

A comprehensive separation of dark matter and baryonic mass components in galaxy clusters – II: an overview of the mass distribution in Abell S1063

Benjamin Beauchesne¹,^{1,2★} Benjamin Clément,³ Marceau Limousin,⁴ Anna Niemiec,⁵ Mathilde Jauzac¹,^{1,2,6,7} Belén Alcalde Pampliega,^{8,9,10} Johan Richard¹,¹¹ Guillaume Mahler,¹² Jose M. Diego¹,¹³ Pascale Hibon¹,⁸ Anton M. Koekemoer¹,¹⁴ Thomas Connor,¹⁵ Jean-Paul Kneib³ and Andreas L. Faisst¹⁶

¹Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

²Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

³Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

⁴Aix Marseille Univ, CNRS, CNES, LAM, 13388Marseille, France

⁵Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC IN2P3, 53, Avenue des Martyrs, F-38000 Grenoble, France

⁶Astrophysics Research Centre, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa

⁷School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa

⁸ESO Vitacura, Alonso de Córdova 3107, Vitacura, Casilla 19001, Santiago de Chile, Chile

⁹Instituto de Estudios Astrofísicos, Facultad de Ingeniería y 455 Ciencias, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile

¹⁰SKA Observatory, Macclesfield Jodrell Bank, SK11 9FT, UK

¹¹Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon (CRAL) UMR5574, F-69230 Saint-Genis-Laval, France

¹²STAR Institute, Quartier Agora - Allée du six Août, 19c B-4000 Liège, Belgium

¹³Instituto de Física de Cantabria (CSIC-UC). Avda. Los Castros s/n. E-39005 Santander, Spain

¹⁴Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218, USA

¹⁵Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA

¹⁶Caltech/IPAC, 1200 E. California Blvd. Pasadena, CA 91125, USA

Accepted 2026 April 12. Received 2026 February 26; in original form 2025 September 9

ABSTRACT

In the first paper of this series, we derived mass constraints on the total mass and the baryonic components of the galaxy cluster Abell S1063. The main focus was to recover stellar masses and kinematics for cluster members, the brightest cluster galaxy (BCG) and the intra-cluster light (ICL). In this second paper, we introduce a multi-probe mass modelling approach that incorporates constraints on both the total mass and the individual baryonic components. We obtain comprehensive mass models of Abell S1063, in which the dark matter distribution is disentangled from the baryonic mass at both cluster and galaxy scales. The best-fitting mass model achieves an RMS of 0.50 arcsec on the positions of strongly lensed multiple images. The kinematic profiles of the BCG & ICL, as well as the X-ray surface brightness of the intra-cluster gas, are accurately reproduced within observational uncertainties. However, a 35 km s^{−1} scatter is required for the cluster member line-of-sight dispersions. This method yields the most complex parametric mass model with consistency among almost all available mass constraints. We find a 1 σ agreement between the inferred stellar-to-subhalo mass relation and that predicted by large-scale cosmological simulations. The ICL stellar mass derived from our model is consistent with estimates from stellar population modelling. We present the first multi-probe mass modelling method capable of disentangling the dark matter from the baryonic mass distributions in massive galaxy clusters. Its results, such as the stellar-to-subhalo mass relation or the distribution of each mass component, can be directly compared to hydrodynamical cosmological simulations such as illustrisTNG.

Key words: gravitational lensing: strong – galaxies: kinematics and dynamics – galaxies: stellar content – galaxies: clusters: general – galaxies: clusters: individual: Abell S1063 – X-rays: galaxies: clusters.

* E-mail: benjamin.e.beauchesne@durham.ac.uk

1 INTRODUCTION

In this series of two papers, we aim to develop a comprehensive mass modelling method in which each main cluster mass component is modelled, i.e. where dark matter (DM) is fully separated from baryons. Such a method would generate some of the most sophisticated parametric mass models to date, representing the first approach to achieve consistency across nearly all available mass probes. This new method extends the framework developed by B. Beauchesne et al. (hereafter, B24 2024), which combines constraints from strong lensing together with the X-ray emission from the intra-cluster gas in a self-consistent framework. Thanks to this combined dataset, the intra-cluster gas or intra-cluster medium (ICM), i.e. the main contributor to the baryonic mass budget, can be disentangled from the total mass. Hence, this series of papers now focuses on the remaining baryonic component, i.e. the stellar mass, which is contained in cluster galaxies and the intra-cluster stars. The latter are responsible for emitting the intra-cluster light (ICL). As in the cluster centre, the brightest cluster galaxy (BCG) is difficult to distinguish from the ICL, we thus consider them as a single component, denoted BCG & ICL.

To test such new method, we are using Abell S1063 (Hereafter AS1063) as a test case. Thanks to an observational dataset of increasing quality over the years, AS1063 has been used as a test case for methodological improvements. It is one of the morphologically simplest clusters with such data. M. Bonamigo et al. (2018) added a fixed mass component representing the X-ray gas based on M. Bonamigo et al. (2017) method. P. Bergamini et al. (2019) and G. Granata et al. (2022) focus on improving the cluster member modelling by calibrating their scaling laws with galaxy kinematics and morphology. The earlier used a scaling based on the Faber & Jackson law (S. M. Faber & R. E. Jackson 1976) while the latter made use of the more accurate fundamental plane of elliptical galaxies (J. B. Hyde & M. Bernardi 2009). M. Limousin, B. Beauchesne & E. Jullo (2022) revisited P. Bergamini et al. (2019) model and tested the addition of a free-form perturbative B-spline component (B. Beauchesne et al. 2021). B24 incorporated elements from G. Granata et al. (2022) and M. Bonamigo et al. (2018) to develop the framework that we intend to extend in this work. That framework allows us to disentangle the DM from the main baryonic component (i.e. intra-cluster gas). It reduces bias on the resulting DM distribution, in particular when the gas distribution does not follow the DM. However, the B24 framework does not account for the baryonic mass in the form of stellar mass. This baryonic contribution is usually small, except locally when a cluster member or the BCG is present. The DM distribution in and around the BCG is of particular interest as cold dark matter (CDM) model and its alternative such as Fuzzy DM (FDM; W. Hu, R. Barkana & A. Gruzinov 2000) or self-interacting DM (D. N. Spergel & P. J. Steinhardt 2000, SIDM;) disagree on the expected DM profile slope. CDM prefers a cuspy profile, while the other two favour a cored distribution. Strong lensing models of galaxy clusters tend to favour cored distributions for the total mass profile (M. Limousin et al. 2022), disentangling the DM distribution up to the BCG would allow us to assess if it is due to the DM profile or an interplay between baryons and DM.

AS1063 has been modelled several other times in the past, with different observational datasets and methods. As part of the Cluster Lensing and Supernova Survey (CLASH; M. Postman et al. 2012), the Hubble Frontier Fields (HFF; J. M. Lotz et al. 2017), the Beyond Ultra-deep Frontier Fields And Legacy Observations

(BUFFALO; C. L. Steinhardt et al. 2020) and JWST GLIMPSE programme (H. Atek et al. 2025), AS1063 has been modelled for each new observational data set. A. Monna et al. (2014) produced a parametric mass model from the CLASH data with GLEE (S. H. Suyu & A. Halkola 2010; S. H. Suyu et al. 2012). In preparation for the upcoming HFF observations, both T. L. Johnson et al. (2014); J. Richard et al. (2014) modelled the cluster with LENSTOOL (E. Jullo et al. 2007) using different mass parametrizations. G. B. Caminha et al. (2016) made a model with LENSTOOL based on the HFF data and new spectroscopic measurements from VLT/MUSE and VLT/VIMOS as part of the CLASH-VLT programme (A. Mercurio et al. 2021). Those data sets were also used by J. M. Diego et al. (2016) to make a free-form model with WSLAP (J. M. Diego et al. 2005, 2007). More recently, with the GLIMPSE data, parametric and free-form models have been presented by H. Atek et al. (2025) and J. M. Diego et al. (2026), respectively.

In B. Beauchesne et al. (2025) (hereafter B25a), we detail the mass constraints required to set up our comprehensive mass modelling of AS1063. In Section 3, we present the existing set of constraints from strong lensing and X-rays taken from B24. In Sections 4 and 5, we focus on the supplementary mass probes, i.e. cluster members and the BCG & ICL. That set of observations is the same for both components, although we apply different treatments adapted to their properties. We recover the cluster member light distribution by fitting all galaxies in the cluster field to reduce bias from the crowded environment. Thanks to this light model, we are able to subtract the galaxy light model to fit the BCG & ICL with a multi-Gaussian expansion (MGE). For both components, we fit their spectral energy density (SED) and their stellar kinematics to obtain stellar mass estimates as well as a total mass probe. We recall a part of the mass constraints presented in B25a, in Fig. 1.

In this second paper, we present the mass modelling methodology and detail how we use the supplementary mass constraints from B25a. In Section 2, we detail the mass model parametrization of each component, where we use the light distribution from cluster members and the BCG & ICL as a tracer of their stellar mass distribution. In Section 3, we define the likelihoods for each mass probe: X-ray surface brightness, strong lensing systems, cluster member kinematics, and the BCG & ICL component. We conclude this section by describing the tailored optimization process required by the complex mass model needed to accommodate all these constraints. We then present the mass model results in Section 4 and discuss the reliability of the mass estimates from the new component in Section 5.

We recall the convention used in B25a, where we adopt a flat Λ CDM cosmology with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Magnitudes are quoted in the AB system. The uncertainties quoted in that article are the median-centred credible intervals (CI) based on the posterior distribution of the considered random variable. These intervals are presented as $n\sigma$ CI where n is an integer such that CI contains $100 \times \text{erf}\left(\frac{n}{\sqrt{2}}\right)$ per cent of the posterior. If we use σ to denote a standard deviation, we explicitly notify it in the text.

2 MASS MODELLING HYPOTHESES

In our model, we aim to disentangle the DM from the baryonic components at galaxy and cluster scales. In the following, we define how we model the main cluster components (DM, BCG &

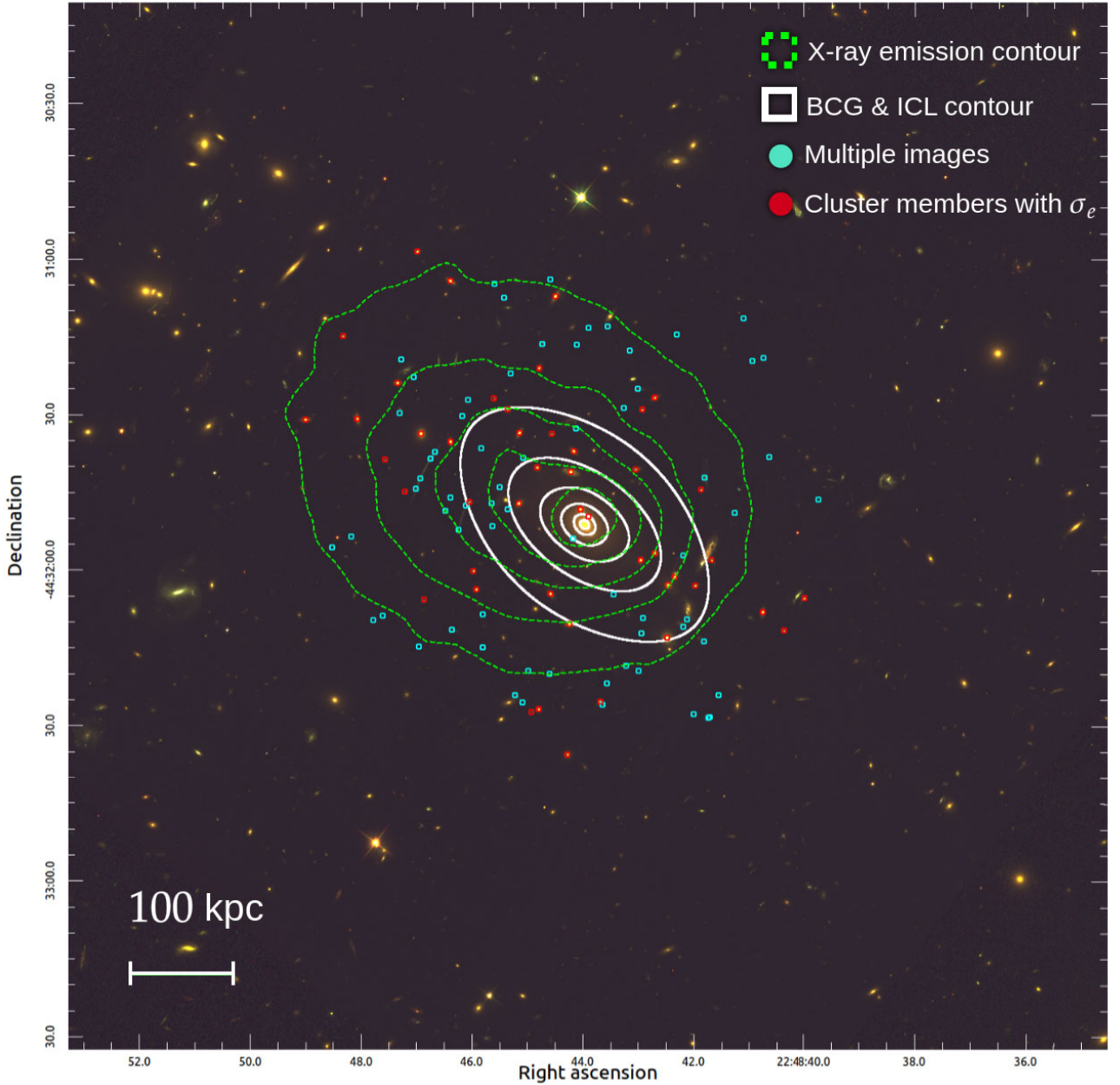


Figure 1. The BUFFALO colour composite image of AS1063 with the following *HST* filters: F435W (blue), F606W (green) and F814W (red). The X-ray surface brightness from the CHANDRA X-RAY OBSERVATORY is shown by the green dashed contours. The set of multiple images used in this work is highlighted by the cyan circles. The white contours present the BCG & ICL light distribution as fitted in B25a. Cluster members for which we measure the velocity dispersion, σ_e , in B25a are highlighted by red circles.

ICL, intra-cluster medium, and cluster members) so that their respective baryonic fractions can be assigned their own mass components.

2.1 Intra cluster medium

To model the ICM, we follow the same method as in B24, as it already fits our requirement of separating the baryonic mass from the DM. The gas distribution is modelled by a sum of dPIE potentials in a free-form way. All potentials are free to move within the cluster field, and we only fix their cut radii at 1.25 Mpc, as in B24.

Indeed, it is ill-constrained when more than one potential is used. This cut radius limit corresponds to the radius at which the observed number counts are approximately equal to the background value.

The number of potentials is determined following the goodness-of-fit procedure of B24, whereby potentials are added iteratively until the best-fitting model adequately reproduces the observed X-ray emission according to the B24 criterion. This criterion is an adaptation of the usual goodness-of-fit assessment with χ^2 1 for the X-ray likelihood defined in Section 3.1. Hence, as in B24, we use three dPIEs to model the hot gas in AS1063.

2.2 Cluster members

The cluster member distribution is one of the less significant mass components of the cluster, although due to the mass concentration of the galaxy, it locally influences the position of multiple images. Hence, they are an important part of a cluster lens mass model, as some systems cannot be reproduced without them. However, due to their large number and the scarcity of lensing constraints, several simplifying assumptions about their parameters are made. These simplifications are usually based on a ‘light traces matter’ principle.

In this work, we deviate from the standard modelling of cluster members used in the LENSTOOL software (J. P. Kneib et al. 1993; E. Jullo et al. 2007) which is based on the Faber & Jackson law (S. M. Faber & R. E. Jackson 1976). It scales the dPIE parameters (r_{core} , r_{cut} and σ_0) of each cluster member based on its luminosity L and a reference galaxy (r_{core}^* , r_{cut}^* , σ_0^* and L^*). During the parameter inference, only the dPIE parameters of the reference galaxy are optimized. The scaling relations are the following:

$$r_{\text{cut}} = r_{\text{cut}}^* \left(\frac{L}{L^*} \right)^{1/2}; r_{\text{core}} = r_{\text{core}}^* \left(\frac{L}{L^*} \right)^{1/2}; \sigma_0 = \sigma_0^* \left(\frac{L}{L^*} \right)^{1/4} \quad (1)$$

the other parameters of the dPIE are usually fixed from the light profile: centre, position angle and ellipticity. Some works prefer circular cluster members rather than the ellipticity of the light distribution. Hence, the DM and baryons of each galaxy are modelled by a single profile.

In this work, we aimed at separating the DM from the baryons in each galaxy using individual mass profiles to be consistent with the cluster-scale model. Since the baryonic distribution of elliptical galaxies is dominated by stars, it can serve as a stellar tracer for kinematic modelling, providing more local mass constraints than lensing. Hence, separating baryons and DM in cluster members not only gives more information to extract from the model, but it also allows the model to use more observational datasets. To make such a framework work, we need relations that link the baryons and DM profiles. So instead of using a relation based on observations such as the Faber & Jackson law or fundamental plane of an elliptical galaxy, we use the Stellar-to-subhalo mass relation (SsHMR). It has the advantages to be easily comparable to simulation results.

In this new framework, we still use dPIE potentials for their computational efficiency rather than combining a Sérsic and a power-law profile as it is commonly done in galaxy-scale lensing analyses (e.g. A. J. Shajib et al. 2021). However, a dPIE profile can approximate a Sérsic light profile relatively well for a Sérsic index between 1 and 4 (A. A. Dutton et al. 2011). Each galaxy component (baryons and DM) is modelled with a separate dPIE profile, subject to several simplifying assumptions to limit the number of free parameters. Hence, we use the same centre coordinates, ellipticity and position angle for both dPIEs. These parameters are assumed to follow the light distribution. Such approximations are poor for disk galaxies, although the majority of the included cluster members are elliptical galaxies or galaxies in transition to elliptical. We fix these parameters to the best-fitting values of the cluster member light profile with a dPIE presented in B25a (Section 4.1). The core and cut radii of the baryonic dPIE ($r_{\text{core,b}}$ and $r_{\text{cut,b}}$), are fixed to the values of the light profile as well. Only the central velocity dispersion $\sigma_{0,b}$ remains free, which we rescale to match the stellar mass from B25a (Section 4.3). Hence, the baryonic component is fixed in the overall model and is de-

termined only from measurements independent of the lensing analysis.

For the DM component, radii, $r_{\text{core,DM}}$ and $r_{\text{cut,DM}}$, and velocity dispersion, $\sigma_{0,DM}$, are scaled by a spatial and concentration scale, α_c and β_c , with respect to their baryonic counterpart, ($r_{\text{core,b}}$, $r_{\text{cut,b}}$ and $\sigma_{0,l}$), such that:

$$r_{\text{core,DM}} = \alpha_c r_{\text{core,b}} \quad (2)$$

$$r_{\text{cut,DM}} = \alpha_c r_{\text{cut,b}} \quad (3)$$

$$\sigma_{0,DM} = \beta_c \sigma_{0,l} \quad (4)$$

The spatial scale, α_c , is assumed to be the same for all cluster members, and is optimized with our set of heterogeneous constraints. This parameter is expected to exhibit a radial, cluster-centric variation, as it reflects the tidal stripping of DM halos experienced by galaxies falling into the cluster potential well (A. Niemiec et al. 2022). We leave the exploration of such variations for future work.

β is different for each cluster member, and is recovered through the modelling of the SsHMR in the model under the form of a double power-law as in A. Niemiec et al. (2022) and B. P. Moster, T. Naab & S. D. M. White (2013), such that:

$$M_{\text{tot}} = 2N \left(\left(\frac{M_*}{M_1} \right)^\delta + \left(\frac{M_*}{M_1} \right)^{-\gamma} \right) M_* \quad (5)$$

where N , M_1 , δ , and γ are optimized, and define the cluster member distribution model with α_c . We could switch roles between α_c and β_c . In that case, we would optimize β_c and set α_c with the SsHMR. However, β_c is more likely to be ill constrained by strong lensing as it mostly affects cluster members inner profiles. It is still not perfect as γ and α_c are not well constrained. We impose an upper bound on them, as a large value of γ and α_c leads to unrealistic mass for the faintest cluster member, which is not constrained. We lack velocity dispersions for these cluster members, and these unrealistic masses are highly concentrated, which severely reduces the ability to penalize them through strong lensing. More generally, α_c does not seem to be well bounded by the data, and as it increases, it changes the mass balance between the other components, which makes it hard to match the measured velocity dispersion of the BCG within the two-step optimization process presented in Section 3.5, for example. We based the upper bound on empirical limits obtained from preliminary models.

In our modelling, the SsHMR is the usual scaling relation used for cluster members modelling with LENSTOOL. However, instead of using an observational law, we use a parametrization oriented towards comparison with cosmological simulations. The SsHMR of our model can be directly compared with simulation results, such as those presented in A. Niemiec et al. (2022). We recover β_c from the SsHMR by writing the total mass of the combined dPIE models, M_{tot} , such that:

$$M_{\text{tot}} = \left(\alpha_c \beta_c^2 \frac{LM_\odot}{L_\odot} + M_* \right) \quad (6)$$

where L is the total luminosity of the dPIE fitted to the light distribution, and M_* is the stellar mass associated with the object.

In contrast to B24, we do not inject the cluster member line-of-sight velocity dispersion (LOSVD) in their associated dPIE parameters. We would rather use them as a new set of constraints, as presented in Section 3.3. In doing so, the measurement uncertainties on the LOSVD are more directly accounted for, and their weight naturally balanced against that of other mass probes, such as strong lensing, within the inference process. In B24, these

quantities are accounted for with the prior on the galaxy scaling law.

As in B25a (Section 4.3), we obtain stellar mass estimates from different SED models used to define three different parametrizations of cluster members. Hence, we gain insights into the influence of SED systematic biases on this type of modelling, particularly on the SsHMR discussed in Section 5.5.

2.3 BCG & ICL

To model the BCG & ICL component, we assume a stellar component to represent its baryonic distribution and a DM component for the BCG. The motivation behind the BCG DM component is to allow for different mass slopes at scales that differ from those of the main DM halo.

We model the BCG & ICL stellar component using an MGE parametrization. We begin with the MGE model of its light distribution in the F160W filter, measured in B25a (Section 5.1). Contours of the MGE model are presented in Fig. 1. We adopt identical parameters for light and mass distributions, except that the Gaussian normalizations now represent the stellar mass-to-light ratio $\Upsilon_{*}^{\text{BCG}}$ of the BCG & ICL component. We designate the estimate of $\Upsilon_{*}^{\text{BCG}}$ by LENSTOOL as $\Upsilon_{*,\text{lt}}^{\text{BCG}}$, in contrast to $\Upsilon_{*,\text{F160W}}^{\text{BCG}}$ that we refer to as $\Upsilon_{*,\text{SED}}^{\text{BCG}}$. We could impose a prior on $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ based on $\Upsilon_{*,\text{SED}}^{\text{BCG}}$ as estimated in B25a (Section 5.3). However, we prefer a uniform prior since SED fits can misestimate this ratio when the assumed initial mass function (IMF) is inappropriate for the object. In particular, stellar mass estimates from kinematic modelling of early-type galaxies present significant differences compared to SED estimates (for a review on this topic see R. J. Smith 2020). To further accommodate such discrepancies, we define three MGE parametrizations which allow for different degrees of variation of $\Upsilon_{*}^{\text{BCG}}$ with the cluster-centric radius. The first parametrization assumes constant $\Upsilon_{*}^{\text{BCG}}$, renormalizing the entire light distribution with a single coefficient $\Upsilon_{*,\text{lt}}^{\text{BCG},1}$. We refer to this parametrization as ‘BCG–ML 1’. The second uses two coefficients: $\Upsilon_{*,\text{lt1}}^{\text{BCG},2}$ for Gaussians with standard deviations smaller than the BCG half-light radius ($R_e = 17$ kpc; L. Tortorelli et al. (2018)), and $\Upsilon_{*,\text{lt2}}^{\text{BCG},2}$ for wider Gaussians. This effectively separates the BCG from the ICL contributions. We refer to this parametrization as ‘BCG–ML 2’. To accommodate wider variations of $\Upsilon_{*}^{\text{BCG}}$, we make three groups of Gaussians through a K-means clustering method as implemented in SCIKIT-LEARN (F. Pedregosa et al. 2011), each represented by one coefficient, $\Upsilon_{*,\text{lt}i}^{\text{BCG},3}$ where $i \in \{1, 2, 3\}$. $\Upsilon_{*,\text{lt1}}^{\text{BCG},3}$ represents the same group of Gaussians as $\Upsilon_{*,\text{lt1}}^{\text{BCG},2}$, while $\Upsilon_{*,\text{lt2}}^{\text{BCG},3}$ and $\Upsilon_{*,\text{lt3}}^{\text{BCG},3}$ allow for more variation of $\Upsilon_{*}^{\text{BCG}}$ within the ICL. We refer to this last parametrization as ‘BCG–ML 3’.

Since LENSTOOL lacks native MGE potential implementation, we incorporate them as fixed external maps of lensing quantities (computed in B25a, Section 5.1) scaled by normalization factors. When we split the MGE parametrization into groups, we apply the same procedure to each group to compute the lensing quantities as if they were separate MGE potentials. We include the map of the lensing quantities by interpolating them with degree 3 B-spline surfaces, allowing us to use the spline implementation of LENSTOOL (B. Beauchesne et al. 2021). In contrast to B. Beauchesne et al. (2021), we do not impose boundary conditions, and there is a different B-spline surface for each quantity: lensing potential, deflection angle, convergence and shear. We simply use the implementation to compute the value of the

B-spline surface. This approach generalizes to any light distribution parametrization, accommodating complex cases such as SMACS J0723.3–7323 (G. Mahler et al. 2023). We chose an MGE parametrization specifically to enable stellar kinematic analysis through Jeans Anisotropic Modelling (JAM), which requires MGE-parametrized mass and light profiles.

To model the DM component of the BCG, we use a dPIE profile. Regarding its parameters, we fix its centre to the BCG centroid and leave the other parameters free to vary. We allow the position angle and ellipticity to vary, as their values may differ from those of the light distribution. We also allow for possible offsets between the DM and light distributions, as parametrized by the smooth cluster-scale DM halo that we define in the following section. We note that this DM component was required in addition to the cluster-scale DM to obtain a sufficiently high velocity dispersion near the BCG and match the observation.

2.4 Dark matter

For the DM cluster-scale smooth component, we use a modelling approach similar to other LENSTOOL models to represent the bulk of the cluster mass. This component is often referred to as the cluster DM component by neglecting baryons contained in the ICM and the ICL. In our case, since we are modelling the baryons with other profiles, it only accounts for the DM mass.

We follow the approach of B24 by using two dPIEs haloes, one representing the main cluster DM halo, and a second one representing the North–East (NE) halo. Regarding the model parameters, we use the same assumptions as B24 for the NE halo position and radii, but we leave its position angle and ellipticity free to vary. We allow all of these parameters to vary for the main clump, and we constrain its central position to a square of 10 arcsec centred around the BCG position. Similarly to B24, we fix the cut radius to 3 Mpc.

These two additional dPIEs complete the cluster mass model. Overall, the model includes 588 different profiles. Cluster members dominate this count, with each galaxy contributing two profiles (one for baryons, one for DM). Despite the large profile count, most free parameters characterize cluster-scale components, such as the main DM halo and ICM, which make fewer assumptions than the cluster member population.

3 LIKELIHOODS

Thanks to our comprehensive model presented in Section 2, we can define likelihoods for our set of heterogeneous constraints: lensing, cluster members, σ_e , X-ray surface brightness, and BCG & ICL, V_{rms} . We recall the set of constraints in Fig. 1, with X-ray surface brightness contours, multiple image systems and cluster members with measured kinematics. Each of these likelihoods constrains different parts of the model, as shown in Fig. 2, where each model component is associated with a different set of likelihoods. We have strong lensing ($\mathcal{L}_{\text{SL}}$), X-rays ($\mathcal{L}_{\text{X-ray}}$), cluster member kinematics ($\mathcal{L}_{\text{CL-kin}}$) and the BCG & ICL kinematics ($\mathcal{L}_{\text{BCG-kin}}$) likelihoods. Only the first one, $\mathcal{L}_{\text{SL}}$, constrains the entire mass distribution, as it is sensitive to the cluster’s total mass. At the exception of $\mathcal{L}_{\text{BCG-kin}}$, which also depends on the total mass in the cluster core, the last two, $\mathcal{L}_{\text{X-ray}}$ and $\mathcal{L}_{\text{CL-kin}}$, only constrain one model component each. We define each of these likelihoods in the following sections. In Section 3.5, we detail our parameter inference procedure.

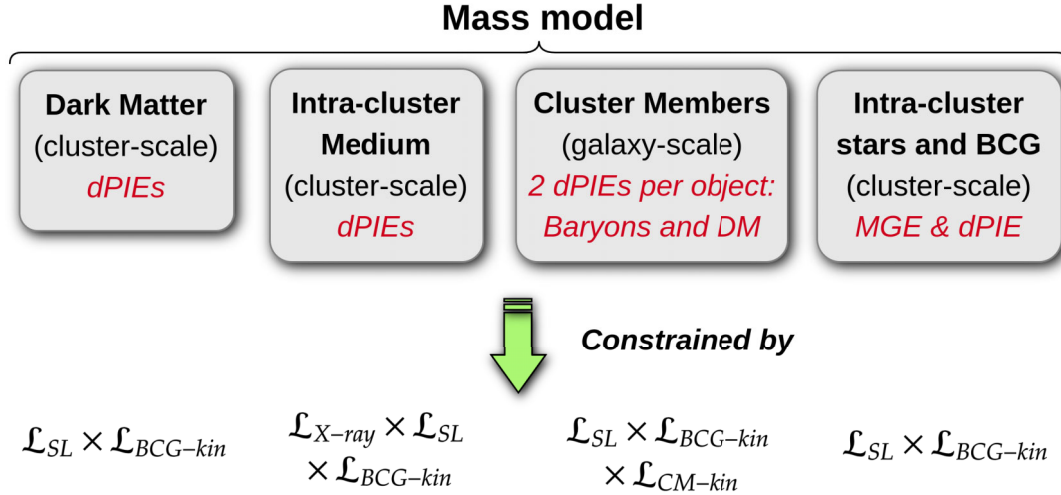


Figure 2. Overview of each model component. Each one of them is associated with a set of analytical profiles and the likelihood that they are affecting.

3.1 X-ray surface brightness

To allow our model to account for the ICM X-ray emission as a constraint, we use the mass distribution of the dPIEs associated with the gas to predict the cluster X-ray surface brightness. We use the same transformation as in B24 to convert the gas mass distribution to its associated electron density, and then its surface brightness. We refer the reader to B24 and the references within for a detailed procedure description.

To define the likelihood and to take into account the limitations of our modelling choices, we incorporate an intrinsic error on the X-ray photon count, σ_X , and we use a likelihood based on the Negative Binomial distribution ($\mathbb{NB}(r = \mu_i^2/\sigma_X^2, p = \mu_i/(\mu_i + \sigma_X^2))$ with μ_i the expected photon count in a bin). The full likelihood is then defined as follows:

$$\mathcal{L}_{X-ray} = \prod_i \frac{\Gamma(k_i + \frac{\mu_i^2}{\sigma_X^2})}{k_i! \Gamma(\frac{\mu_i^2}{\sigma_X^2})} \left(\frac{\mu_i}{\mu_i + \sigma_X^2} \right)^{\frac{\mu_i^2}{\sigma_X^2}} \left(\frac{\sigma_X^2}{\mu_i + \sigma_X^2} \right)^{k_i} \quad (7)$$

where μ_i is the expected photon counts from the model in each pixel of the map, k_i is the observed number of photon counts in a bin, and Γ represents the well-known gamma function.

In contrast to B24, we do not use a constant value of σ_X^2 for the whole field of AS1063. We use an intrinsic error proportional to the expected value of the surface brightness model, such that $\sigma_X = a\mu_i$, where a is the error parameter to optimize. It is better suited to the motivation behind adding this error. The inclusion of σ_X^2 was motivated in B24 by the inability of the model to fully reproduce gas mass variations on scales below that of the cluster. The gas sloshing pattern identified in B24 is an example of such variations. Hence, we expect such phenomena to have an amplitude correlated to the gas density as reflected by the expression of σ_X . If a constant error is still needed, it would likely indicate a misestimation of the background sky.

3.2 Multiple images positions

For the strong lensing constraints, we also follow the same definition as in B24, particularly in choosing the source position of each multiply-imaged system. Indeed, we do not rely on a weighted

barycentre of the positions of each image in the source plane but rather use it as a starting point that we optimize to find the best-available source position for a given system. We assume that the uncertainties on the positions of multiply-imaged systems are well represented by independent Gaussian distributions, which gives the following likelihood definition:

$$\mathcal{L}_{SL} = \prod_j \frac{1}{\prod_i^{N_{im,j}} \sigma_{ij} \sqrt{2\pi}} \exp(-\chi_j^2/2) \quad (8)$$

with N_{sys} , the total number of multiply-imaged systems, $N_{im,j}$, the number of images in the j^{th} system, and σ_{ij} , the observational error on the i^{th} images of the j^{th} system. χ_j^2 is the χ^2 statistics associated with the j^{th} system and reads as follows:

$$\chi_j^2 = \sum_i \frac{\|\vec{\theta}_{i,j}^{obs} - \vec{\theta}_{i,j}^{pred}\|^2}{\sigma_{ij}^2} \quad (9)$$

A positional uncertainty of 0.55 arcsec is adopted to account for both the observational uncertainties in image position identification and the systematic limitations of the model. This value is determined by a maximum-likelihood estimate, calibrated such that the best-fitting model achieves a reduced χ^2 of approximately 1.

3.3 Cluster member kinematics

Thanks to our detailed cluster members modelling, we can access both their light and mass distributions in terms of dPIEs. We use a similar approach to B24 to model the LOSVD of the cluster members at the half-light radius, σ_e , from the light and mass distributions. The main difference is that we do not try to obtain σ_0 from σ_e , but σ_e as a function of dPIEs parameters.

Following P. Bergamini et al. (2019); B. Beauchesne et al. (2024), the LOSVD, σ_{LOS} , is expressed as follows:

$$\sigma_{LOS}^2(R) = \frac{2G}{I(R)} \int_R^\infty \nu(r) M_{3D}(r) \mathcal{F}(r) r^2 \beta_{aniso}^{-2} dr \quad (10)$$

where I and ν are the surface density and density of the luminosity distribution, respectively. M_{3D} is the total mass enclosed in a sphere of radius r . It leaves us with two terms, $\mathcal{F}$ and β_{aniso} , which

are linked to the anisotropy of stars' orbits. For simplicity, we assume that these orbits are isotropic, which gives the following formulas (M. Cappellari 2008):

$$\mathcal{F}(r) = \sqrt{r^2 - R^2} \text{ and } \beta_{\text{aniso}} = 0 \quad (11)$$

To match the actual measurements of σ_{ap} and their extraction apertures, σ_{LOS} is averaged in circular apertures of radius R' . We obtain the following expression:

$$\sigma_{\text{ap}}^2(R') = \frac{2\pi}{L(R')} \int_0^{R'} \sigma_{\text{LOS}}^2(R) I(R) R dR \quad (12)$$

where $L(R')$ is the luminosity enclosed in a circle of radius R' . In our case, we have three dPIEs, one for the light ($\sigma_{0,l}, r_{\text{core},l}, r_{\text{cut},l}$), one for the baryons ($\sigma_{0,b}, r_{\text{core},b}, r_{\text{cut},b}$) and one for DM ($\sigma_{0,\text{DM}}, r_{\text{core},\text{DM}}, r_{\text{cut},\text{DM}}$) masses. We assume that the mass and light of the considered cluster members are the only components that significantly affect the LOSVD. Thus, we obtain the following expressions for the different terms presented in the previous equations:

$$\nu(r) = \frac{\sigma_{0,l}^2(r_{\text{core},l} + r_{\text{cut},l})}{2\Upsilon_{\odot} \pi G r_{\text{core},l}^2 r_{\text{cut},l}} \frac{1}{\left(1 + \left(\frac{r}{r_{\text{core},l}}\right)^2\right) \left(1 + \left(\frac{r}{r_{\text{cut},l}}\right)^2\right)} \quad (13)$$

$$I(r) = \frac{2\sigma_{0,l}^2 r_{\text{cut},l}}{G(r_{\text{cut},l} - r_{\text{core},l})} \Upsilon_{\odot} \left(\frac{1}{\sqrt{r^2 + r_{\text{core},l}^2}} - \frac{1}{\sqrt{r^2 + r_{\text{cut},l}^2}} \right) \quad (14)$$

$$M_{3D}(r) = \sum_{i \in [b, \text{DM}]} \left(\frac{2\sigma_{0,i}^2 r_{\text{cut},i}}{G(r_{\text{cut},i} - r_{\text{core},i})} \times \right. \quad (15)$$

$$\left. \left(r_{\text{cut},i} \arctan\left(\frac{r}{r_{\text{cut},i}}\right) - r_{\text{core},i} \arctan\left(\frac{r}{r_{\text{core},i}}\right) \right) \right) \quad (16)$$

$$L(r) = \frac{2\sigma_{0,l}^2 r_{\text{cut},l}}{G(r_{\text{cut},l} - r_{\text{core},l})} \Upsilon_{\odot} \times \quad (17)$$

$$\left(\sqrt{r^2 + r_{\text{core},l}^2} + r_{\text{cut},l} - \sqrt{r^2 + r_{\text{cut},l}^2} - r_{\text{core},l} \right) \quad (18)$$

In our case, we extract the LOSVD in a radius equal to the half-light radius of the considered cluster member, which in the case of the dPIE is (Á. Elíasdóttir et al. 2007):

$$R_e = \frac{3}{4} \sqrt{r_{\text{core},l}^2 + \frac{10}{3} r_{\text{core},l} r_{\text{cut},l} + r_{\text{cut},l}^2} \quad (19)$$

We have now settled everything to compute $\sigma_e = \sigma_{\text{ap}}(R_e)$, and define the likelihood associated with cluster member kinematics, $\mathcal{L}_{\text{CM-kin}}$. We assume that each σ_e measurement is well represented by a Gaussian distribution, which gives the following definition:

$$\mathcal{L}_{\text{CM-kin}} = \prod_j^{N_{\text{gal}}} \frac{1}{\delta_{\sigma_e,i} \sqrt{2\pi}} \exp\left(-\frac{(\sigma_{e,i} - \sigma_{e,\text{model},i})^2}{2\delta_{\sigma_e,i}^2}\right) \quad (20)$$

where N_{gal} is the number of cluster members with a measured LOSVD, and $\delta_{\sigma_e,i}$ is the observational error on that measurement. $\sigma_{e,i}$ and $\sigma_{e,\text{model},i}$ are the values of the measured and model-predicted LOSVD, respectively.

As our model parameterization includes several hypotheses to simplify the model, we cannot reproduce each LOSVD within its error bars. Similarly to the lensing constraints, we add an intrinsic error, such that the associated reduced χ^2 is approximately 1 for the mode of the posterior. We estimate the systematic error to be 35 km s^{-1} based on a preliminary maximum-likelihood estimate.

3.4 BCG & ICL kinematics

To model the kinematics of the BCG & ICL, we rely on JAM to obtain a 2D model and the JAMPY package (M. Cappellari 2008). We already have an axisymmetric MGE representation of the light distribution as measured in B25a (Section 4.1). Thus, we need to define a similar representation for the mass so we can compute the JAM kinematic model. To be computationally efficient, we base the mass model on the 1D mass profile of the galaxy cluster. We compute the average mass density in spherical shells, which we fit with a 1D MGE representation. We include all profiles that define cluster-scale distributions, such as the smooth DM, the gas and the BCG & ICL components, in the mass density. In preliminary tests, we find that adding some cluster members is necessary to account for all the mass required by the stellar kinematics. Hence, we include each cluster member within the kinematics extraction area. This selection is correct under the assumption of a single mass plane, although in reality, we may include cluster members that would not be in the spherical shells used to estimate the 1D mass profile. We build the 2D MGE mass by assuming the same position angle as the light distribution. This is required by the JAM method. To estimate an ellipticity as close as possible to the underlying mass distribution, we uniformly sample points in the combined area of all extraction regions of the kinematics. We compute the second moment of the mass distribution in the x and y axes of the LENSTOOL reference frame and estimate the ellipticity with the same formulae as SEXTRACTOR¹. In preliminary tests, we tried to use the 2D mass density and obtain the 2D MGE representation directly. However, the requirement of having the same position angle led to a poor fit of the mass by the 2D MGE.

Given a representation of light and mass as MGEs, we need to specify the remaining parameters of the JAM method to properly define the kinematic model. These parameters are the inclination with respect to the LOS, and the anisotropy of the stars' orbits. We marginalize over the inclination by uniformly sampling it within the range that yields a viable kinematic model. We define the minimal inclination, θ_{min} , with the following equation from the JAM code parameter restriction:

$$q_{\text{obs,min}} = \sqrt{\cos^2(\theta_{\text{min}}) + q_{\text{min}} \sin^2(\theta_{\text{min}})} \quad (21)$$

where $q_{\text{obs,min}}$ is the minimal axis ratio among the Gaussians of both MGE representations (i.e. light and mass), and q_{min} is the minimal axis ratio of the deprojected Gaussians in 3D. As in the JAM code, we fix q_{min} to 0.05. $\theta_{\text{max}} = 90$ degree represents the edge-on case. In the JAM code, there are multiple choices for the anisotropy of the stars' orbits: cylindrical alignment of the velocity ellipsoid (M. Cappellari 2008), spherical alignment (M. Cappellari 2020) and custom functions such as a logistic (D. A. Simon, M. Cappellari & J. Hartke 2024) or a constant function. In the preliminary test, we found no significant difference between the two alignment types when integrating the kinematic model across the observational bins. We try to use a logistic function, as in D. A. Simon et al. (2024), to constrain the anisotropy using the data. We used a gradient descent method to optimize the anisotropy in a constant-mass model, but the parameters remained poorly constrained and stayed near the starting points. Hence, we use the cylindrical alignment, as it is less computationally costly than the spherical one. Once the JAM 2D model is defined in the area

¹<https://sextractor.readthedocs.io/en/latest/Position.html>

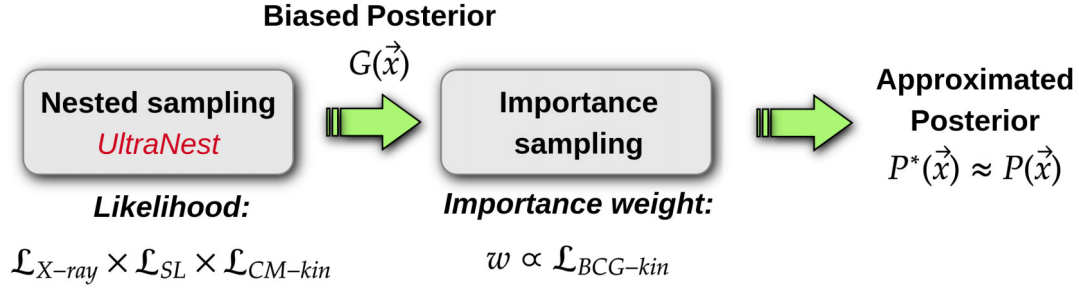


Figure 3. Diagram of the inference workflow in two steps. A biased posterior distribution is obtained on the combined likelihood excluding the BCG & ICL kinematics. In the second step, we resample the biased posterior through importance sampling to obtain the final posterior estimate.

where V_{rms} is measured, we integrate it over the elliptical annuli used for extraction using a trapezoidal method.

In the observation bins, we then define the BCG & ICL kinematics likelihood $\mathcal{L}_{\text{BCG-kin}}$ as a Gaussian likelihood on the V_{rms} measurements such that:

$$\mathcal{L}_{\text{BCG-kin}} = \prod_j^{N_{\text{bin}}} \frac{1}{\delta_{V_{\text{rms},i}} \sqrt{2\pi}} \exp\left(-\frac{(V_{\text{rms},i} - V_{\text{rms},i}^{\text{model}})^2}{2\delta_{V_{\text{rms},i}}^2}\right) \quad (22)$$

where N_{bin} is the number of extraction bins where $V_{\text{rms},i}$ has been measured with the observational error $\delta_{V_{\text{rms},i}}$. $V_{\text{rms},i}^{\text{model}}$ is the model predicted LOS velocity moment.

3.5 Optimization process

In the case of our modelling, the computational cost of the combined likelihood is too high to use regular sampling methods for the posterior, such as Markov Chain Monte Carlo or Nested sampling. However, this cost begins to be too expensive only when we consider the kinematic model of the BCG & ICL. Indeed, this likelihood constrains only a subset of the model parameters, yet its inclusion increases the computational cost of each likelihood evaluation, leading to a significant slowdown.

The BCG & ICL kinematics are particularly interesting as they complete the strong lensing constraints in the cluster core. Hence, we do not want to exclude them. We converge on the optimization process presented in Fig. 3, which uses an approximation to obtain the posterior of the total likelihood. We start the inference scheme by running a regular sampling method to obtain the posterior of the incomplete likelihood $\mathcal{L}_{\text{CM-kin}} \times \mathcal{L}_{SL} \times \mathcal{L}_{X\text{-ray}}$. In our case, we rely on the nested sampling Monte Carlo algorithm, ML-FRIENDS (J. Buchner et al. 2014; J. Buchner 2019) implemented in the PYTHON package ULTRANEST² (J. Buchner 2021). We then consider this posterior as a biased distribution in the context of importance sampling. We aim to obtain an approximation to the posterior defined by our total likelihood, which serves as the target distribution for the importance sampling. In that case, we can determine the importance weight, w , as follows:

$$w \propto \frac{\mathcal{L}_{\text{CM-kin}} \times \mathcal{L}_{SL} \times \mathcal{L}_{X\text{-ray}} \times \mathcal{L}_{\text{BCG-kin}} P(\theta)}{\mathcal{L}_{\text{CM-kin}} \times \mathcal{L}_{SL} \times \mathcal{L}_{X\text{-ray}} P(\theta)} \quad (23)$$

$$\propto \mathcal{L}_{\text{BCG-kin}} \quad (24)$$

where $P(\theta)$ represents the priors of the whole model parameters. As we can see, the importance weights vastly simplify in our case, such that they are only proportional to $\mathcal{L}_{\text{BCG-kin}}$. We then use

the systematic resampling algorithm presented in J. D. Hol, T. B. Schon & F. K. Gustafsson (2006) to obtain our approximated posterior.

We note that such a procedure works only if the biased and target posteriors are not too different, such that the resampling keeps as many samples as possible from the biased posterior. We address this issue in our case in Section 4.2.4, where we present the resulting posterior on the kinematic model of the BCG & ICL.

Due to the complexity of the current model, particularly the number of dPIE potentials, we have to speed up the first inference step further. Due to limitations in multi-threading/processing in the Bayesian inference code presented above, the first step would take approximately a month on 60 CPU cores spread across several computing nodes. Hence, we use a particle swarm optimizer (M. R. Bonyadi & Z. Michalewicz 2017) to obtain a maximum likelihood estimate and a sample of models. From this set of points, we model the likelihood using a multivariate normal distribution, fixing the mean to the best fit and scaling the log-likelihood so that the true and surrogate likelihoods have the same value at the best-fitting point. We sample from the surrogate likelihood posterior and estimate a 1D posterior for each parameter using a Gaussian kernel density estimator (KDE). We ensure that the cumulative distribution function (cdf) of the Gaussian KDE encompasses the whole prior by constraining the cdf to be 0 and 1 at the lower and higher bounds of the prior. We use these 1D Gaussian KDE estimations as priors for the ULTRANEST run.

Thanks to this process, we can perform the first step of the inference in 4 to 5 d, including the particle swarm optimizer, the estimation of the approximated prior and the nested sampling run with the true likelihood. In particular, this approach can bias the estimation of the Bayesian evidence (i.e. marginal likelihood) of the nested sampling run. However, it is not usable in our case as we miss $\mathcal{L}_{\text{BCG-kin}}$ from its computation. We rely on nested sampling for its robustness in finding the mode of the posterior distribution and not the marginal likelihood estimate. To ensure that this process does not bias the posterior, we provide in Appendix A the density plots of the approximated posterior from the surrogate, the posterior from the nested sampling run and the approximated posterior obtained via the final importance sampling step.

A possible issue with the importance sampling step is the undersampling of the true posterior. Such a case happens when a small number of samples have a non-negligible weight. In particular, the inclination relative to the plane of the sky influences the resulting mass models. In the case of large ellipticities (i.e. $A/B \gtrsim 2.0$ with A and B the major and minor axis sizes), different

²<https://johannesbuchner.github.io/UltraNest/>

Table 1. Widely applicable information criterion (WAIC) for the different BCG & ICL parametrizations and estimation of the cluster member stellar masses with their associated root mean squared error (RMS) on the multiple image positions. We highlight in bold the two models favoured by the WAIC.

BCG model	SED cluster member	WAIC mean $\pm$ std	RMS (arcsec)
BCG–ML 1	LEPHARE	-33538.94 ± 0.65	0.50''
BCG–ML 1	Delayed	-33548.34 ± 1.47	0.54''
BCG–ML 1	Dbl. power-law	-33536.99 ± 1.05	0.50''
BCG–ML 2	lePhare	-33523.48 ± 0.94	0.53''
BCG–ML 2	Delayed	-33526.09 ± 1.16	0.53''
BCG–ML 2	Dbl. power-law	-33536.19 ± 1.35	0.53''
BCG–ML 3	LEPHARE	-33552.96 ± 1.75	0.53''
BCG–ML 3	Delayed	-33554.16 ± 1.87	0.54''
BCG–ML 3	Dbl. power-law	-33544.20 ± 1.48	0.53''

inclinations can have a V_{rms} difference of more than 20 km/s. Such differences are enough for a set of parameters from the mass model to be located outside of the final posterior. A solution to avoid undersampling is to increase the sampling of the inclination relative to the mass models. Hence, we are sampling 3 different inclinations for each mass model.

4 RESULTS

4.1 Model discrimination

As we defined three parametrizations for the BCG & ICL mass components (Section 2.3) and three estimates of cluster member stellar masses (B25a, Section 4.3), we produce 9 models combining each possibility. To distinguish between them, we rely on the Widely Applicable Information Criterion (WAIC; S. Watanabe 2010), which uses the posterior distribution to compute the criterion value. It is quite similar to the Bayesian Information Criterion (BIC G. Schwarz 1978) or Akaike Information Criterion (AIC H. Akaike 1998), and is a readaptation of such criteria in a Bayesian inference context. It replaces the maximum-likelihood estimate with its average over the posterior distribution, and switches to an effective number of parameters instead. The number of effective parameters is also estimated from the parameters posterior distribution.

Table 1 presents the WAIC estimate performed for each model through the implementation in the ARVIZ library (R. Kumar et al. 2019). We estimate the mean and standard deviation of the WAIC by randomly selecting 1000 models from the posterior, and repeating the operation 1000 times. Focusing on the BCG & ICL parametrization, ‘BCG–ML 2’ models are the best-performing models when averaging over the cluster member SED estimates. They are followed by the ‘BCG–ML 1’ models and then the ‘BCG–ML 3’ models. Except for the model with a double power-law SFH and ‘BCG–ML 2’ parametrization, there is an improvement at a fixed specific SED model from a model with a ‘BCG–ML 3’ to a ‘BCG–ML 1’ and from ‘BCG–ML 1’ to ‘BCG–ML 2’ parametrization. Hence, according to the WAIC and our set of constraints, the best parametrization of the BCG & ICL component is to assume different stellar-mass-to-light ratios for the BCG and the ICL. Providing further freedom to the varying $\Upsilon_{*}^{\text{BCG}}$ does not improve the models, and the last coefficient, $\Upsilon_{*, \text{ICL}}^{\text{BCG}, 3}$, is not constrained.

Regarding cluster member stellar masses, there is no clear preference for one SED model that would perform better for each BCG & ICL parametrization. The SED model from LEPHARE is the best, on average, among all BCG & ICL parametrization.

Although models with LEPHARE do not outperform the other models with the same BCG & ICL parametrization. For ‘BCG–ML 2’ parametrization, it is equally favoured by the WAIC as the model with a delayed SFH, while for ‘BCG–ML 1’ and ‘BCG–ML 3’ parametrizations, it is the double power-law SFH model that has the best WAIC. In contrast to the BCG & ICL parametrization, which can provide a significant increase in the WAIC for most SED models, these cluster member parametrizations have a smaller effect on the overall models and constraints reproduction. Another way to distinguish between these SED models would be to test their prediction of $\Upsilon_{*}^{\text{BCG}}$ after selecting the best BCG & ICL parametrization. This aspect is discussed in Section 5.1, together with the biases on $\Upsilon_{*}^{\text{BCG}}$. Focusing on models with a ‘BCG–ML 2’ parametrization, the Delayed SFH and LEPHARE SED models are the two best models according to the WAIC. The uncertainties on their WAICs are too large to allow us to discriminate between them. To simplify the discussion and focus on a single model, we present the mass distribution results on the LEPHARE SFH model in the following sections, as it has the best-fitting models among all. We provide the posteriors and 1σ CI for all models with a ‘BCG–ML 2’ parametrizations in Appendix A.

As shown in Table 1, the root mean squared error (RMS) on the multiple image positions is similar among all models. We observe the same behaviour for X-ray and the cluster-member likelihood values, which are of similar order across all models. The most impactful likelihood in the WAIC results is the BCG & ICL ones, which account for most of the combined likelihood difference between each model. As this likelihood particularly constrains the BCG & ICL component, it explains why we can discriminate between its parametrizations but not the cluster member one.

4.2 Reproduction of the constraints

In this section, we present how the selected mass model (i.e. LEPHARE SFH and ‘BCG–ML 2’) reproduces the constraints on each dataset. We start with lensing and X-rays in Sections 4.2.1 and 4.2.2, respectively. We finish with the two new datasets, i.e. the BCG & ICL and cluster member stellar kinematics, in Sections 4.2.4 and 4.2.3, respectively.

4.2.1 Lensing

We quantify the reproduction of the lensing constraints with the root mean square (RMS) error on the position of the multiply-imaged constraint between the observation and model prediction. Our best model, which uses the ‘BCG–ML 2’ parametrization,

tion and the SED model with a LEPHARE SFH, has achieved a RMS of 0.53 arcsec. This is close to the RMS achieved by the other tested models. In comparison, our worst model according to the WAIC (i.e. ‘BCG–ML 3’ parametrization and SED model with a delayed SFH) has a RMS of 0.54 arcsec. Hence, lensing constraints are not the dominant dataset to distinguish between the models presented here, as they all perform equivalently. As we use the same set of images as B24, it is the closest comparison we can make with pre-existing models. In that case, the B24 model without a B-spline perturbation achieves an RMS of 0.60 arcsec, which highlights that we improve the reproduction of the lensing constraints with our new treatment of the BCG & ICL and cluster members. Our present model has a higher RMS than the B24 model with a B-spline perturbation, although both are close, with RMS values of 0.44 arcsec. We did not try to include such a component here as the B-spline perturbation as defined in B. Beauchesne et al. (2021) are only defined on projected quantities in the plane of the sky, although the BCG & ICL stellar kinematics require the definition of a deprojection in 3D space.

Interestingly, AS1063 has a multiply imaged system with a central image close to the BCG (I. Balestra et al. 2013), which is more sensitive to the BCG & ICL baryons than other systems. On that specific model, the system RMS is of 0.40 arcsec, slightly lower than the overall RMS and the central image has an error of 0.22 arcsec. In comparison, B24 model had a RMS of 0.50 arcsec for that system and 0.51 arcsec for the central image. This system is a good argument in favour of our modelling approach.

4.2.2 X-rays

The mass model of the intra-cluster gas is almost identical to the model obtained in B24. The two parametrizations of the likelihood intrinsic error lead to similar likelihood values according to their variations. Mock distribution of 10^5 likelihoods based on the best-fitting model yields a standard deviation of 87, which is way larger than the likelihood difference between this work and B24, with a $\mathcal{L}_{X\text{-ray}}$ of -33208 and -33189 , respectively. The log-likelihood distribution among the pixels is also close, though with the new error parametrization, it is slightly wider, with a few pixels with worse and better likelihoods. Hence, we can estimate that the addition of the other likelihoods has not affected the estimate of the gas distribution as presented in B24.

4.2.3 Cluster member kinematics

Fig. 4 presents the model-predicted velocity dispersion σ_e as a function of the observational measurements. To appreciate the variation due to the SED model, we provide the results for each model with the ‘BCG–ML 2’ parametrization and their combined results. Each mass model presents similar results in terms of scatter and uncertainties, though some galaxies show significant differences depending on the SED model. This result is reflected by the reduced χ^2 of the best-fitting model of each mass modelling that varies from 0.92 to 1.03 between these three models with an assumed scatter of 35 km/s.

According to the bottom right panel of Fig. 4, the difference between the σ_e estimates relying on different SED models is larger than the statistical uncertainty of the mass models. It is unclear how much the discrepancies are due to the cluster member stellar mass or the interplay between the different components of the mass model. Our results highlight the importance of the selected

SED models, as the mass-model uncertainty is too low to blend them. These discrepancies in σ_e estimates appear to correlate with increasing values, as points with $\sigma_e > 200 \text{ km s}^{-1}$ present larger error bars. To assess such a correlation, we analyse this dataset with a Spearman correlation (C. Spearman 1904) based on σ_e and $\delta_{e,\text{sys}}/\sigma_e$, with $\delta_{e,\text{sys}}$ being half of the width of the 1σ CI. We report a slight anti-correlation of -0.12 with a large p-value of 0.42, suggesting no correlation. The same test as with σ_e and $\delta_{e,\text{sys}}$ gives a correlation coefficient of 0.86, with p-values inferior to 10^{-13} . Hence, the amplitude of the systematic error appears to correlate with the accessible mass range of a cluster member within our parametrization.

4.2.4 BCG & ICL kinematics

Fig. 5 presents the kinematic model of the BCG & ICL component through its V_{rms} . As shown by the cyan distribution, mass models are already in good agreement with the observed stellar kinematics even when not fitted to those data. This agreement is a necessary condition to validate the importance sampling procedure, but it also increases confidence in the mass model parameterization. Indeed, as the inner mass model is driven by prior assumptions on the mass profile, it could have deviated from the true mass without direct constraints. Thanks to the importance sampling procedure, we can significantly reduce the width of the kinematic posterior, particularly in the BCG and its vicinity, as highlighted by the dashed black line representing the BCG half-light radius.

When looking more closely at the reduced χ^2 of each model in the resampled posterior, we obtain a value of $3.60^{+0.23}_{-0.18}$ for its 1σ CI with a minimum of 3.24, while the best-fitting model among all data sets reaches a reduced χ^2 of 4.02. More precisely, the median distance to the data points from the model is below 2 standard deviations for every point, except the fifth and sixth ones, which are at 3.36 and 3.38, respectively. As these two points deviate from the trend adopted by the other data points, their uncertainties are likely not accounting for systematic biases. If we exclude them, we obtain $2.72^{+0.45}_{-0.35}$ for the reduced χ^2 1σ CI with a minimum at 2.02. If our kinematic model posterior agrees with the observed data, it does not include a model that agrees with the measurement uncertainties. It may highlight a limitation of our kinematic modelling or an undersampling of the posterior due to our two-step optimization procedure.

4.3 Mass distribution and profiles

Fig. 6 presents the normalized surface mass density or convergence of the mass model. Each cluster-scale mass component is highlighted by dashed contours. Thanks to our highly detailed mass modelling, we can disentangle each mass component. Similarly to B24, we observe a discrepancy between the intra-cluster gas and DM. The gas distribution presents a smaller ellipticity and an asymmetry toward the NE halo. It highlights the importance of considering the gas, even in slowly perturbed clusters. Regarding the BCG & ICL baryons, their contribution is negligible compared to the gas and DM, except in the vicinity of the BCG. However, within a few kpc of the cluster centre, the ellipticity of the BCG & ICL highest surface mass density contour closely matches that of the DM. Moving inward, the two innermost contours of the BCG & ICL show progressively lower ellipticities, indicating a systematic deviation from the DM ellipticity toward the BCG one.

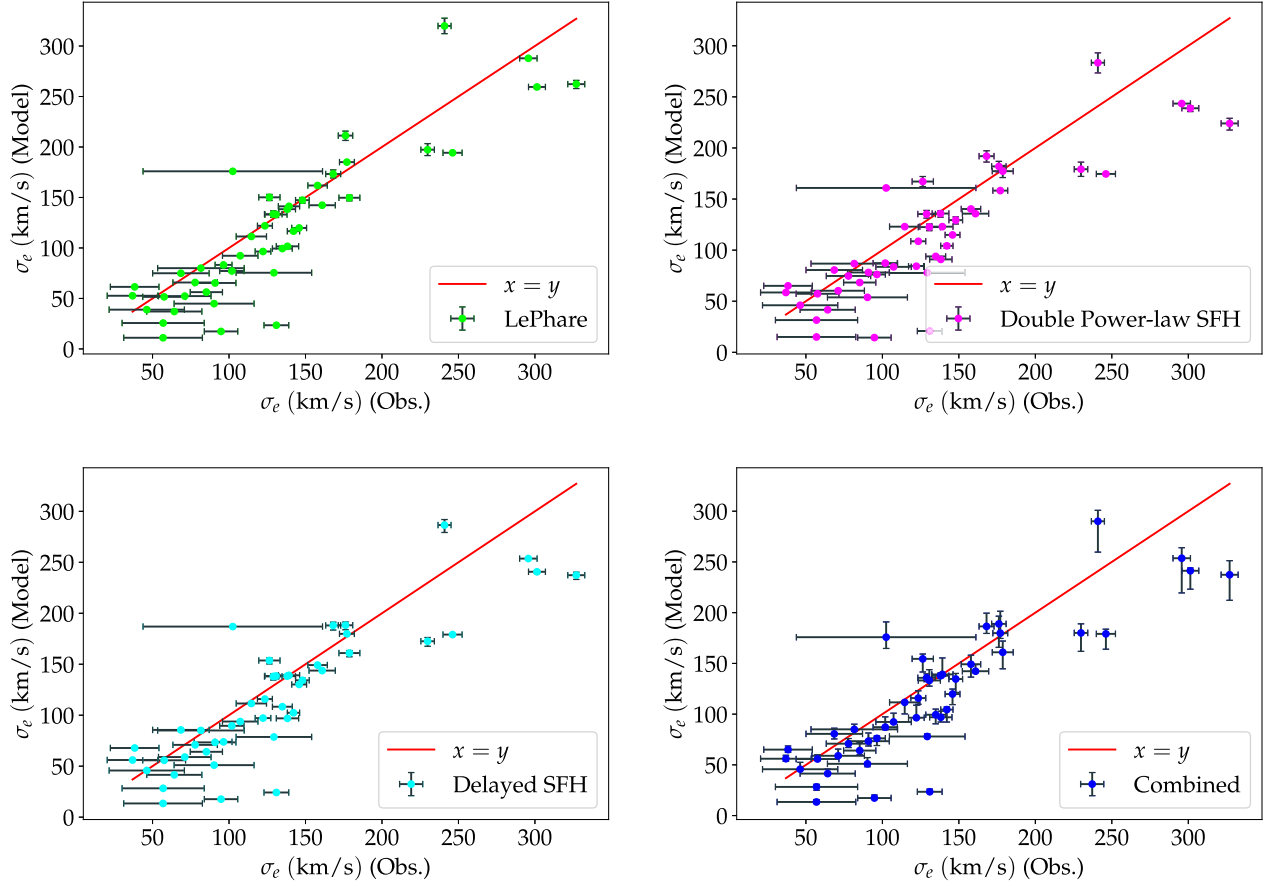


Figure 4. Model predicted σ_e as a function of the observed one, for each SED model. The uncertainties for the observed σ_e are the measurement standard deviations. For the model predicted σ_e , the uncertainties show the 1σ CI from the mass model posterior. The *top left*, *top right*, and *bottom left* panels present the results from the mass model with LEPHARE, the double power-law SFH and the delayed SFH SED models, respectively. The *bottom-right* panel combines the predictions from the three mass models.

Overall, the mass distribution is similar to that of B24 and other lensing studies on AS1063 (G. B. Caminha et al. 2016; P. Bergamini et al. 2019; G. Granata et al. 2022; M. Limousin et al. 2022), which are mostly unimodal with an asymmetry toward the NE halo. As found in B24, it is likely the remnant of a small cluster/group of galaxies infalling in AS1063. Such an event induces gas sloshing, as confirmed by a specific pattern in the gas properties.

Fig. 7 presents the surface mass density profile of each mass component with the total mass in the left panel, and the component profiles relative to the total mass in the right panel. We present the DM distribution through the cluster-scale DM component in the left panel and the overall DM, including the cluster member DM, in the right panel. The cluster member distribution includes both DM and baryons. Thanks to our detailed modelling, we have access to the baryonic fraction and the contribution to each baryonic component. We can see that they reach a maximum of ≈ 40 per cent within the BCG, where the baryons are dominated by the BCG & ICL stellar masses. Beyond a radius of 20 kpc, the ICM becomes the dominant baryonic component, having reached the contribution of the BCG & ICL at that radius, each accounting for 8 per cent of the overall mass. The bulk of the baryonic mass in the form of ICM gas represents ≈ 10 per cent of the total mass up to the end of the strong lensing region at ≈ 250

kpc. The cluster member baryons contribute only locally to the overall baryonic content. Although their total mass can represent as much as the gas distribution when the density of the galaxy is large enough. In particular, there are three cluster members around the BCG which contribute up to 8 per cent of the overall mass locally.

Our mass profile also shows lower statistical uncertainties than the B24 results. Within the strong lensing area ($R < 250$ kpc), our total mass profile presents a 1σ CI of $1.45^{+0.54}_{-0.42}$ per cent of error among the considered radial bins. It represents a 21 per cent reduction compared to the same measurement from the B24 model, which has a 1σ CI of $1.82^{+1.06}_{-0.58}$. Similar error reductions are observed within the BCG and its close surroundings for $R < 30$ kpc. We obtain $3.08^{+0.88}_{-0.52}$ and $4.06^{+3.87}_{-1.39}$ for our model and B24’s one, respectively. In addition to the significant error reduction of the median, the width of the 1σ CI has been divided by a factor ≈ 4 , highlighting the benefit of the BCG stellar kinematics constraints. Thanks to these reduced uncertainties, we obtain an error of $2.15^{+1.16}_{-0.41}$ and $4.68^{+0.24}_{-0.34}$ per cent for $R < 250$ kpc and $R < 30$ kpc, respectively, for the DM-only profiles.

To complete the picture of the mass profiles, we present the mass density of each cluster component with the total mass density in the left panel of Fig. 8. Their relative fraction with respect to the total mass is shown in the right panel of Fig. 8. We add a fit

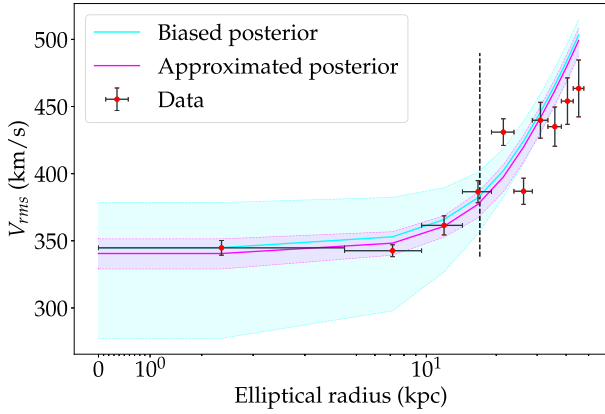


Figure 5. Stellar kinematic data and model of the BCG & ICL component as a function of the elliptical radius in kpc. The cyan distribution represents the biased posterior before the importance sampling step, while the magenta one is the approximated posterior after the resampling procedure. The scattered points represent the observational data points. The elliptical radius is defined with the same ellipticity as used for the fit of the V_{rms} . The shaded area represents the 3σ uncertainties of each posterior, while the plain lines highlight the median. The errors of the data points represent the standard deviation of the V_{rms} fit. The dashed black vertical lines show the BCG half-light radius as measured by L. Tortorelli et al. (2018) of 17 kpc.

by a generalized Navarro–Freu–White profile (gNFW; H. Zhao 1996) of the total mass density as a qualitative comparison to the model mass slope. In contrast to the projected profile presented in Fig. 7, there is a clear domination of the baryonic mass of the BCG up to around 20 kpc, which is approximately the BCG half-light radius (i.e. 17 kpc). The ICM becomes the main baryonic contributor at radii larger than the projected counterparts, i.e. around 50 kpc. Considering the total mass, the gNFW profile can provide a good fit to the total mass from 20 kpc and up to the end of the constrained area. Hence, it is unable to account for the BCG mass distribution when it dominates in the inner core, which can be put into perspective with the results of A. B. Newman et al. (2013), who added a dPIE to account for the BCG in addition to the gNFW profile. In our case, we obtain a slope $\beta = 0.14^{+0.03}_{-0.03}$, significantly lower than 1, where 1 is the slope of a NFW profile. Similarly, C. Cerny et al. (2025) presented a sample of cluster mass models where the BCG kinematics was used along with strong lensing to constrain the model parameters. In contrast to this work, in their models the mass components are all modelled with dPIEs, and they rely on a scaling for the cluster member based on the Faber & Jackson law (S. M. Faber & R. E. Jackson 1976). We reproduce their slope measurement $\gamma = -\frac{d \log_{10}(\rho)}{d \log_{10}(r)}$ over 5 to 50 kpc for the DM distribution ($\gamma = 0.182^{+0.008}_{-0.009}$) and the total mass with the BCG & ICL subtracted ($\gamma = 0.68^{+0.040}_{-0.049}$). We obtain a similar discrepancy as C. Cerny et al. (2025) results, which is mostly due to the cluster members distribution, as the cluster scale DM and gas distribution combined leads to $\gamma = 0.186^{+0.008}_{-0.007}$. Hence, our new cluster-member scheme produces results similar to those of the regular Faber & Jackson law scaling. However, this result, as our mass model assumes that all cluster members lie in the same plane in 3D, likely overestimates the influence of cluster members near the BCG, where their projected density is the highest. In particular, their contribution is larger than the cluster-scale and BCG DM component in the inner 15 kpc toward the BCG, which likely highlights the de-projection effect.

Our results highlight the importance of providing an accurate census of the baryonic components to robustly recover the DM profile down to the few kpc of the inner core. In particular, the DM profile’s shape in that area is sensitive to alternatives to the cold DM (CDM) paradigm, such as self-interacting or Fuzzy DM models that feature a cored distribution rather than the cuspy CDM profile. The parameters of these DM models are correlated with the shape and size of these cores, making our modelling well-suited to robustly measure them (A. Robertson et al. 2021).

Another advantage of our self-consistent approach, which relies on disentangling each cluster component, is its ability to assess which baryonic component is the best tracer of DM. It is particularly useful in the context of clusters or groups that do not present lensing features, but deep observations can map some of their baryonic components and then recover an estimated DM profile. For example, the previous analysis by M. Montes & I. Trujillo (2019) provided evidence that the ICL is a good tracer of the DM profile slope.

However, we leave analyses such as measuring DM properties or assessing which baryonic components are the best DM tracers to future work. Before concluding, we must address the reliability of our stellar mass estimates, as inaccurate BCG & ICL stellar content could introduce significant bias into the analysis presented above.

5 DISCUSSION

We choose to focus the discussion on the estimate of the new baryonic components added to the self-consistent multi-probe approach developed in B24. We start in Section 5.1 by analysing our estimate of the BCG & ICL stellar mass in light of the SED estimates provided in B25a (Section 5.3). We follow up in Sections 5.2 and 5.3, where we discuss the probable biases of BCG & ICL stellar mass estimates with the neglected BCG super massive black hole (SMBH) and assumption of the BCG & ICL stellar kinematics. We complete the discussion on the BCG & ICL component with its degeneracy with the DM mass profile at Section 5.4. We then discuss our estimate of the SsHMR and its dependence on the initial stellar mass estimated from cluster members in Sections 5.5 and 5.6.

5.1 BCG & ICL stellar masses estimate

Fig. 9 presents the estimates of the stellar-mass-to-light ratio of the BCG & ICL component. It compares the estimates from each mass model with a ‘BCG–ML 2’ parametrization with the SED fit results provided in B25a (Section 5.3). For each mass model, there is an apparent discrepancy between the estimate from LENSTOOL and the SED fit within the BCG (i.e. $R_e = 17$ kpc). In contrast, at larger radii, both estimates agree within their 1σ CI for the LEPHARE SED model. The models with a double power-law and a delayed SFH present a less good agreement with the SED estimate, although $\Upsilon_{*,SED}^{BCG}$ is included in the 3σ CI of $\Upsilon_{*,lt}^{BCG}$. These results are in line with the stellar-mass-to-light ratio mismatch observed in early-type galaxies (ETG) between SED and stellar kinematics estimate (see R. J. Smith (2020) for a review). These mismatches are observed when the kinematic analysis relies on a constant value for the BCG, Υ_* (S. Posacki et al. 2015; S. Lu et al. 2024). We are in a similar case as our varying $\Upsilon_{*,lt}^{BCG}$ parametrization has only one coefficient associated with the BCG. L. Oldham & M. Auger (2018) and T. E. Collett et al. (2018) find evidence for a steep gradient of Υ_* within ETGs

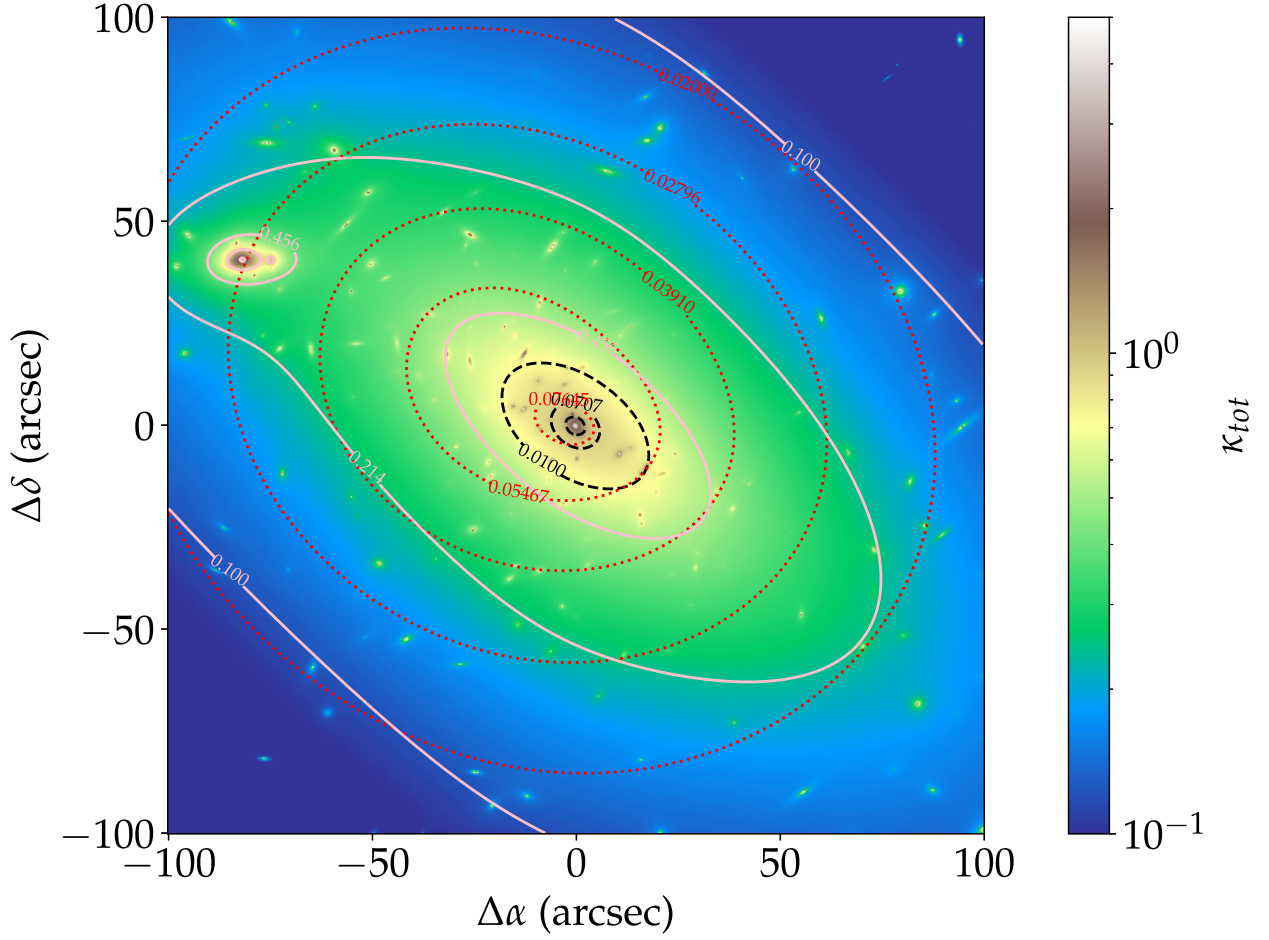


Figure 6. 2D normalized surface mass density (convergence) associated with the total mass distribution in the same field of view as the RGB image presented in Fig. 1. Each cluster-scale mass component is highlighted by contours. The BCG & ICL baryons are highlighted in dashed black, while the intra-cluster gas and DM are in dotted red and solid pink contours, respectively.

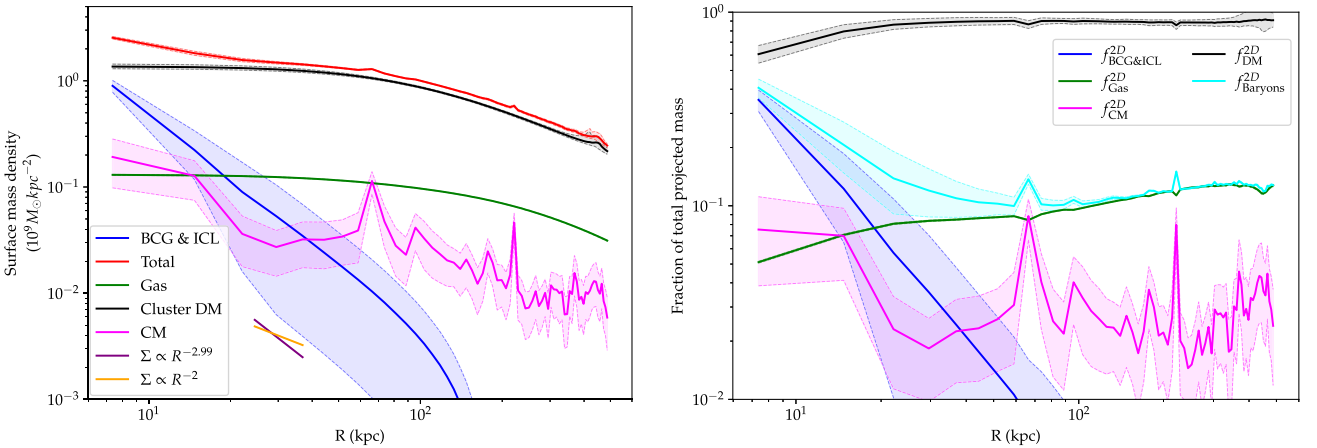


Figure 7. *Left panel:* Surface mass density of the whole cluster (red) and each of its components: the cluster-scale DM component (black), the BCG & ICL baryons (blue), the gas (green) and the cluster members (magenta). The plain lines represent the median of each distribution among the model posterior, while the shaded area shows its 3σ CI. As a qualitative comparison to BCG & ICL profile slope, we present two power-law profiles with a slope of -2 and -2.99 (best fitting solution on the BCG & ICL profile), respectively. *Right panel:* Projected fraction of each cluster mass component relative to the total mass. It presents the mass profiles of each component divided by the total mass density profile. We present BCG & ICL baryons (blue), the gas (green) and the cluster members (magenta) with the same profile as the left panel. For the DM component (black) and baryonic (cyan) mass profiles, we present the overall DM including the cluster-scale and cluster member counterpart, while the baryons include the stellar mass from the BCG & ICL, the cluster member as well as the X-ray gas.

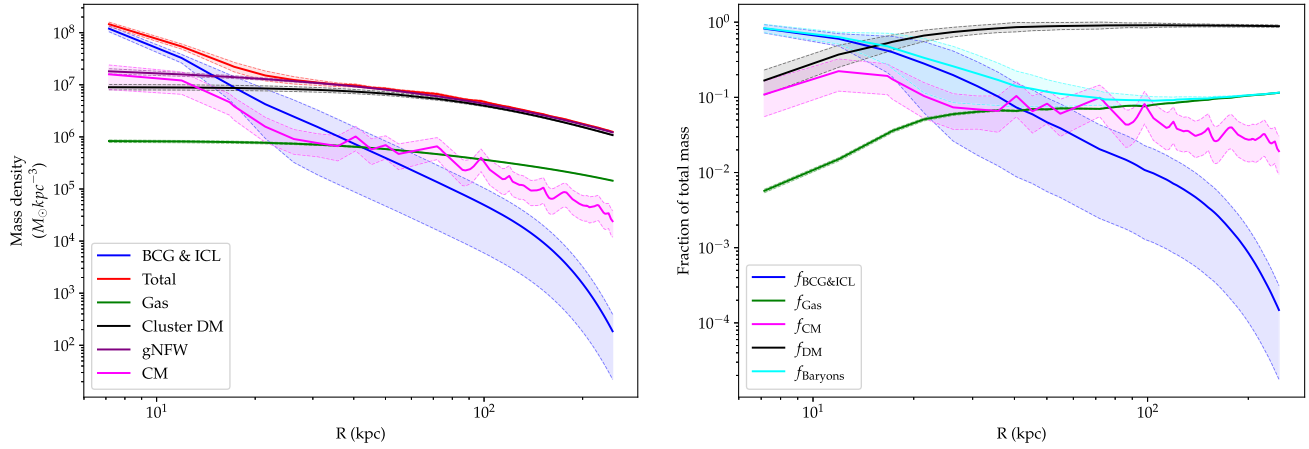


Figure 8. *Left panel:* Mass density of the whole cluster (red) and each of its components: the cluster-scale DM component (black), the BCG & ICL baryons (blue), the gas (green) and the cluster member (magenta). A gNFW fit to the total mass distribution is presented as a comparison to model the mass slope. Plain lines represent the median of each distribution among the model posterior, while the shaded area shows its 3σ CI. *Right panel:* Fraction of each cluster mass component relative to the total mass. For the BCG & ICL (blue) and the gas (green), we present the same profile as in the left panel. For the DM (black) and baryonic (cyan) fraction, we account for the DM and baryons in each component similarly to the right panel of Fig. 7. The colour scheme and meaning of the plain line and shaded area are the same as in the left panel. The combined baryonic profile is presented in cyan.

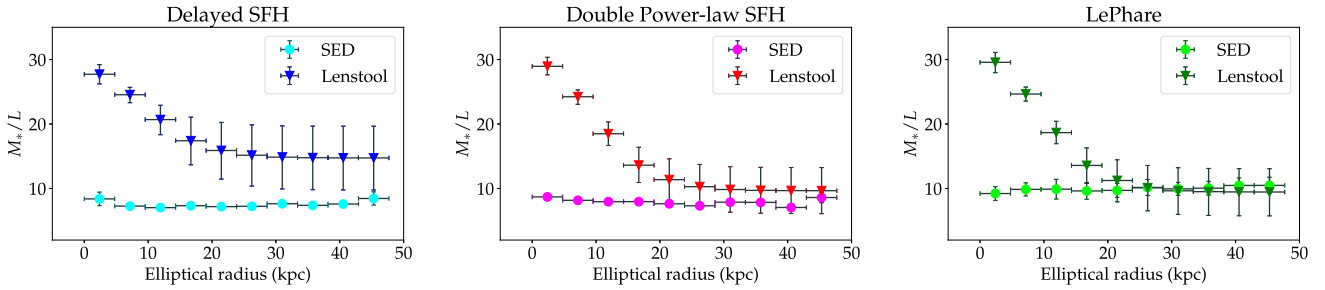


Figure 9. Stellar-mass-to-light ratio of the BCG & ICL component, $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ for the three models with a ‘BCG-ML 2’ parametrization with their equivalent estimated through SED fitting, $\Upsilon_{*,\text{SED}}^{\text{BCG}}$ as a function of the elliptical radii defined in B25a (Section 5.2). Error bars represent the 1σ CI for both $\Upsilon_{*,\text{BCG}}^{\text{BCG}}$ estimates, while errors on the radius represent the width of the elliptical bins in which it is averaged. Each model uses a different SED model. They are presented in the following order from *left to right*: Delayed SFH, double power-law SFH and LEPHARE. The colour schemes highlight $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ (blue, red, and green) with darker colours than $\Upsilon_{*,\text{SED}}^{\text{BCG}}$ (cyan, magenta, and lime). Similar colours regroup models that are associated with the same SED model.

associated with an IMF variation with radius from a E. E. Salpeter (1955) IMF to a Milky-way like IMF that would produce such a mismatch. As reported by K. Mehrgan et al. (2024), such gradients are visible for very small radii, $R \lesssim 1$ kpc, which cannot be probed with our data sets. In addition to this picture, our results suggest that the ICL in AS1063 follows a Milky-way like IMF. Hence, the BCG and the ICL would present a similar IMF at their common border.

These discrepancies in the estimate of the stellar-mass-to-light ratio imply a larger stellar mass for the BCG & ICL component than estimated in B25a (Section 5.3). We obtain 1.70 ± 0.32 , 1.41 ± 0.24 and $1.41 \pm 0.24 \times 10^{12} M_{\odot}$ for the model with a delayed SFH, double power-law SFH and LEPHARE SED models, respectively. These masses are estimated within the same elliptical bins as presented in B25a (Section 5.2). These new estimates are from 57 to 149 per cent higher than their SED counterparts, which are mostly due to the underestimation of the BCG stellar mass. The models with a delayed SFH and a double power-law SFH, which exhibit discrepancies larger than 100 per cent, also yield higher estimates of the stellar mass within the ICL.

To further compare our results and in particular the Υ_* estimate mismatches, we define the following mismatch parameter α :

$$\alpha = \frac{\Upsilon_{*,\text{lt}}^{\text{BCG}}}{\Upsilon_{*,\text{SED}}^{\text{BCG}}} \quad (25)$$

We estimate the posterior distribution of α by propagating the distribution of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ from mass models with a single evaluation of $\Upsilon_{*,\text{SED}}^{\text{BCG}}$ represented by its median. We do not include the uncertainty of $\Upsilon_{*,\text{SED}}^{\text{BCG}}$ as it is negligible for two SED models, and we do not have the full posterior shape for every SED fit, in particular the one made with LEPHARE. We compare α with the results from S. Lu et al. (2024), who analysed the IMF variation for the MaNGA DynPop project. This project aimed to observe 10 000 early-type galaxies (ETGs) to perform dynamical and stellar population synthesis analyses. In particular, they estimate α as a function of the galaxy velocity dispersion, σ_e , for a E. E. Salpeter (1955) IMF. We convert their results to a Milky-way like IMF following the conversion factor provided in A. H. Wright et al. (2017), and we use a value of 341 km s^{-1} for σ_e , which has been measured in the most central bins of our BCG & ICL stellar kinematics measurement.

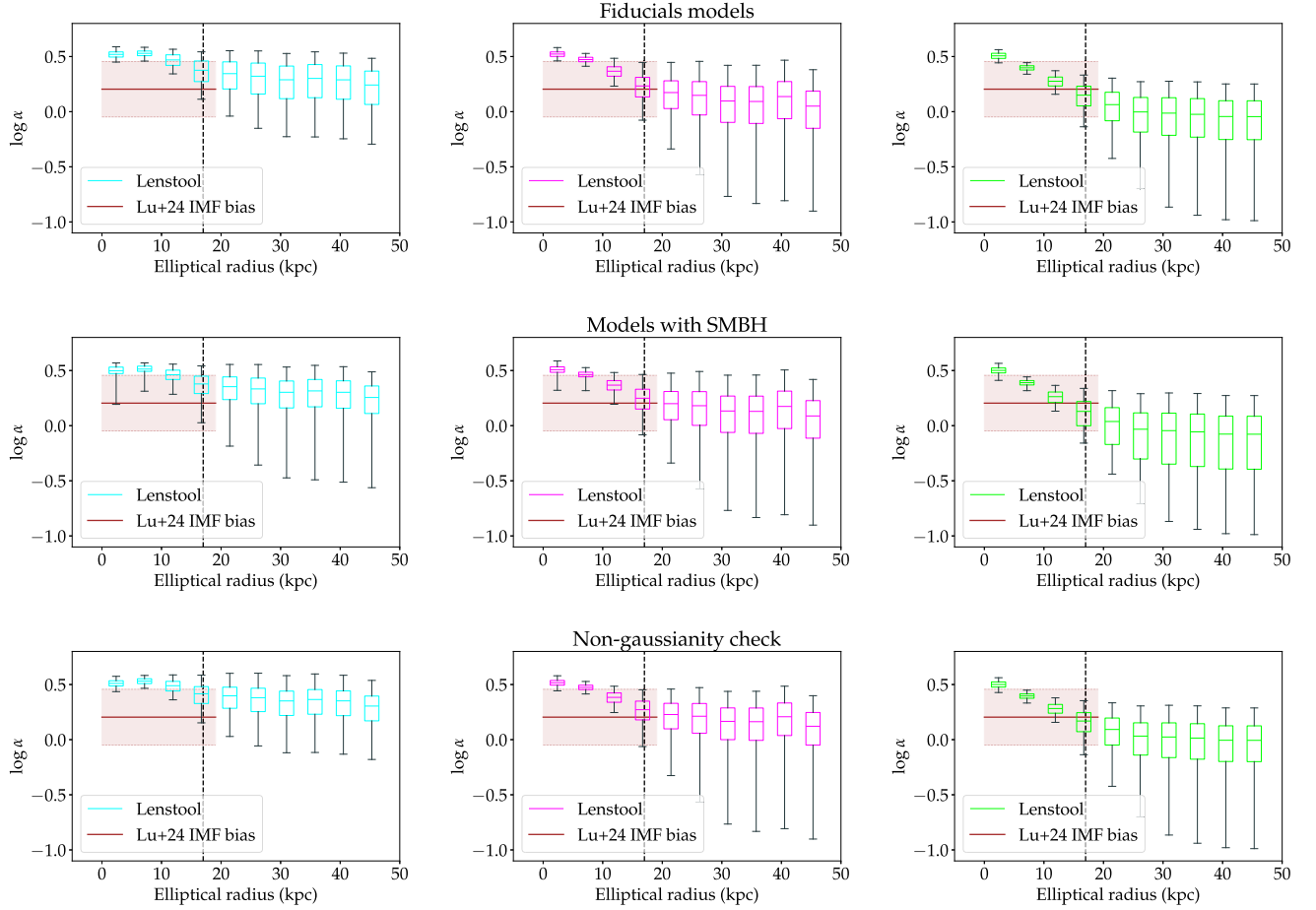


Figure 10. Stellar-mass-to-light ratio mismatch parameter α as a function of the elliptical radius defined in B25a (Section 5.2). In each plot, we highlight the BCG size using its half-light radius, represented by the dashed black line. The shaded brown area represents the expected mismatch due to IMF variation in early-type galaxies as reported by S. Lu et al. (2024). The width of this area represents 3 times the standard deviations from S. Lu et al. (2024) fit, while the plain line represents their best-fitting solution. The 1σ CI and 3σ CI of the posterior distribution of α are represented by the boxes and the associated error bars, respectively. The median is highlighted by the horizontal bars within each box. Each column represents a model with a ‘BCG–ML 2’ parametrization that used different SED models for the cluster member components. They are ordered from *left to right*: BAGPIPES Delayed SFH, BAGPIPES double power-law SFH and LEPHARE. The colour scheme of the boxes represents the associated SED models. Each row represents a different test. *Top row*: Mass models with the methodology outlined in this work. *Middle row*: Mass models with a central SMBH. *Bottom row*: Mass models where the kinematic models of the BCG & ICL component are compared against the observed kinematics, in which we account for the deviation from Gaussian stellar velocity distributions.

The top row of Fig. 10 presents the estimation of α compared to the results of S. Lu et al. (2024). The 3σ CI of each model agrees with the expected mismatch from S. Lu et al. (2024) for the data points within the BCG. However, the agreement gets weaker towards the cluster centre. Indeed, the 1σ CI of only the two/three last data points of the BCG agree with the expected mismatch. It may indicate an overestimate of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ at the first data point if the BCG is not an outlier of the S. Lu et al. (2024) relations. This could be due to the presence of a SMBH with a mass above the $M_{\bullet}-\sigma_e$ relation as we discuss in Section 5.2. The assumption behind the BCG & ICL stellar kinematic models may also have an influence. A JAM model assumes a single ellipticity and a Gaussian line-of-sight velocity distribution. A lower ellipticity would increase the V_{rms} of the kinematic model at a fixed mass, which may reduce the $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ required at the first data point. However, a scheme with varying ellipticity should be employed to avoid an increase of V_{rms} within the ICL. Similarly, a misalignment with the light tracer

and mass distribution that is forbidden by the axisymmetric requirement of the JAM method would result in deviation from a Gaussian distribution for the line-of-sight velocity distribution. A proper assessment of such bias requires a 3D modelling method, such as Schwarzschild’s orbit-superposition (J. Bovy 2015), but is beyond the scope of this work. In Section 5.3, we try to correct the effect of possible deviation from a Gaussian line-of-sight velocity distribution by refitting these velocity distributions without the extra-moments, h_3 and h_4 .

5.2 BCG SMBH

In our modelling, we neglect the influence of an SMBH which mass can be degenerated with the BCG stellar mass. Hence, to assess the reliability of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ estimates to that bias, we add a point mass profile with a fixed position at the BCG centre in the lensing mass model. We modify our stellar kinematic model by adding a

SMBH with its mass contained in the point mass profile of the lensing model. We apply these modelling changes to each ‘BCG–ML 2’ model and reproduce the whole optimization procedure presented in Section 3.

We apply this procedure for two SMBH priors, one assuming the mass from the $M_\bullet$ – σ_e relation compiled in R. C. E. den Bosch (2016) and another wider prior to test the deviation from this relation. Using the relation of R. C. E. den Bosch (2016), we obtain $M_\bullet = 3.66^{+0.62}_{-0.53} \times 10^9 M_\odot$ using the uncertainties from their relation fit and our best-fitting velocity dispersion in the first bin (i.e. $\sigma_{\text{BCG}} = 341 \text{ km s}^{-1}$). The associated prior is a Gaussian distribution with the $M_\bullet$ median as the mean and half of the distance between the two quantiles as a standard deviation. For the prior over the wider range, we use a particularly wide log-uniform prior from 10^7 to $10^{11} M_\odot$.

When the SMBH is assumed from the $M_\bullet$ – σ_e relation, we observe mostly negligible changes to the model parameters and the reproduction of the constraints. The posterior on $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ is mostly the same with a slight decrease of the 1σ CI width within the ICL. It shows that neglecting the BCG SMBH, which would follow the $M_\bullet$ – σ_e relation, does not significantly affect the estimate of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$.

With the larger prior on the SMBH mass, we obtain better reproduction of the stellar kinematics only with the model featuring a double power-law SFH, as in the previous case. However, there are larger changes in the model parameter, and in particular with $\Upsilon_{*,\text{lt}}^{\text{BCG}}$. These changes appear in the width of the 3σ CI, while the median and 1σ CI tend to remain the same as in the model without a SMBH. As our constraints on the SMBH are weak, the modification of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ can be overestimated. The constraints are too weak to set a higher bound on the SMBH mass across all three models. The middle row of Fig. 10 presents the mismatch parameter, α , for these three models. The data points within the BCG highlight the modification of the 3σ CI, which overlap more favourably with the expected mismatch from S. Lu et al. (2024). It indicates that an SMBH with masses above the $M_\bullet$ – σ_e relation can lead to an overestimation of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ within the BCG. However, our dataset is insufficient to assess the presence of such a SMBH.

Another aspect of the SMBH addition is the slight modification of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ within the ICL, as the 3σ CI of the double power-law SFH and LEPHARE SED models are closer to each other than previously. In the case of the double power-law SFH, $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ are in better agreement with $\Upsilon_{*,\text{SED}}^{\text{BCG}}$ than before, as the 1σ CI is now including $\log \alpha = 0$. In contrast, the model with a delayed SFH, presents a narrower 3σ CI which barely includes the horizontal line at $\log \alpha = 0$. Hence, it points toward a more complex correlation with other model components rather than a degeneracy of the mass between the BCG & ICL and SMBH.

5.3 Non-gaussianity of the BCG stellar velocity distribution

In B25a (Section 5.2), we fit the line-of-sight velocity distribution assuming a Gaussian modified by a Gauss–Hermite expansion. In our parametrization, only the coefficients for the third and fourth Hermite polynomials are non-zero (i.e. parameters h_3 and h_4). To account for the V_{rms} differences if the velocity distribution is purely Gaussian, we fit a Gaussian distribution to the velocity distribution obtained in B25a. It mostly affects points in the ICL, as they exhibit larger h_4 values of around 0.05. Hence, we observe an opposite pattern from R. Gavazzi (2005), as the deviation from

Gaussianity appears at a larger radius than closer to the cluster centre.

We reproduce the importance sampling procedure with those modified V_{rms} , and the impact on the mismatch parameter is presented in the bottom row of Fig. 10. It does not change the agreement with the (S. Lu et al. 2024) expected mismatch for the BCG, although models tend to prefer slightly higher $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ for the ICL. In particular, the model with LEPHARE has a narrower 1σ CI as its lower tails lean towards higher $\Upsilon_{*,\text{lt}}^{\text{BCG}}$. Hence, the deviation from a Gaussian stellar velocity distribution does not impact $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ in the BCG according to this test. However, it leaves other deviations from the JAM method, such as varying ellipticity, as potential biases for future work.

5.4 Degeneracy between $M_\bullet/L$ and the DM profiles

An additional potential source of the observed discrepancies in the stellar mass-to-light ratio is the assumed DM profile. In the case of the BCG, it is represented by two dPIEs, one for the BCG DM and the other for the main DM halo. As shown by the parameter posterior in Appendix A, the r_{core} of these two profiles converge to values larger than 100 kpc (i.e. 20 arcsec), yielding a cored profile around the BCG. As our mass probes are constraining the total mass, the BCG & ICL stellar mass may account for some DM to accommodate the limitation of dPIEs.

To test such degeneracies, we use the mass profiles presented in Fig. 8, and we swap the DM profiles for other parametrizations such as a gNFW and a power-law profile and allow for variations of the BCG & ICL stellar-mass-to-light ratio. The modelled total density, $\rho_{\text{tot}}^{\text{mod}}$, as the following form:

$$\rho_{\text{tot}}^{\text{mod}}(r) = \rho_{\text{Gas}}(r) + a\rho_{\text{BCG}}(r) + \rho_{\text{DM}}^{\text{mod}}(r) \quad (26)$$

where ρ_{Gas} and ρ_{BCG} are the gas and BCG & ICL stellar mass density from the mass model. We do not account for the cluster member in that test, aside from the BCG. a represent the variation of the stellar-mass-to-light ratio of BCG & ICL while $\rho_{\text{DM}}^{\text{mod}}$ is the DM density from an analytical profile. We are fitting $\rho_{\text{tot}}^{\text{mod}}$ on the total density of the model without the cluster members in log space, and we assume an error of 0.1 dex. We deliberately choose an error larger than the model uncertainty to account for systematic error due to the mass model D. We perform that fitting for each ‘BCG–ML 2’ model with three different parametrizations of the DM profile. Two of these parametrizations are a gNFW and a power-law profile. For the third parametrization, we use a similar formula as for the BCG & ICL stellar mass density but with the DM, which we denote as ‘fiducial’. This last parametrization allows us to measure the effect of the fitting procedure, in particular, the uncertainty expected from the assumed error.

Fig. 11 presents the results for each ‘BCG–ML 2’ model for the three parametrization where each column refers to a different ‘BCG–ML 2’ model and each row a different DM parametrization. When comparing the first and second rows of the figure, adopting a gNFW DM profile reduces the discrepancy between the SED and mass-model stellar mass-to-light ratios, yet the results still indicate an underestimation of the stellar mass by the SED fitting. The power-law profile in the last row shows the largest modification, with the required stellar mass almost halved. In particular, the SED estimation is now consistent with the 3σ CI of the mass model estimation. However, the power-law profile is still not enough to recover the SED estimation within the 1σ CI.

Comparing the uncertainty from the top row of Figs 11 and 9, it is likely that a proper model with a power-law or gNFW profile

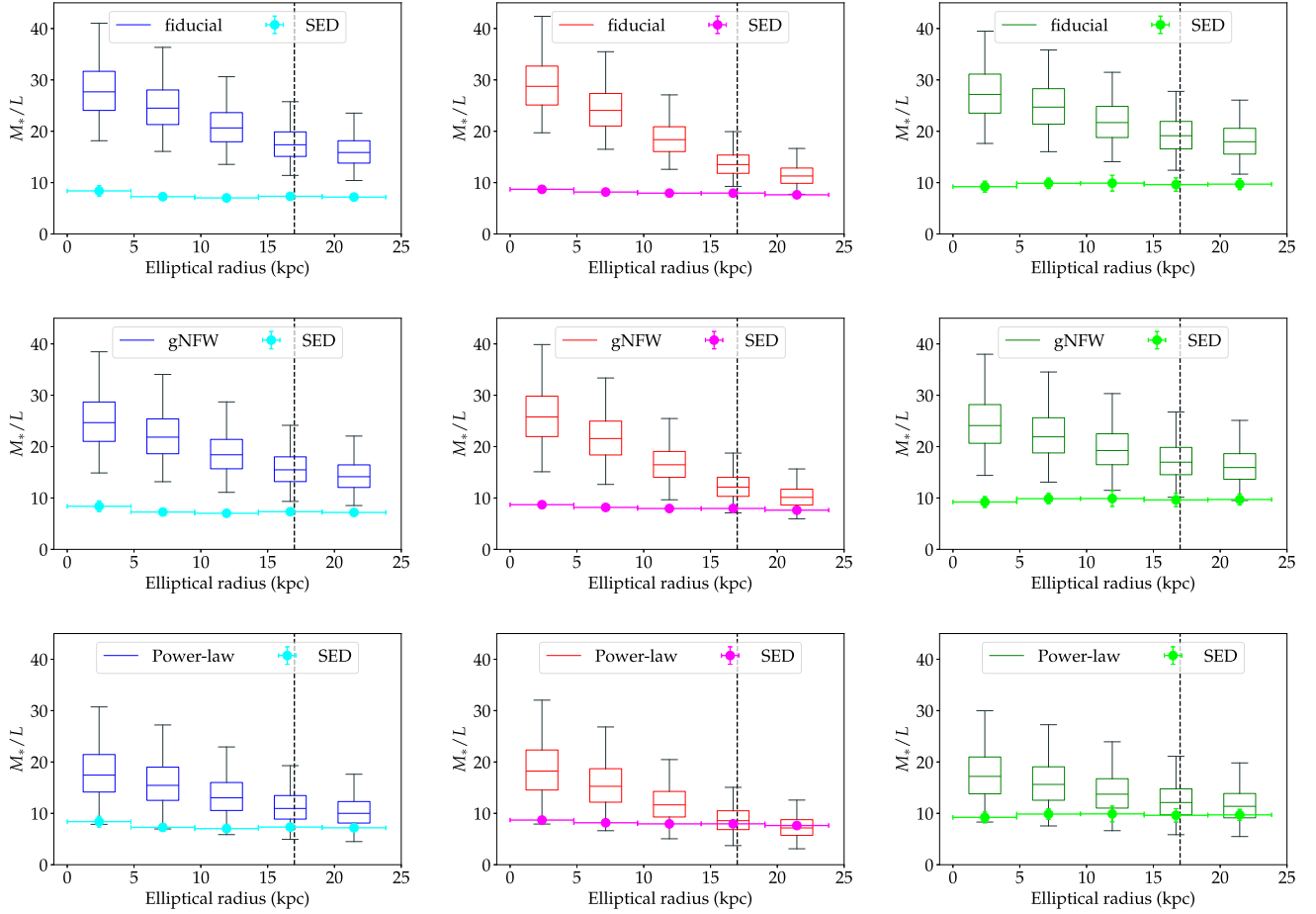


Figure 11. Stellar-mass-to-light ratio M_*/L as a function of the elliptical radius defined in B25a (Section 5.2). In each plot, we highlight the BCG size using its half-light radius, represented by the dashed black line. The 1σ CI and 3σ CI of the posterior distribution of M_*/L from the mass model are represented by the boxes and the associated error bars, respectively. The median is highlighted by the horizontal bars within each box. For the SED estimation of M_*/L , the error bars represent the 1σ CI. Each column represents a model with a ‘BCG–ML 2’ parametrization that used different SED models for the cluster member components. They are ordered from *left to right*: BAGPIPES Delayed SFH, BAGPIPES double power-law SFH and LEPHARE. The colour scheme of the boxes represents the associated SED models. Each row represents a different test: fiducial fit (*Top row*), gNFW fit (*Middle row*) and power-law fit (*Bottom row*).

will have smaller 3σ CI and 1σ CI. The agreement within 3σ CI is unlikely to hold for most models, as the SED estimation is at the limit of that interval.

It appears that the DM profile is likely the largest source of mis-estimation of the stellar mass from the BCG & ICL component. Although using a steeper mass profile, such as a gNFW or power law, is not enough to make both estimates agree. As presented in Fig. 10, the overestimation of the BCG stellar mass is at the 3σ limit of the discrepancies observed by S. Lu et al. (2024). Hence, a steeper DM profile still tends to favour a discrepancy, but shows better agreement with S. Lu et al. (2024) results.

These three tests regarding our estimate of $\Upsilon_{*,\text{lt}}^{\text{BCG}}$ highlight that the main biases are the degeneracy between the DM and stellar mass of the BCG, as the two other tests did not show significant influence on the final results. Such biases were also found in the kinematic studies of ETG, although they have been mitigated by constraining the mass distribution using a 2D mapping of stellar kinematics (M. Cappellari et al. 2012). Future work should therefore tackle this issue by incorporating 2D kinematic constraints, which appear most promising given that baryons are

more centrally concentrated than DM, thereby helping break the degeneracy between the stellar and DM mass components.

5.5 Stellar-to-subhalo mass relation

Thanks to our new cluster member modelling, we can estimate the stellar-to-subhalo mass relations (SsHMR) with the double power-law defined in Section 2.2 for satellite galaxies in cluster cores. In the context of the SsHMR in cluster, satellite galaxies refer to gravitationally bound subhaloes of a cluster-size halo, which in our case translate to the cluster members without the central galaxy/BCG. Such relations are measured in cosmological simulations, and they should vary during the infall of galaxies within their host cluster (A. Niemiec et al. 2019): the increase of stellar mass stops as galaxies are quenched by the cluster environment, while the outside-in stripping of DM by tidal forces of the cluster decreases sub-halo masses. In addition, self-interacting DM could cause an additional ‘evaporation’, thus further decreasing sub-halo masses during their infall (J. Bhattacharyya et al. 2022; E. L. Sirks et al. 2022). As galaxies that have spent a longer

time in their host cluster tend to be located closer to the core, precisely measuring the radial evolution of the SsHMR could allow us to quantify the impact of tidal stripping, as well as possible evaporation due to DM self-interactions. Stacked galaxy–galaxy weak lensing analyses have placed some constraints on the SsHMR (C. Sifón et al. 2015; R. Li et al. 2016; A. Niemiec et al. 2017; C. Sifón et al. 2018; C. Wang et al. 2024), but these measurements do not reach the cluster core, where stripping effects should be highest. To make the fairest comparisons, the types of galaxies (i.e. early or late type galaxies) and their distances to the cluster core should be considered (A. Niemiec et al. 2022). In particular, A. Niemiec et al. (2022) reported a double power-law fit of the SsHMR for satellite galaxies in the IllustrisTNG (R. Weinberger et al. 2017; A. Pillepich et al. 2018) simulation. They reported the fit for the whole sample of satellite galaxies, as well as for different subsets based on galaxy type, redshift bin, or distance to the cluster core. Hence, to match the sample of galaxies considered from IllustrisTNG to our observational measurements, we combine different SsHMR to match the properties of our sample. We sample 1000 SsHMRs, where 19 per cent of the sample follows the SsHMR for satellite late-type galaxies. This proportion matches the proportion of late-type galaxies in our cluster member sample as determined by visual inspection. We classify cluster members as late-type galaxies if they exhibit complex morphologies with possible spiral arms or colour difference from the bulk of the population. The remaining cluster members are modelled assuming the SsHMR derived from satellite early-type galaxies with radii smaller than the maximum cluster-centric distance of our cluster member selection. Hence, the radial dependencies of the SsHMR are only included for early-type galaxies, as this dependence has not been probed by A. Niemiec et al. (2022) for late-type galaxies.

The left panel of Fig. 12 compares the SsHMR from IllustrisTNG with our measurement for each model with the ‘BCG–ML 2’ parametrization. The 1σ CI of the SsHMR of our three models and IllustrisTNG overlap over the entire stellar mass range, from $10^{8.5}$ to $10^{11.5} M_{\odot}$. We notice that our results are systematically above the median SsHMR measured from the simulation. It may highlight discrepancies between our model and the simulation SsHMR on subsamples, such as selections based on galaxy type, e.g. only early- or late-type galaxies. As in the current modelling, we do not distinguish between the two galaxy types; we leave any further analyses for future work.

Each SED model used to estimate the cluster member M_* yields slightly different SsHMRs, which agree at most for $M_* \approx 10^{10} M_{\odot}$ and diverge as M_* increases or decreases. Hence, we observe that the double power-law SFH model provides higher M_* for the most massive cluster members and lower M_* for the least massive. The LEPHARE SED model shows an opposite pattern, while the delayed SFH model is closer to the double power-law one to a lesser extent. Even if their 1σ CI overlap with IllustrisTNG, they present different slopes. It may indicate some discrepancies in the details of the different galaxy populations within AS1063. Increasing the mass range of M_* above $10^{11.5} M_{\odot}$ would likely lead to a disagreement between our results and the simulation, particularly when considering the double power-law SFH model.

5.6 Cluster member IMF variation

In our SsHMR estimate, we use a fixed estimation of the cluster member M_* as the posterior median from the SED fit. These SED models assume a Milky-Way like IMF, which does not agree with

the kinematic modelling of early-type galaxies (S. Posacki et al. 2015; S. Lu et al. 2024). We use the best-fitting relation from S. Lu et al. (2024) for the expected mismatch between SED and kinematic estimate of M_* to correct our SED estimate of M_* . We estimate σ_e for each cluster member based on their mass and light distribution as we do for the cluster member likelihood (see Section 3.3) and assume the best-fitting mass model. We apply this correcting factor to each cluster member in our sample without distinguishing between early- and late-type galaxies.

The right panel of Fig. 12 presents the SsHMR with the correcting factor applied to M_* estimates with the SsHMR from IllustrisTNG. Similarly to the uncorrected SsHMR, the 1σ CI of the observation and simulation overlap. We also observe that the SsHMR followed in AS1063 is systematically higher than the median SsHMR from IllustrisTNG. Although at the massive end, there is a better agreement with the trend between observations and simulations, as there is no tendency for discrepancies at $M_* \approx 10^{10.75} M_{\odot}$ or above.

We try to optimize new mass models with the correct M_* , although the obtained best-fit model σ_e are not in agreement with the ones used to estimate the correcting factor for M_* . These differences in correction factors exceed 50 per cent for some cluster members. Hence, we reject these models. The correcting formula from S. Lu et al. (2024) as a polynomial of degree 1 could be added to our modelling of cluster members with a split between early and late-type galaxies. It could be a new way of estimating IMF variations within the cluster galaxy population.

Our new modelling of cluster members appears to be a robust method for estimating SsHMR, as the overall trend aligns with large-scale simulation results. The estimates of the total masses of cluster members seem consistent across SED models, as the slope difference between them is expected from the difference in stellar masses. As highlighted by the slight differences between our results and IllustrisTNG, anomalies may be found between observations and simulations, which can highlight differences in how the cluster environment strips DM in real cases.

6 CONCLUSION

In this series of two papers, we present the first mass modelling method that allows us to fully disentangle DM at the galaxy- and cluster-scale, but also each baryonic component. To achieve such details, we rely on multiple mass probes to specifically constrain the baryonic components with X-ray surface brightness or multi-band photometry. We combine total mass constraints from the BCG & ICL stellar kinematics and strong lensing to constrain the mass profile from ≈ 5 to 250 kpc, and use cluster member velocity dispersions to measure their total masses. We detail how we gather these sets of constraints in B25a.

Here, we focus on the mass modelling methodology and results. We test different parameterizations for the BCG & ICL components, which favour a varying $\Upsilon_{*}^{\text{BCG}}$ with radius. In our case, a simple model with two coefficients, one for the BCG and one for the ICL, is favoured. In addition to providing a comprehensive mass model of AS1063, with each cluster component disentangled, we reduce the uncertainties on the total mass profile by 20–30 per cent, depending on the considered radii. It allows us to measure the DM only profile with an error of less than 5 per cent over the entire considered region. We obtain these results while presenting one of the most complex parametric mass models.

We then discuss the robustness of the estimates of the stellar masses contained in the BCG & ICL component, as well as cluster

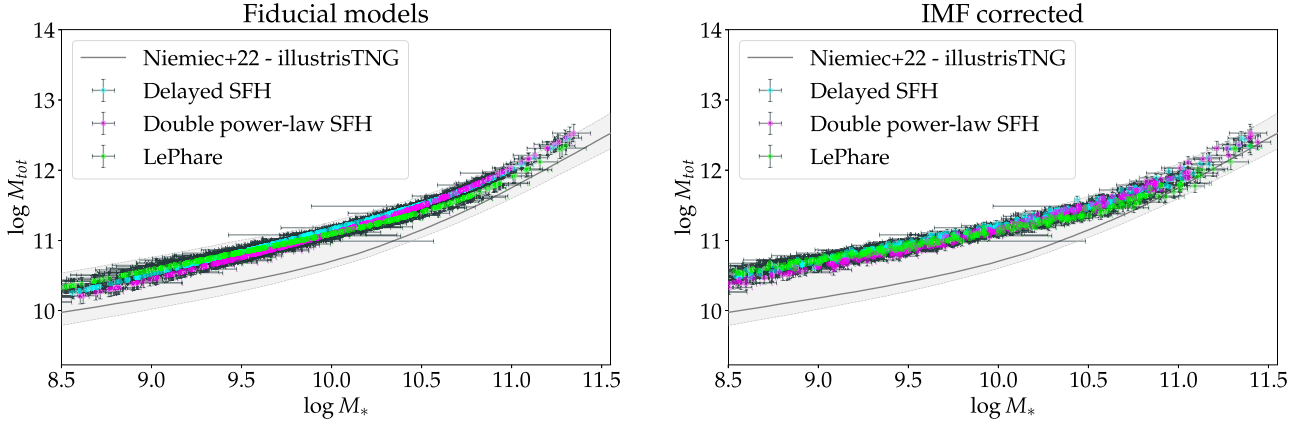


Figure 12. Stellar-to-subhalo mass relation measured in this work compared to the results from the illustrisTNG simulations as reported in A. Niemic et al. (2022). We represent our estimate through each cluster member. The positions of the colour points represent the median of their stellar masses, as determined by the SED fitting, or the total mass from the model posterior distribution. The error bars show the 1σ CI for those distributions. Each colour represents different SED models with the Delayed SFH, double power-law SFH and LEPHARE models shown in cyan, magenta and lime, respectively. The results from the illustrisTNG simulation are represented by the grey area, which shows the 1σ CI, while the plain lines show the median. The *left panel* presents stellar-to-subhalo mass relations as defined in the mass models, while the *right panel* presents the same relation with the stellar masses being corrected for the IMF variation as reported in S. Lu et al. (2024). The IMF correction is based on the model-predicted velocity dispersion, σ_e , for the best-fitting model.

members. We found that the stellar mass of the ICL agrees at 3σ with a Milky-Way like IMF as fitted through three different SED models. We find that the BCG requires a larger $\Upsilon_{*}^{\text{BCG}}$ that can be explained by the IMF variation observed in early-type galaxies, although AS1063 BCG is in the tails of the early-type galaxies distribution modelled by S. Lu et al. (2024). We look for over- or under-estimates of $\Upsilon_{*}^{\text{BCG}}$ by adding a SMBH in the BCG, accounting for deviations from a Gaussian velocity distribution and looking at its degeneracy with the DM mass profile. We found that the most significant bias is the degeneracy with the DM profile, although a gNFW or power-law profile still yields $\Upsilon_{*}^{\text{BCG}}$ larger than the SED estimation. However, with those steeper DM profiles, the discrepancies are in better agreement with S. Lu et al. (2024) results. Besides this degeneracy, we found that only a SMBH with a mass above the $M_{\bullet}-\sigma_e$ relation, we find no significant influence on $\Upsilon_{*}^{\text{BCG}}$.

Regarding cluster members, we analyse our estimates of their total and stellar masses by comparing the SsHMR included in the model with results from the large-scale simulation IllustrisTNG (A. Niemic et al. 2022). We find a 1σ CI agreement between the observational and simulated relations, although with some discrepancy at the higher mass end ($> 10^{10.75} M_{\odot}$) for the delayed and double power-law SFH models. When accounting for the IMF variations with the S. Lu et al. (2024) correction formula, we find a slightly better agreement for the massive end, as no data points are outside the 1σ CI of the IllustrisTNG result. More detailed comparisons in future work will enable us to refine the comparisons and potentially identify discrepancies between observations and simulations for subsets of the galaxy population.

In this series, we extend the B24 framework to obtain a complete picture of the cluster mass distribution. This method can be further improved via a better treatment of the constraints or modelling hypotheses. With the development of hardware accelerators, the next logical step appears to be considering the full light information provided by multiple images, similar to what is done routinely with galaxy-scale lenses (e.g. J. Nightingale et al. 2021). At cluster-scale, it is already possible to reconstruct a

multiple image system at fixed mass model (e.g. S. Sharma et al. 2021), although the combination of both the reconstruction and mass model optimization has been out of reach. A. Acebron et al. (2024) is paving the way in that direction as they used a hybrid approach, where they combined the reconstruction of a single lensing system with the usual point-like approximation for the other systems.

Beyond the methodological considerations, applying our method to a sample of clusters, ideally relaxed, will be instrumental in constraining DM properties by allowing for the precise estimation of the DM only profile. Crucially, our approach is the first step toward reproducing the treatment of A. Robertson et al. (2021) for observations aimed at measuring the self-interacting DM cross-sections, as we have an accurate census of the baryonic distribution. This method can also be leveraged to develop similar analyses for alternative DM models (e.g. fuzzy DM profile from H. Y. J. Chan et al. 2022) but also modified gravity, allowing us to set constraints beyond the single scope of the self-interacting DM paradigm.

The broader application of this method depends on the availability of comprehensive multiwavelength datasets, which currently exist for only a limited number of clusters, such as the Hubble Frontier Fields. This work advances cluster mass modelling to new methodological and observational frontiers, establishing the observational framework required for the precise characterization of the DM distribution.

ACKNOWLEDGEMENTS

The authors are grateful to the referee for carefully reading the manuscript and for valuable suggestions and comments, which helped improve and clarify it. The authors thank Massimo Meneghetti, Raphaël Gavazzi, Michaela Hirschmann, and Mireia Montes for their helpful discussions at various stages of the project. The computations were performed at the University of Geneva using Yggdrasil and Boabab HPC services. BB acknowledges the Swiss National Science Foundation (SNSF) for supporting this work. ML acknowledges the Centre National

de la Recherche Scientifique (CNRS) and the Centre National des Etudes Spatiale (CNES) for support. MJ is supported by the United Kingdom Research and Innovation (UKRI) Future Leaders Fellowship ‘Using Cosmic Beasts to uncover the Nature of Dark Matter’ (grant number MR/X006069/1).

DATA AVAILABILITY

The mass models with a ‘BCG–ML 2’ parametrization are released publicly with the paper, and are available at the Mikulski Archive for Space Telescopes as High Level Science Products via <https://doi.org/10.17909/t9-w6tj-wp63>.

REFERENCES

- Acebron A. et al., 2024, *ApJ*, 976, 14
- Akaike H., 1998, *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer New York, New York, NY, p. 199, https://doi.org/10.1007/978-1-4612-1694-0_15
- Atek H. et al., 2025, preprint ([arXiv:2511.07542](https://arxiv.org/abs/2511.07542))
- Balestra I. et al., 2013, *A&A*, 559, L9
- Beauchesne B., Clément B., Richard J., Kneib J.-P., 2021, *MNRAS*, 506, 2002
- Beauchesne B. et al., 2024, *MNRAS*, 527, 3246
- Beauchesne B. et al., 2025, preprint ([arXiv:2509.07762](https://arxiv.org/abs/2509.07762))
- Bergamini P. et al., 2019, *A&A*, 631, A130
- Bhattacharyya J., Adhikari S., Banerjee A., More S., Kumar A., Nadler E. O., Chatterjee S., 2022, *ApJ*, 932, 30
- Bonamigo M. et al., 2017, *ApJ*, 842, 132
- Bonamigo M. et al., 2018, *ApJ*, 864, 98
- Bonyadi M. R., Michalewicz Z., 2017, *Evolutionary Comput.*, 25, 1
- Bovy J., 2015, *ApJS*, 216, 29
- Buchner J., 2019, *PASP*, 131, 108005
- Buchner J., 2021, *J. Open Source Softw.*, 6, 3001
- Buchner J. et al., 2014, *A&A*, 564, A125
- Caminha G. B. et al., 2016, *A&A*, 587, A80
- Cappellari M., 2008, *MNRAS*, 390, 71
- Cappellari M., 2020, *MNRAS*, 494, 4819
- Cappellari M. et al., 2012, *Nature*, 484, 485
- Cerny C., Jauzac M., Lagattuta D., Niemiec A., Mahler G., Edge A., Massey R., 2025, *MNRAS*, 541, 25
- Chan H. Y. J., Ferreira E. G. M., May S., Hayashi K., Chiba M., 2022, *MNRAS*, 511, 943
- Collett T. E. et al., 2018, *Science*, 360, 1342
- Diego J. M., Protopapas P., Sandvik H. B., Tegmark M., 2005, *MNRAS*, 360, 477
- Diego J. M., Tegmark M., Protopapas P., Sandvik H. B., 2007, *MNRAS*, 375, 958
- Diego J. M., Broadhurst T., Wong J., Silk J., Lim J., Zheng W., Lam D., Ford H., 2016, *MNRAS*, 459, 3447
- Diego J. M. et al., 2026, preprint ([arXiv:2601.11704](https://arxiv.org/abs/2601.11704))
- Dutton A. A. et al., 2011, *MNRAS*, 417, 1621
- Elíasdóttir Á. et al., 2007, preprint ([arXiv:0710.5636](https://arxiv.org/abs/0710.5636))
- Faber S. M., Jackson R. E., 1976, *ApJ*, 204, 668
- Gavazzi R., 2005, *A&A*, 443, 793
- Granata G. et al., 2022, *A&A*, 659, A24
- Hol J. D., Schon T. B., Gustafsson F. K., 2006, *2006 IEEE Nonlinear Statistical Signal Processing Workshop*, p. 79
- Hu W., Barkana R., Gruzinov A., 2000, *Phys. Rev. Lett.*, 85, 1158
- Hyde J. B., Bernardi M., 2009, *MNRAS*, 396, 1171
- Johnson T. L., Sharon K., Bayliss M. B., Gladders M. D., Coe D., Ebeling H., 2014, *ApJ*, 797, 48
- Jullo E., Kneib J. P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, *New J. Phys.*, 9, 447
- Kneib J. P., Mellier Y., Fort B., Mathez G., 1993, *A&A*, 273, 367
- Kumar R., Carroll C., Hartikainen A., Martin O., 2019, *J. Open Source Softw.*, 4, 1143
- Li R. et al., 2016, *MNRAS*, 458, 2573
- Limousin M., Beauchesne B., Jullo E., 2022, *A&A*, 664, A90
- Lotz J. M. et al., 2017, *ApJ*, 837, 97
- Lu S., Zhu K., Cappellari M., Li R., Mao S., Xu D., 2024, *MNRAS*, 530, 4474
- Mahler G. et al., 2023, *ApJ*, 945, 49
- Mehrgan K., Thomas J., Saglia R., Parikh T., Neureiter B., Erwin P., Bender R., 2024, *ApJ*, 961, 127
- Mercurio A. et al., 2021, *A&A*, 656, A147
- Monna A. et al., 2014, *MNRAS*, 438, 1417
- Montes M., Trujillo I., 2019, *MNRAS*, 482, 2838
- Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121
- Newman A. B., Treu T., Ellis R. S., Sand D. J., 2013, *ApJ*, 765, 25
- Niemiec A. et al., 2017, *MNRAS*, 471, 1153
- Niemiec A., Jullo E., Giocoli C., Limousin M., Jauzac M., 2019, *MNRAS*, 487, 653
- Niemiec A., Giocoli C., Cohen E., Jauzac M., Jullo E., Limousin M., 2022, *MNRAS*, 512, 6021
- Nightingale J. et al., 2021, *J. Open Source Softw.*, 6, 2825
- Oldham L., Auger M., 2018, *MNRAS*, 474, 4169
- Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
- Pillepich A. et al., 2018, *MNRAS*, 473, 4077
- Posacki S., Cappellari M., Treu T., Pellegrini S., Ciotti L., 2015, *MNRAS*, 446, 493
- Postman M. et al., 2012, *ApJS*, 199, 25
- Richard J. et al., 2014, *MNRAS*, 444, 268
- Robertson A., Massey R., Eke V., Schaye J., Theuns T., 2021, *MNRAS*, 501, 4610
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Schwarz G., 1978, *Ann. Statist.*, 6, 461
- Shajib A. J., Treu T., Birrer S., Sonnenfeld A., 2021, *MNRAS*, 503, 2380
- Sharma S., Richard J., Yuan T., Patricio V., Kewley L., Rigby J. R., Gupta A., Leethochawalit N., 2021, *MNRAS*, 505, L1
- Sifón C. et al., 2015, *MNRAS*, 454, 3938
- Sifón C., Herbonnet R., Hoekstra H., van der Burg R. F. J., Viola M., 2018, *MNRAS*, 478, 1244
- Simon D. A., Cappellari M., Hartke J., 2024, *MNRAS*, 527, 2341
- Sirks E. L., Oman K. A., Robertson A., Massey R., Frenk C., 2022, *MNRAS*, 511, 5927
- Smith R. J., 2020, *ARA&A*, 58, 577
- Spearman C., 1904, *Am. J. Psychol.*, 15, 72
- Spergel D. N., Steinhardt P. J., 2000, *Phys. Rev. Lett.*, 84, 3760
- Steinhardt C. L. et al., 2020, *ApJS*, 247, 64
- Suyu S. H., Halkola A., 2010, *A&A*, 524, A94
- Suyu S. H. et al., 2012, *ApJ*, 750, 10
- Tortorelli L. et al., 2018, *MNRAS*, 477, 648
- van den Bosch R. C. E., 2016, *ApJ*, 831, 134
- Wang C. et al., 2024, *MNRAS*, 528, 2728
- Watanabe S., 2010, preprint ([arXiv:1004.2316](https://arxiv.org/abs/1004.2316))
- Weinberger R. et al., 2017, *MNRAS*, 465, 3291
- Wright A. H. et al., 2017, *MNRAS*, 470, 283
- Zhao H., 1996, *MNRAS*, 278, 488

APPENDIX A: MASS MODEL PARAMETERS AND POSTERIORS

In this appendix, we present the model posteriors for the three models with a ‘BCG–ML 2’ parametrization. The 1D posterior distributions of each model parameter are presented in Fig. A1, Figs A2 and A3. It shows the models with a delayed SFH, a double power-law SFH and the LEPHARE SED model, respectively. In our optimization procedure detailed in Section 3.5, we use three different distributions that are presented in different colours. To speed up the inference process, we use a first approximation of

Table A1. Model parameters distributions of each mass component for the model with a ‘BCG–ML 2’ parametrization and the delayed SFH model. Starting from the top, the first two tables present the mass component modelled with dPIEs, the third table shows the MGE component, and the last one shows the cluster member population parameters. Each parameter is represented as median^{84percent limit}_{16percent limit} (best-fit) except the ones fixed a priori that are shown in brackets. $\Delta\alpha$ and $\Delta\delta$ are the coordinates of dPIE haloes relative to the BCG centre at $\alpha(\text{J2000}) = 342.183213$ and $\delta(\text{J2000}) = -44.530897$. The lt label refers to the velocity dispersions of dPIEs as implemented in LENSTOOL.

ID	$\Delta\alpha$ kpc	$\Delta\delta$ kpc	ϕ degree	ϵ	
DM-main	$5.0^{+1.0}_{-1.1}(5.7)$	$-1.7^{+0.8}_{-0.8}(-2.5)$	$-36.4^{+0.2}_{-0.2}(-36.3)$	$0.636^{+0.012}_{-0.011}(0.631)$	
DM-BCG	[0.0]	[0.0]	$-93.7^{+4.2}_{-3.9}(-91.6)$	$0.597^{+0.048}_{-0.052}(0.67)$	
DM-NE	$-399.8^{+14.5}_{-5.5}(-399.3)$	$190.5^{+6.2}_{-5.6}(201.0)$	$1.3^{+9.4}_{-8.8}(20.4)$	$0.624^{+0.075}_{-0.072}(0.643)$	
Gas-1	$-5.0^{+10.7}_{-7.0}(-2.0)$	$-86.2^{+14.3}_{-16.8}(-100.9)$	$107.2^{+7.9}_{-6.1}(96.7)$	$0.175^{+0.027}_{-0.032}(0.163)$	
Gas-2	$-0.8^{+1.4}_{-1.8}(-0.7)$	$-5.5^{+1.2}_{-0.8}(-5.0)$	$-12.4^{+1.3}_{-1.4}(-12.9)$	$0.326^{+0.014}_{-0.017}(0.332)$	
Gas-3	$-106.9^{+4.4}_{-3.6}(-112.3)$	$88.6^{+3.1}_{-4.3}(94.2)$	$-27.6^{+3.1}_{-2.1}(-27.5)$	$0.147^{+0.018}_{-0.018}(0.128)$	
ID	r_{core} kpc	r_{cut} kpc	$\sigma_{0,\text{lt}}$ km s^{-1}		
DM-main	$99.9^{+3.2}_{-2.8}(94.1)$	[3000]	$1179^{+7}_{-7}(1167)$		
DM-BCG	$178.1^{+24.4}_{-22.6}(205.2)$	$420^{+44}_{-48}(419)$	$416^{+20}_{-25}(430)$		
DM-NE	[2.5]	$93^{+44}_{-37}(79)$	$419^{+52}_{-57}(513)$		
Gas-1	$557.7^{+16.5}_{-18.4}(537.0)$	[1250]	$335^{+6}_{-7}(330)$		
Gas-2	$69.9^{+2.6}_{-3.1}(74.7)$	[1250]	$224^{+7}_{-10}(235)$		
Gas-3	$159.3^{+2.7}_{-3.5}(158.3)$	[1250]	$342^{+10}_{-9}(333)$		
ID	$\gamma_{\text{BCG}, 2}^{*, \text{lt} 1}$	$\gamma_{\text{BCG}, 2}^{*, \text{lt} 2}$			
BCG	$30.22^{+2.52}_{-2.23}(32.7)$	$14.94^{+4.8}_{-5.03}(11.09)$			
ID	α_c	δ	γ	M_1	N
				$10^{10} \text{ M}_{\odot}$	
Gal.	$18.62^{+1.36}_{-1.75}(18.49)$	$0.38^{+0.04}_{-0.03}(0.35)$	$1.06^{+0.28}_{-0.22}(0.97)$	$8.31^{+1.09}_{-1.18}(8.1)$	$2.69^{+0.34}_{-0.38}(2.74)$

the posterior based on a multivariate normal distribution, shown in blue. We use the 1D posterior of this multivariate to transform our parameter space to increase the volume represented by the model posteriors with respect to the priors. The posteriors obtained by the nested sampling runs are presented in red. In this run, parameters are not constrained by the BCG & ICL kinematics. The last distribution presented in green is the final approxi-

mation of the posterior obtained after the importance sampling step to incorporate the BCG & ICL constraints.

We are presenting tabulated values of each parameter with their 1σ CI and best-fitting values in Tables A1, A2, and A3. These are associated with the models having a delayed SFH, a double power-law SFH or LEPHARE SED model, respectively.

Table A2. Same as Table A1 for the model with the double power-law SFH.

ID	$\Delta\alpha$ kpc	$\Delta\delta$ kpc	ϕ degree	ϵ	
DM-main	$4.7^{+1.1}_{-1.2}(3.8)$	$-1.4^{+0.8}_{-0.8}(-0.5)$	$-36.3^{+0.2}_{-0.2}(-36.3)$	$0.63^{+0.009}_{-0.008}(0.616)$	
DM-BCG	[0.0]	[0.0]	$-93.3^{+3.9}_{-4.0}(-95.8)$	$0.61^{+0.063}_{-0.066}(0.691)$	
DM-NE	$-408.9^{+6.7}_{-4.8}(-416.7)$	$205.0^{+3.5}_{-2.9}(209.3)$	$8.4^{+8.2}_{-8.1}(13.6)$	$0.559^{+0.096}_{-0.1}(0.572)$	
Gas-1	$-11.6^{+9.3}_{-9.6}(0.8)$	$-70.7^{+13.2}_{-13.3}(-83.6)$	$114.3^{+6.1}_{-5.4}(113.2)$	$0.203^{+0.024}_{-0.034}(0.192)$	
Gas-2	$-3.2^{+2.0}_{-2.0}(-2.9)$	$-4.0^{+1.1}_{-1.0}(-3.3)$	$-12.4^{+1.5}_{-1.3}(-12.2)$	$0.307^{+0.018}_{-0.018}(0.308)$	
Gas-3	$-114.0^{+2.5}_{-3.4}(-114.1)$	$96.5^{+3.4}_{-3.2}(94.1)$	$-24.0^{+3.2}_{-3.3}(-26.8)$	$0.123^{+0.014}_{-0.014}(0.146)$	
ID	r_{core} kpc	r_{cut} kpc	$\sigma_{0,\text{lt}}$ km s^{-1}		
DM-main	$97.0^{+2.2}_{-2.4}(93.6)$	[3000]	$1174^{+5}_{-5}(1172)$		
DM-BCG	$153.2^{+17.7}_{-20.9}(184.4)$	$442^{+37}_{-36}(447)$	$398^{+23}_{-19}(406)$		
DM-NE	[2.5]	$124^{+38}_{-34}(110)$	$387^{+40}_{-51}(421)$		
Gas-1	$562.9^{+16.8}_{-15.8}(582.1)$	[1250]	$351^{+5}_{-5}(355)$		
Gas-2	$75.3^{+2.1}_{-1.8}(72.5)$	[1250]	$241^{+6}_{-5}(234)$		
Gas-3	$152.9^{+2.7}_{-3.1}(155.1)$	[1250]	$319^{+5}_{-7}(327)$		
ID	$\Upsilon_{*,\text{lt}1}^{\text{BCG},2}$	$\Upsilon_{*,\text{lt}2}^{\text{BCG},2}$			
BCG	$32.99^{+1.91}_{-2.16}(33.38)$	$9.7^{+3.58}_{-3.64}(1.95)$			
ID	α_c	δ	γ	M_1 $10^{10} \text{ M}_{\odot}$	N
Gal.	$16.09^{+2.21}_{-2.21}(13.77)$	$0.39^{+0.05}_{-0.05}(0.37)$	$0.82^{+0.25}_{-0.25}(0.9)$	$5.72^{+1.33}_{-1.67}(6.44)$	$2.57^{+0.36}_{-0.38}(2.75)$

Table A3. Same as Table A1 for the model with LEPHARE SED model.

ID	$\Delta\alpha$ kpc	$\Delta\delta$ kpc	ϕ degree	ϵ	
DM-main	$4.4^{+1.2}_{-1.0}(4.5)$	$-1.2^{+0.9}_{-0.7}(-1.3)$	$-36.2^{+0.2}_{-0.2}(-36.0)$	$0.63^{+0.011}_{-0.01}(0.623)$	
DM-BCG[0.0]	[0.0]	$-91.2^{+3.9}_{-3.8}(-87.4)$	$0.661^{+0.038}_{-0.055}(0.634)$		
DM-NE	$-401.4^{+32.6}_{-7.0}(-399.4)$	$200.5^{+4.7}_{-4.6}(200.0)$	$4.4^{+10.2}_{-10.6}(-10.4)$	$0.457^{+0.124}_{-0.117}(0.6)$	
Gas-1	$-0.3^{+11.4}_{-11.0}(0.7)$	$-67.5^{+14.6}_{-15.5}(-55.6)$	$104.6^{+9.1}_{-9.3}(98.9)$	$0.115^{+0.035}_{-0.036}(0.046)$	
Gas-2	$-96.9^{+2.1}_{-2.1}(-98.3)$	$73.8^{+0.0}_{-0.0}(73.8)$	$-30.7^{+1.4}_{-1.7}(-32.1)$	$0.179^{+0.011}_{-0.011}(0.196)$	
Gas-3	$4.3^{+1.4}_{-1.4}(3.1)$	$-6.3^{+0.9}_{-0.9}(-6.6)$	$-11.5^{+1.9}_{-1.9}(-9.1)$	$0.317^{+0.018}_{-0.024}(0.291)$	
ID	r_{core} kpc	r_{cut} kpc	$\sigma_{0,\text{lt}}$ km s^{-1}		
DM-main	$95.0^{+2.9}_{-2.6}(95.6)$	[3000]	$1173^{+7}_{-8}(1171)$		
DM-BCG	$229.3^{+32.4}_{-33.0}(231.0)$	$430^{+49}_{-42}(498)$	$444^{+25}_{-29}(432)$		
DM-NE	[2.5]	$62^{+24}_{-19}(63)$	$514^{+55}_{-61}(461)$		
Gas-1	$556.9^{+12.0}_{-13.9}(534.2)$	[1250]	$325^{+6}_{-7}(323)$		
Gas-2	$160.2^{+2.7}_{-2.9}(158.0)$	[1250]	$366^{+5}_{-5}(363)$		
Gas-3	$59.8^{+1.5}_{-1.6}(59.0)$	[1250]	$192^{+3}_{-3}(191)$		
ID	$\Upsilon_{*,\text{lt}1}^{\text{BCG},2}$	$\Upsilon_{*,\text{lt}2}^{\text{BCG},2}$			
BCG-1	$33.79^{+2.26}_{-2.3}(31.8)$	$9.41^{+3.74}_{-3.61}(8.73)$			
ID	α_c	δ	γ	M_1 $10^{10} M_{\odot}$	N
Gal.	$17.68^{+2.1}_{-2.19}(19.99)$	$0.49^{+0.04}_{-0.04}(0.47)$	$1.15^{+0.26}_{-0.28}(1.37)$	$6.98^{+1.42}_{-1.25}(7.77)$	$1.96^{+0.33}_{-0.38}(2.57)$

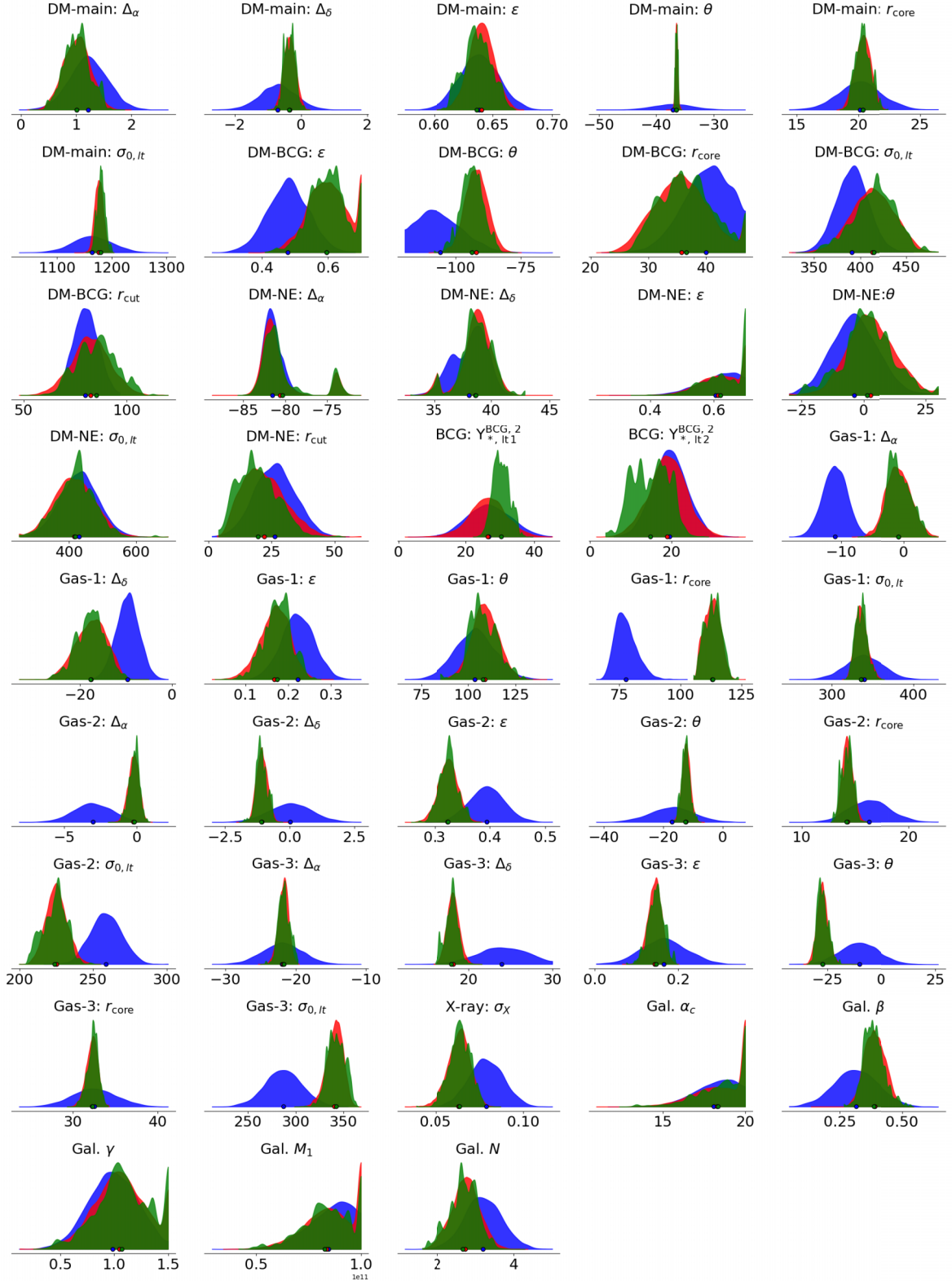


Figure A1. 1D posterior density of each model parameter. The model considered here has a ‘BCG-ML 2’ parametrization and the delayed SFH model. Each colour represents the distribution at a different step of our optimization process as detailed in Section 3.5. The blue distribution represents the initial approximation based on a multivariate normal distribution. The red and green posteriors show the biased and final approximate posteriors or the distributions before and after the importance sampling step. dPIEs parameters are presented with the relative x-coordinate Δ_α (arcsec), the relative y-coordinate Δ_δ (arcsec), ellipticity ϵ , position angle θ (degree), core radius r_{core} (arcsec), cut radius r_{cut} (arcsec), and central velocity dispersion $\sigma_{0,lt}$ (km s^{-1}). For the BCG & ICL component, only its normalization ($\Upsilon_{*,lt,1/2}^{\text{BCG},2}$) is presented. The cluster member parameters are shown by the spatial scale, α_c , the power-law slopes, δ and γ , the turn-off mass, M_1 ($M_\odot$), and the normalization, N .

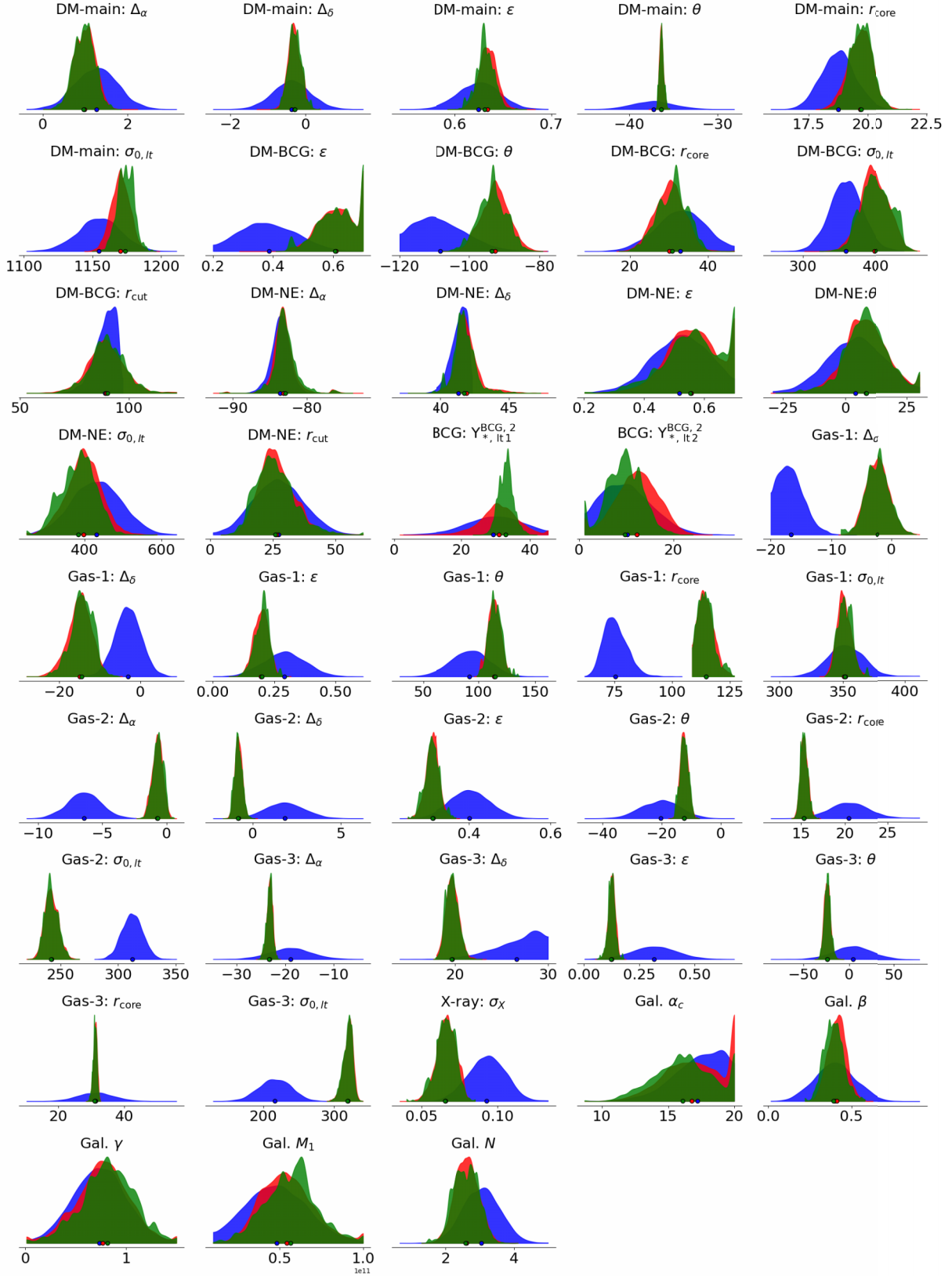


Figure A2. Same as Fig. A1 for the model with the double power-law SFH.

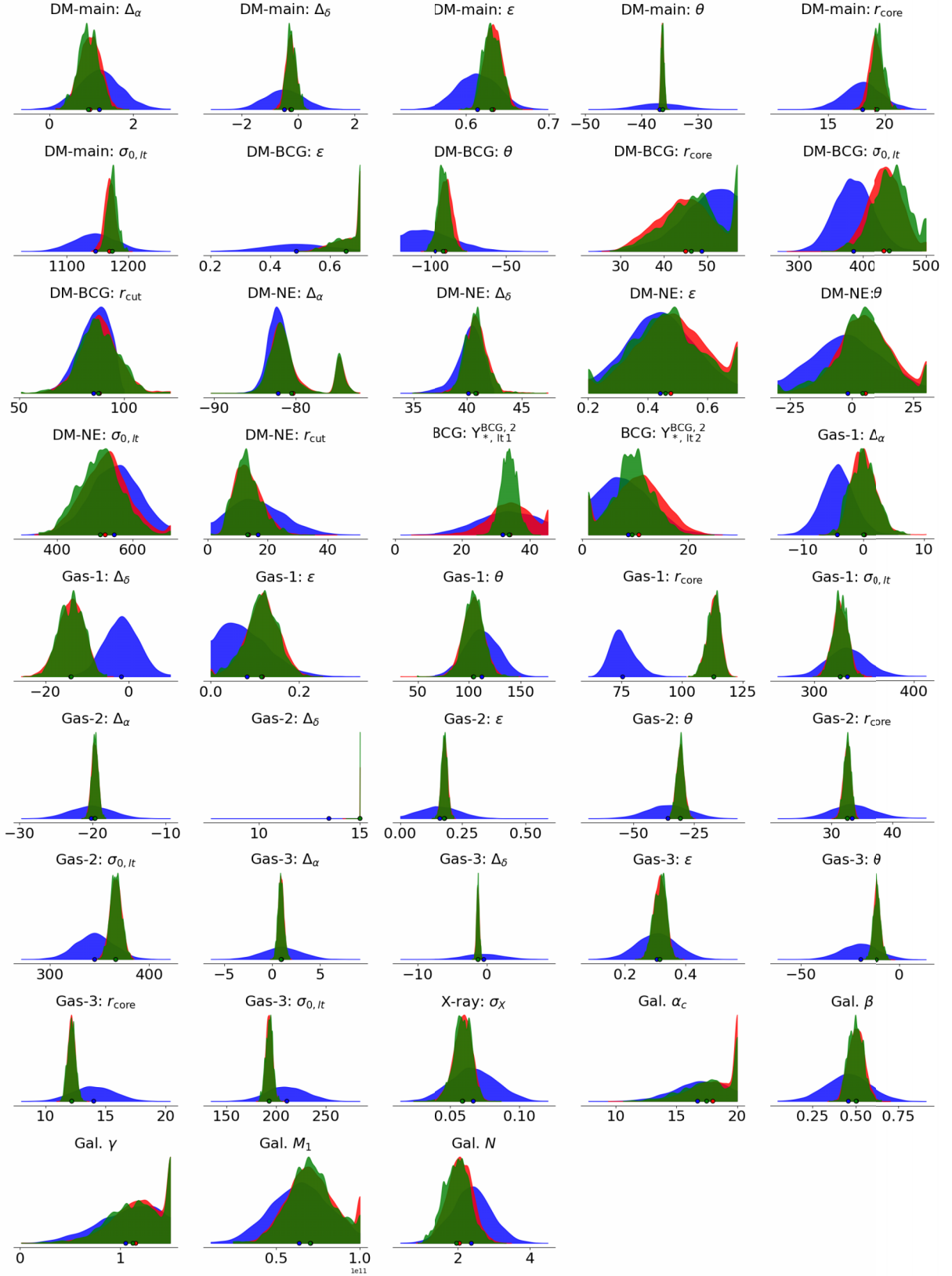


Figure A3. Same as Fig. A1 for the model with LEPHARE SED model.

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