future, the method we have used from Zhu et al. (2016) has performed well and we intend to use it on other galaxies with extensive GC systems.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for helping to improve the paper. The models were accomplished on the ‘Zen’ cluster at National Astronomical Observatories, Chinese Academy of Sciences (NAOC), and the ‘Venus’ cluster at Tsinghua University. We acknowledge useful discussions with Eric Emsellem and Hongyu Li. This work is partly supported by the National Key Basic Research and Development Program of China (Grant No. 2018YFA0404501 to SM), by the National Science Foundation of China (Grant Nos 11821303, 11333003, 11390372, and 11761131004 to SM). LZ acknowledges support from Shanghai Astronomical Observatory, Chinese Academy of Sciences under Grant No. Y895201009. GvdV acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under Grant No. 724857 (Consolidator Grant ArcheoDyn). AL acknowledges the support by the Centre National d’etudes Spatiales, CNES.

## REFERENCES

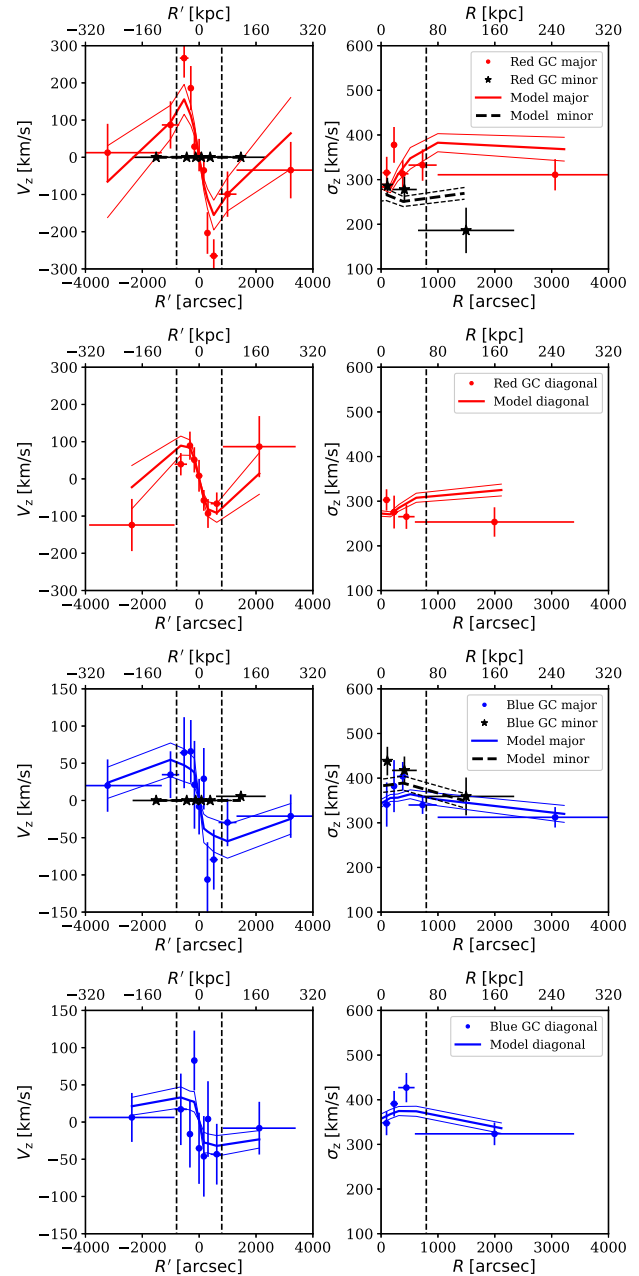
- Agnello A., Evans N. W., Romanowsky A. J., Brodie J. P., 2014, *MNRAS*, 442, 3299
- Alabi A. B. et al., 2017, *MNRAS*, 468, 3949
- Amorisco N. C., 2018, *MNRAS*, 482, 2978
- Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379
- Binggeli B., Tammann G. A., Sandage A., 1987, *AJ*, 94, 251
- Binney J., Tremaine S., 1987, *Galactic dynamics*. Princeton Univ. Press, Princeton, NJ
- Binney J., Tremaine S., 2008, *Galactic Dynamics*, 2nd edn. Princeton Univ. Press, Princeton, NJ
- Blakeslee J. P. et al., 2009, *ApJ*, 694, 556
- Brockamp M., Küpper A. H. W., Thies I., Baumgardt H., Kroupa P., 2014, *MNRAS*, 441, 150
- Brodie J. P. et al., 2014, *ApJ*, 796, 52
- Brodie J. P., Strader J., 2006, *ARA&A*, 44, 193
- Cappellari M., 2002, *MNRAS*, 333, 400
- Cappellari M., 2008, *MNRAS*, 390, 71
- Cappellari M. et al., 2013, *MNRAS*, 432, 1709
- Chen C.-W., Côté P., West A. A., Peng E. W., Ferrarese L., 2010, *ApJS*, 191, 1
- Côté P., Marzke R. O., West M. J., 1998, *ApJ*, 501, 554
- Das P., Gerhard O., Churazov E., Zhuravleva I., 2010, *MNRAS*, 409, 1362
- de Zeeuw P. T. et al., 2002, *MNRAS*, 329, 513
- Durrell P. R. et al., 2014, *ApJ*, 794, 103
- Dutton A. A., Macciò A. V., 2014, *MNRAS*, 441, 3359
- Eke V. R. et al., 2004, *MNRAS*, 348, 866
- Emsellem E., Monnet G., Bacon R., 1994, *A&A*, 285, 723
- Emsellem E., Krajnović D., Sarzi M., 2014, *MNRAS*, 445, L79
- Ferrarese L. et al., 2016, *ApJ*, 824, 10
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Gebhardt K., Thomas J., 2009, *ApJ*, 700, 1690
- Gebhardt K., Adams J., Richstone D., Lauer T. R., Faber S. M., Gültekin K., Murphy J., Tremaine S., 2011, *ApJ*, 729, 119
- Geller M. J., Huchra J. P., 1983, *ApJS*, 52, 61
- Huchra J. P., Geller M. J., 1982, *ApJ*, 257, 423
- Ko Y. et al., 2017, *ApJ*, 835, 212
- Li H., Li R., Mao S., Xu D., Long R. J., Emsellem E., 2016, *MNRAS*, 455, 3680
- Li H. et al., 2017, *ApJ*, 838, 77
- Long R. J., Mao S., 2010, *MNRAS*, 405, 301
- Longobardi A., Arnaboldi M., Gerhard O., Pulsoni C., Söldner-Rembold I., 2018a, *A&A*, 620, A111
- Longobardi A. et al., 2018b, *ApJ*, 864, 36
- Lyubenova M. et al., 2016, *MNRAS*, 463, 3220
- McLaughlin D. E., 1999, *ApJ*, 512, L9
- Mei S. et al., 2007, *ApJ*, 655, 144
- Montes M., Trujillo I., Prieto M. A., Acosta-Pulido J. A., 2014, *MNRAS*, 439, 990
- Muñoz R. P. et al., 2014, *ApJS*, 210, 4
- Murphy J. D., Gebhardt K., Adams J. J., 2011, *ApJ*, 729, 129
- Napolitano N. R., Pota V., Romanowsky A. J., Forbes D. A., Brodie J. P., Foster C., 2014, *MNRAS*, 439, 659
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Nulsen P. E. J., Bohringer H., 1995, *MNRAS*, 274, 1093
- Oldham L., Auger M., 2018, *MNRAS*, 474, 4169
- Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, *ApJ*, 725, 2312
- Peng E. W. et al., 2008, *ApJ*, 681, 197
- Pointecouteau E., Arnaud M., Pratt G. W., 2005, *A&A*, 435, 1
- Pota V. et al., 2015, *MNRAS*, 450, 3345
- Prieto J. L., Gnedin O. Y., 2008, *ApJ*, 689, 919
- Ramos-Almendares F., Sales L. V., Abadi M. G., Doppel J. E., Muriel H., Peng E. W., 2019, preprint ([arXiv:1906.11921](https://arxiv.org/abs/1906.11921))
- Romanowsky A. J., Strader J., Brodie J. P., Mihos J. C., Spitler L. R., Forbes D. A., Foster C., Arnold J. A., 2012, *ApJ*, 748, 29

- Röttgers B., Naab T., Oser L., 2014, *MNRAS*, 445, 1065  
 Sarzi M., Spiniello C., La Barbera F., Krajnović D., van den Bosch R., 2018, *MNRAS*, 478, 4084  
 Starkenburg E. et al., 2009, *ApJ*, 698, 567  
 Strader J. et al., 2011, *ApJS*, 197, 33  
 Syer D., Tremaine S., 1996, *MNRAS*, 282, 223  
 Urban O., Werner N., Simionescu A., Allen S. W., Böhringer H., 2011, *MNRAS*, 414, 2101  
 van den Bosch R. C. E., van de Ven G., 2009, *MNRAS*, 398, 1117  
 Vesperini E., Zepf S. E., 2003, *ApJ*, 587, L97  
 Vesperini E., Zepf S. E., Kundu A., Ashman K. M., 2003, *ApJ*, 593, 760  
 Vesperini E., McMillan S. L. W., D'Antona F., D'Ercole A., 2013, *MNRAS*, 429, 1913  
 Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518  
 Wasserman A. et al., 2018, *ApJ*, 863, 130  
 Watkins L. L., Evans N. W., An J. H., 2010, *MNRAS*, 406, 264  
 Watkins L. L., van de Ven G., den Brok M., van den Bosch R. C. E., 2013, *MNRAS*, 436, 2598  
 Wood R. A. et al., 2017, *ApJ*, 847, 79  
 Wu X., Gerhard O., Naab T., Oser L., Martinez-Valpuesta I., Hilz M., Churazov E., Lyskova N., 2014, *MNRAS*, 438, 2701  
 Zhang H.-X. et al., 2015, *ApJ*, 802, 30  
 Zhu L. et al., 2014, *ApJ*, 792, 59  
 Zhu L. et al., 2016, *MNRAS*, 462, 4001

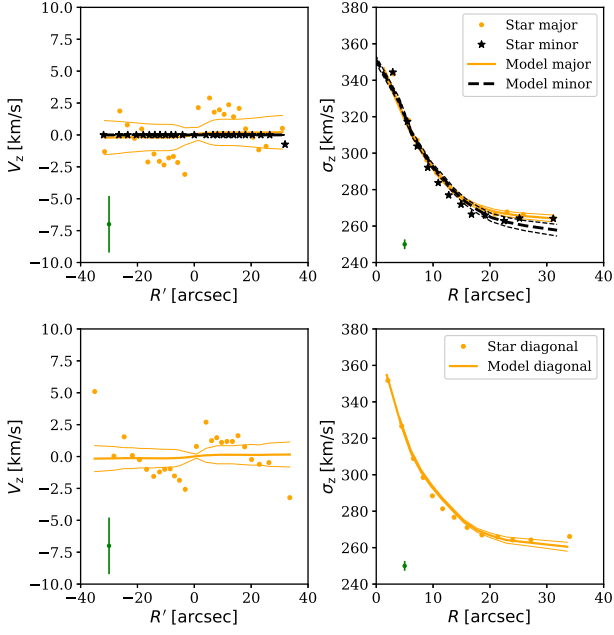
## APPENDIX A: CUSPED DARK MATTER MODELS

In the main text of the paper, we adopt a cored DM profile. In this appendix, we show the results from a cusped model. Most results are unchanged (for more discussion, see Section 5.1).

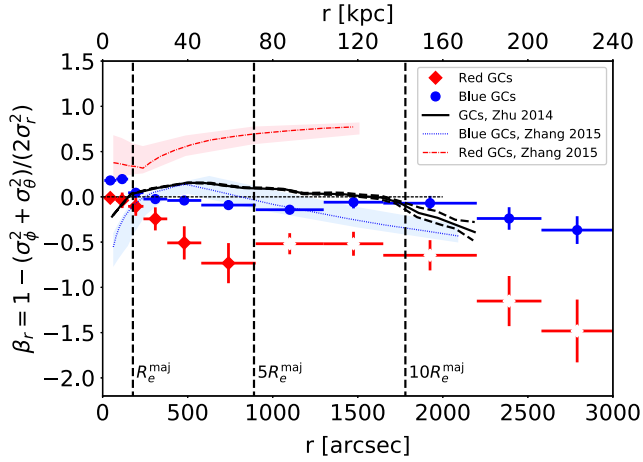
Figs A1 and A2 show the comparison of the observed data with the model predictions. The data and model predictions are binned in the same way as in Fig. 8. The best-fitting cusped model matches the GC data with a mean  $\chi^2$  of 1.15 for the red population and that of 0.51 for the blue population. The mean  $\chi^2$  of the central galaxy directly from the model is 0.33. The velocity dispersion anisotropy profile obtained by the cusped model is shown in Fig. A3.



**Figure A1.** Comparison of the observed GC data with model predictions of the best-fitting cusped model. Fig. 9 shows the corresponding cored model.



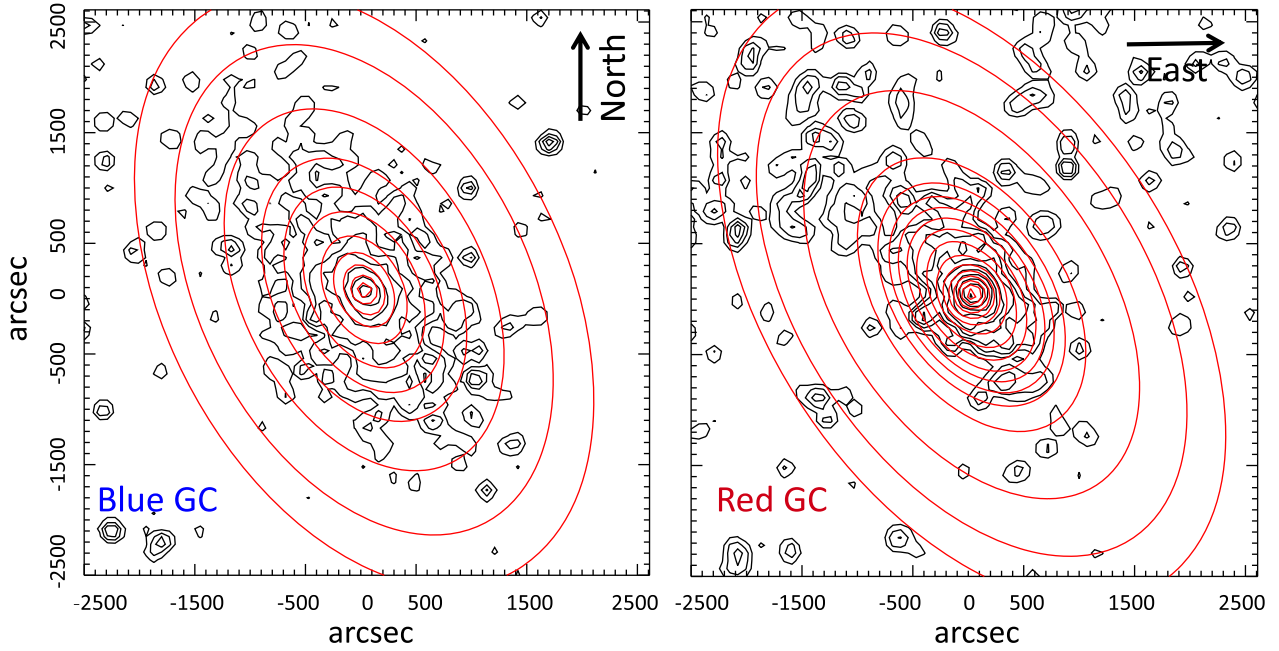
**Figure A2.** Comparison of the stellar data with model predictions of the best-fitting cusped model. Fig. 10 shows the corresponding cored model.



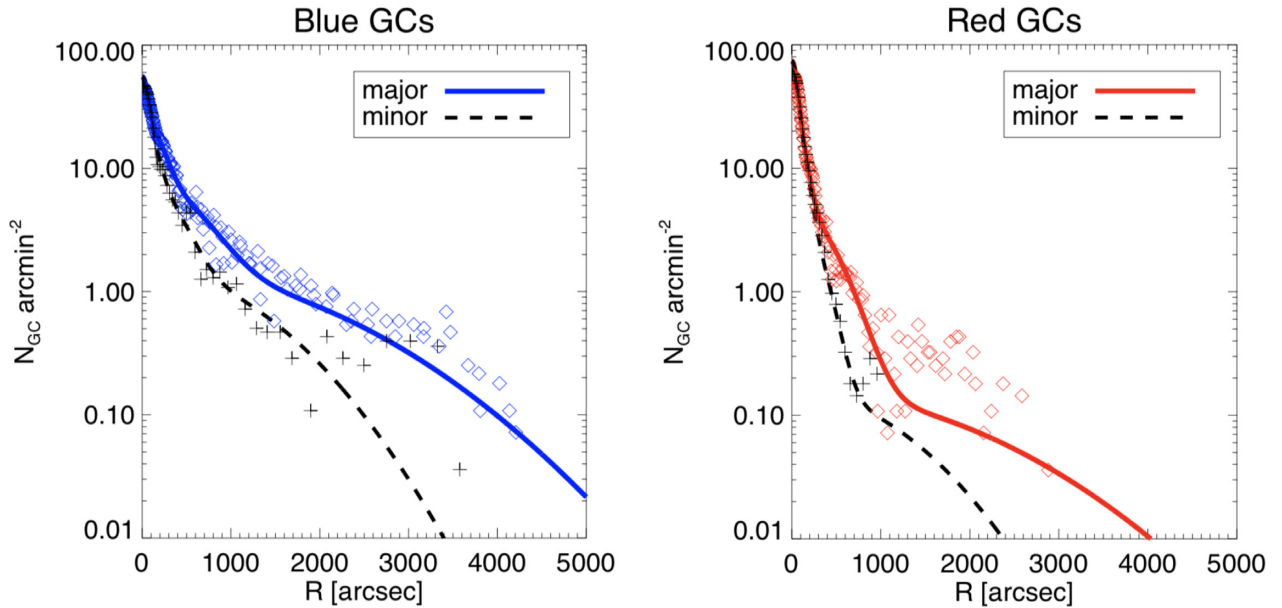
**Figure A3.** The velocity dispersion anisotropy  $\beta_r$  profiles from the cusped model.  $\beta_r$  for red GCs at  $r > 1000$  arcsec should have large uncertainties, and we mark these points with the open symbols. Fig. 13 shows the corresponding cored model.

## APPENDIX B: MGE FITTING TO THE GC POPULATIONS

In Fig. B1, we show the 2D MGEs fitted to the surface number density maps of the blue (left) and red (right) GCs. The black contours are the smoothed surface number density, the red ellipses are the best-fitting MGE models. In Fig. B2, we show how well the MGE models match the surface density profiles along the major and minor axes.



**Figure B1.** The 2D MGEs fitting to the surface number density maps of blue (left) and red (right) GCs. The black contours are the smoothed surface number density, the red ellipses are the best-fitting MGE models. Notice that for the red GCs, there is an asymmetric structure in the north at  $R > 1000$  arcsec, which our model does not match. Instead, our model matched the south side of the data.

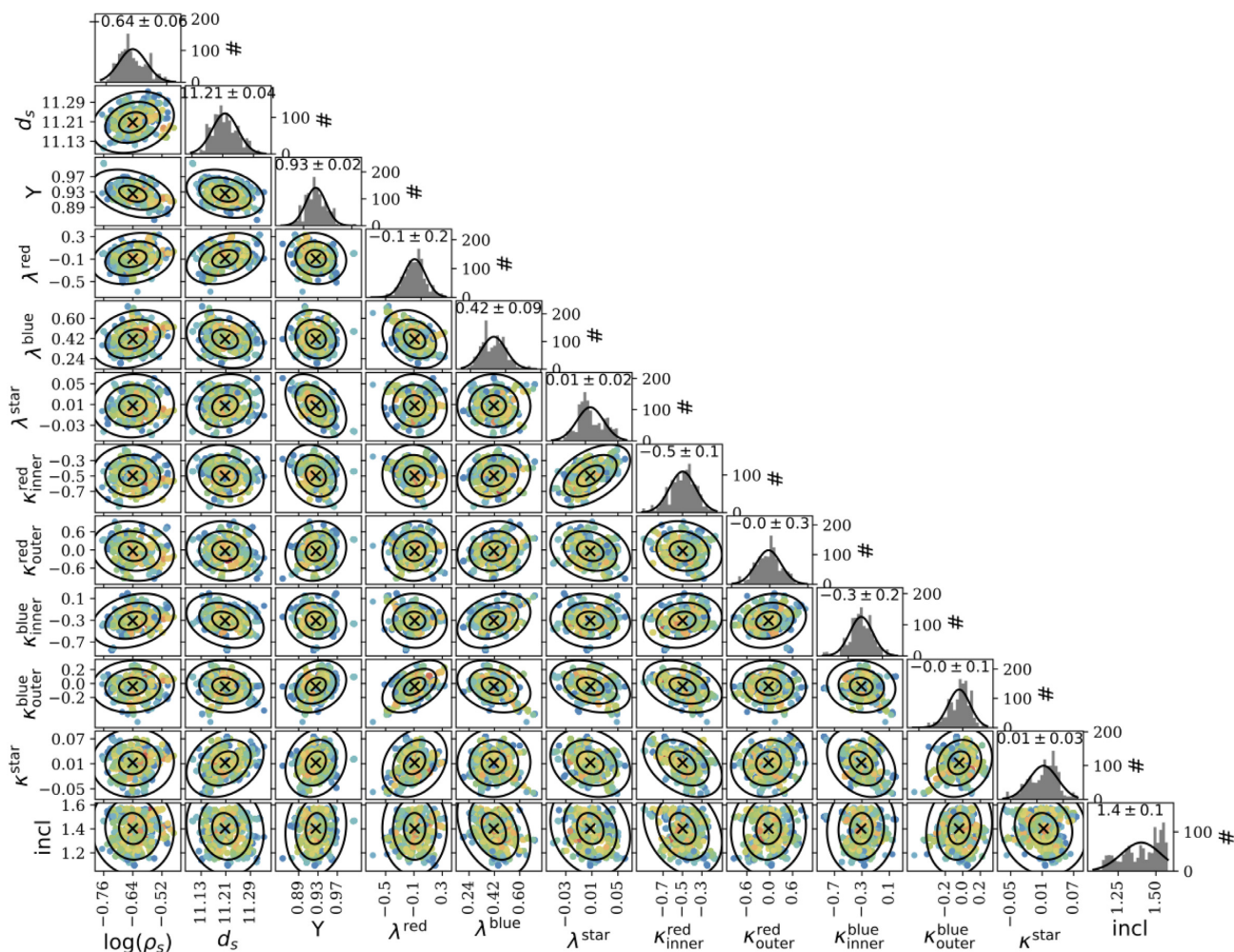


**Figure B2.** The diamonds and the pluses are the surface brightness along the major and minor axes constructed from the photometric data. The solid and dashed curves are the MGE fits along the major and minor axes, respectively.

### APPENDIX C: CORED DARK MATTER MODEL WITH FREE INCLINATION

We created a model with a cored DM halo and leaving the inclination angle  $\vartheta$  free. The best-fitting model parameters are shown in Fig. C1. The model prefers  $\vartheta = 90^\circ$ , but has large uncertainties. There is a weak degeneracy between the inclination angle and the velocity

anisotropy  $\beta_z$  for the blue GCs.  $\beta_z$  tends to be larger when the system is more face-on. This degeneracy is intrinsic in JAM models constrained by 2D kinematic maps, and the degeneracy is stronger for systems with positive  $\beta_z$  (Cappellari 2008). However, overall, the degeneracy is weak in our case, and we obtained similar results for the cored model with either the inclination angle free or fixed with  $\vartheta = 90^\circ$ .



**Figure C1.** The posterior marginalized distributions for the cored model using a free inclination angle. The preferred inclination angle is  $90^\circ$ , but has large uncertainties. The rest of the parameters converge very similar to the model with inclination angle fixed at  $90^\circ$  (Fig. 6).

<sup>1</sup>National Astronomical Observatories, Chinese Academy of Sciences, A 20 Datum Road, Chaoyang District, Beijing 100012, China

<sup>2</sup>School of Astronomy and Space Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China

<sup>4</sup>Department of Astronomy and Tsinghua Center for Astrophysics, Tsinghua University, Beijing 100084, China

<sup>5</sup>Department of Physics and Astronomy, Jodrell Bank Centre for Astrophysics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

<sup>6</sup>Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

<sup>7</sup>Department of Astronomy, Peking University, Beijing 100871, China

<sup>8</sup>Armagh Observatory and Planetarium, College Hill, Armagh BT61 9DG, UK

<sup>9</sup>Centre for Astrophysics Research, University of Hertfordshire, Hatfield, Herts AL1 9AB, UK

<sup>10</sup>Department of Astrophysics, University of Vienna, Türkenschanzstrasse 17, Vienna A-1180, Austria

<sup>11</sup>Department of Astronomy, CAS Key Laboratory for Research in Galaxies and Cosmology, University of Science and Technology of China, Hefei 230026, China

<sup>12</sup>CNRS, CNES, Laboratoire d'Astrophysique de Marseille UMR 7326, Aix Marseille Université, Marseille F-13388, France

<sup>13</sup>National Research Council of Canada, Herzberg Astronomy and Astrophysics Program, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada

<sup>14</sup>Department of Physics and Astronomy, Center for Astronomy and Astrophysics, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>15</sup>Shanghai Key Lab for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>16</sup>Department of Physics and Astronomy, Astronomy Program, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

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