

Does NGC 6397 contain an intermediate-mass black hole or a more diffuse inner subcluster?

Eduardo Vitral and Gary A. Mamon

Institut d'Astrophysique de Paris (UMR 7095: CNRS & Sorbonne Université), 98 bis Bd Arago, 75014 Paris, France
e-mail: vitral@iap.fr

Received 12 October 2020 / Accepted 24 November 2020

ABSTRACT

We analyze proper motions from the *Hubble* Space Telescope (HST) and the second *Gaia* data release along with line-of-sight velocities from the MUSE spectrograph to detect imprints of an intermediate-mass black hole (IMBH) in the center of the nearby, core-collapsed, globular cluster NGC 6397. For this, we use the new MAMPOSST-PM Bayesian mass-modeling code, along with updated estimates of the surface density profile of NGC 6397. We consider different priors on velocity anisotropy and on the size of the central mass, and we also separate the stars into components of different mean mass to allow for mass segregation. The velocity ellipsoid is very isotropic throughout the cluster, as expected in post-core collapsed clusters subject to as strong a Galactic tidal field as NGC 6397. There is strong evidence for a central dark component of 0.8 to 2% of the total mass of the cluster. However, we find robust evidence disfavoring a central IMBH in NGC 6397, preferring instead a diffuse dark inner subcluster of unresolved objects with a total mass of 1000 to 2000 M_{\odot} , half of which is concentrated within 6 arcsec (2% of the stellar effective radius). These results require the combination of HST and *Gaia* data: HST for the inner diagnostics and *Gaia* for the outer surface density and velocity anisotropy profiles. The small effective radius of the diffuse dark component suggests that it is composed of compact stars (white dwarfs and neutron stars) and stellar-mass black holes, whose inner locations are caused by dynamical friction given their high progenitor masses. We show that stellar-mass black holes should dominate the mass of this diffuse dark component, unless more than 25% escape from the cluster. Their mergers in the cores of core-collapsed globular clusters could be an important source of the gravitational wave events detected by LIGO.

Key words. black hole physics – stars: kinematics and dynamics – stars: statistics – methods: data analysis – proper motions – globular clusters: individual: NGC 6397

1. Introduction

When the Laser Interferometer Gravitational-Wave Observatory (LIGO) first detected gravitational waves coming from a stellar-mass black hole merger (Abbott et al. 2016) and then the Event Horizon Telescope (EHT) released the first image of the supermassive black hole (SMBH) in M 87 (Event Horizon Telescope Collaboration 2019), astronomers obtained the most compelling evidence about the existence of those intriguing and particular objects. Yet, black holes (BHs) have been treated as more than a theoretical object for a considerable amount of time, starting in 1939, when Oppenheimer & Snyder (1939) proposed them to be the final step of the life of massive stars ($\geq 10 M_{\odot}$) after their final gravitational collapse into stellar-mass black holes, with masses in between $\sim 3 M_{\odot}$ (Thompson et al. 2020) and $\approx 52 M_{\odot}$ (Woosley 2017), but also when Hoyle & Fowler (1963) identified the then recently discovered quasars (Schmidt 1963) as SMBHs. This latter class of BHs, which reside in the centers of massive galaxies, are responsible for extremely luminous sources in the Universe, such as quasars and active galactic nuclei (AGN), sometimes unleashing powerful jets of relativistic matter, along with outflows that have a profound impact on star formation and galaxy evolution (e.g., Croton et al. 2006; Hopkins et al. 2006). In addition, some black holes may have formed during the early moments of the Universe (e.g., Zel'dovich & Novikov 1966; Hawking 1971), and these primordial BHs may constitute an important mass fraction of black holes below the SMBH mass.

Since there is no theoretical constraint to the mass of a black hole, it would be reasonable to believe that intermediate-mass black holes (IMBHs) could exist, filling the considerable gap between stellar-mass black holes and SMBHs (i.e., with masses between 100 and $10^5 M_{\odot}$). Furthermore, SMBHs are understood to grow by mergers, where the first seeds are stellar-mass BHs (Madau & Rees 2001) or metal-free primordial gas clouds (Loeb & Rasio 1994), so IMBHs may be a transitory stage in the growth of BHs. However, there is currently little evidence for IMBHs (see reviews by Volonteri 2010; Greene et al. 2020), with some important candidates highlighted (e.g., Kaaret et al. 2001; Chilingarian et al. 2018 in dwarf galaxies, and recently Lin et al. 2020 in a globular cluster) and one gravitational wave confirmation (The LIGO Scientific Collaboration & the Virgo Collaboration 2020). Furthermore, IMBHs could help to explain many enigmas in astrophysics, such as filling up part of the dark matter mass budget (e.g., Haehnelt & Rees 1993; Loeb & Rasio 1994) or providing massive seeds for high redshift quasars, whose high masses at such early times represent a challenge to current theories (Haiman 2013). Therefore, great efforts have been undertaken to detect IMBHs to better understand their origin and evolution.

Globular clusters (GCs) appear to be a unique laboratory to test the existence of IMBHs. These quasi-spherical star clusters are known to have old stellar populations, indicating that they formed at early epochs. Their high stellar number densities provide an excellent environment to increase stellar interactions that

could give birth to compact objects. More precisely, contrary to galaxies, the rates of stellar encounters in the inner parts of GCs containing half their stellar mass are sufficiently high to statistically affect the orbits of their stars by two-body relaxation (Chandrasekhar 1942). Moreover, after several relaxation times, the interplay between the negative and positive heat capacities of the inner core and outer envelope causes transfers of energy. This in turn leads to the gravothermal catastrophe, where the core collapses resulting in a steep inner density profile, while the envelope expands. Roughly one-fifth of GCs are believed to have suffered such core-collapse (Djorgovski & King 1986).

Several scenarios have been proposed for the existence of IMBHs in GCs. One is the direct collapse of population III stars (Madau & Rees 2001), but the link of population III stars with GCs is not clear. Another is the accretion of residual gas on stellar-mass BHs formed in the first generation of stars (Leigh et al. 2013), but the availability of the gas is unclear as the first massive stars will blow it out of the GC by supernova explosions. Stellar mergers are a popular mechanism for IMBH formation. Portegies Zwart & McMillan (2002) proposed a runaway path to IMBH formation in GCs, where an initially massive star suffers multiple physical collisions with other stars during the first few Myr of the GC, before they have time to explode as supernovae or simply lose mass. During these collisions, the most massive star will lose linear momentum, ending up at the bottom of the gravitational potential well. At the same time, its mass will grow during the successive stellar mergers, to the point that it will end up as a BH, possibly reaching 0.1% of the GC stellar mass. Miller & Hamilton (2002) proposed a slower process for IMBH formation, where dynamical friction (Chandrasekhar 1943) causes the most massive stellar remnant BHs to sink to the center of the gravitational well over Gyr. Thus a $\geq 50 M_{\odot}$ stellar remnant BH, sufficiently massive to avoid being ejected from the GC by dynamical interactions, would grow in mass through mergers with these other massive BHs as well as other typically massive stars, reaching a mass of $1000 M_{\odot}$ over the Hubble time, which they argued generates IMBHs in some ten percent of GCs. Finally, Giersz et al. (2015) proposed that hard binaries containing stellar-mass BHs merge with other stars and binaries, which can be a fast or slow process.

These models present, however, drawbacks: The short relaxation time needed in the Portegies Zwart & McMillan (2002) scenario usually requires primordial mass segregation in order not to eliminate too many GCs candidates, while the assumption by Miller & Hamilton (2002) of BH seeds above $\approx 50 M_{\odot}$ is not expected as the massive progenitors are fully exploded in pair-instability supernovae (e.g., Woosley 2017).

Unfortunately, attempts to detect IMBHs have been somewhat inconclusive: Dynamical modeling is still dependent on the assumptions concerning the confusion between the IMBH and a central subcluster of stellar remnants (e.g., den Brok et al. 2014; Mann et al. 2019; Zocchi et al. 2019). Furthermore, these analyses usually rely on too few stars inside the sphere of influence of the IMBH, which can lead to false detections (Aros et al. 2020). Besides, searches for signs of accretion indicate no strong evidence for $>1000 M_{\odot}$ black holes in galactic GCs (e.g., Tremou et al. 2018).

In this paper, we analyze the core-collapsed Milky Way globular cluster NGC 6397, and search for kinematic imprints of a central IMBH or subcluster of unresolved objects (hereafter, CUO). We use very precise and deep *Hubble* Space Telescope (HST) proper motions (PMs) from Bellini et al. (2014) and line-of-sight (LOS) velocities from the Multi Unit Spectroscopic Explorer (MUSE) instrument on the Very Large Telescope, to

which we added PMs from the *Gaia* Data Release 2 (hereafter, *Gaia* or *Gaia* DR2).

NGC 6397 has been broadly investigated in previous studies. The PMs lead to plane-of-sky (POS) velocities with equal radial and tangential dispersions, leading to the appearance of isotropic orbits (Heyl et al. 2012; Watkins et al. 2015a with HST for the inner orbits, and Jindal et al. 2019 with *Gaia* for the outer orbits). But this projected velocity isotropy can hide three-dimensional (3D) variations of the anisotropy of the 3D velocity ellipsoid. Kamann et al. (2016) used the JAM mass-orbit modeling code (Cappellari 2008) to fit the LOS velocities obtained with the MUSE spectrograph, and found that a $600 \pm 200 M_{\odot}$ black hole best fits their data, agreeing with the radio emission upper limit of $610 M_{\odot}$ for the same cluster provided in Tremou et al. (2018). On the other hand, numerical *N*-body modeling by Baumgardt (2017) strongly excluded the possibility of an IMBH in NGC 6397, not to say that much of the non-luminous mass measured in this cluster could actually be in the form of unresolved white dwarfs and low-mass stars (Heggie & Hut 1996).

Many of the mass modeling studies performed with this cluster assumed or constrained its surface density parameters to values estimated long ago (e.g., Trager et al. 1995), that may suffer from problems such as radial incompleteness, which is better addressed by missions such as HST and *Gaia*. This lack of accuracy on the surface density can strongly impact the dynamical analysis of NGC 6397.

In the present study, we performed state-of-the-art mass-orbit modeling to analyze the presence of an IMBH or a CUO in the center of NGC 6397 and better measure the radial variation of its velocity anisotropy. Among the different popular mass/orbit modeling (e.g., Mamon et al. 2013; Watkins et al. 2013; Read & Steger 2017; Vasiliev 2019a; see Chap. 5 of Courteau et al. 2014 for a review), we used an extension of the MAMPOSST: Modeling Anisotropy and Mass Profiles of Observed Spherical Systems (Mamon et al. 2013) Bayesian code, which now takes into account both LOS velocities and PMs (MAMPOSST-PM, Mamon & Vitral, in prep.). MAMPOSST-PM was found to perform extremely well in a code challenge using mock data for dwarf spheroidal galaxies, with many similarities to GCs (Read et al. 2020).

We profit from the versatility of MAMPOSST-PM to fit not only cases with a single population of stars, but also with two mass populations, thus allowing for mass segregation. We also provide new fits to the surface density parameters of this cluster by jointly modeling HST and *Gaia* data.

2. Method

2.1. MAMPOSST-PM

We summarize here the main aspects of MAMPOSST-PM (details in Mamon & Vitral, in prep.). MAMPOSST-PM is a Bayesian code that fits parametrized forms of the radial profiles of mass and velocity anisotropy to the distribution of observed tracers (here, GC stars) in two-, three- or four-dimensional projected phase space (PPS): Projected distances to the GC center (hereafter projected radii), combined with LOS velocities and/or PMs. The mass profile is the sum of observed tracers, as well as a possible central BH, and additional dark components. By working on discrete stars, MAMPOSST-PM can probe the inner regions better than methods where velocity or PM data are binned in radial intervals (e.g., van der Marel & Anderson 2010), and an analysis of dark matter on mock dwarf spheroidals

has shown that the conclusions can depend on how the data are binned (Richardson & Fairbairn 2014).

The probability density that a tracer at projected radius R has a velocity vector \mathbf{v} is

$$p(\mathbf{v}|R) = \frac{g(R, \mathbf{v})}{\int g(R, \mathbf{v}) d\mathbf{v}} = \frac{g(R, \mathbf{v})}{\Sigma(R)} \quad (1)$$

where $\Sigma(R)$ is the surface density (SD), while $g(R, \mathbf{v})$ is the density in PPS, such that

$$\begin{aligned} \int \int 2\pi R g(R, \mathbf{v}) dR d\mathbf{v} &= \Delta N_p \\ &= N_p(R_{\max}) - N_p(R_{\min}), \end{aligned} \quad (2)$$

with $N_p(R) = \int_0^R 2\pi R' \Sigma(R') dR'$ being the projected number of points (e.g., stars in a star cluster) in a cylinder of radius R . The likelihood is then calculated as

$$\mathcal{L} = \prod_i p(\mathbf{v}_i | R_i), \quad (3)$$

where the i indices represent individual points (stars).

The observed stars tend to be located in in the GC. But, inevitably, some observed stars will be interlopers (here field stars, FS) mostly foreground in our case, but possibly some background too. MAMPOSST-PM splits the PPS distribution into separate GC and FS components. Equation (1) becomes

$$\begin{aligned} p(\mathbf{v}|R) &= \frac{g_{GC}(R, \mathbf{v}) + g_{FS}(R, \mathbf{v})}{\Sigma_{GC}(R) + \Sigma_{FS}(R)} \\ &= \frac{g_{GC}(R, \mathbf{v}) + \Sigma_{FS} f_{LOS}(v_{LOS}) f_{PM}(\mathbf{PM})/\eta^2}{\Sigma_{GC}(R) + \Sigma_{FS}}, \end{aligned} \quad (4)$$

where \mathbf{PM} is the vector of PMs ($\mu_{\alpha,*}$, μ_δ), $f_{LOS}(v_{LOS})$ and $f_{PM}(\mathbf{PM})$ are the respective distribution functions of field star LOS velocities and PM vectors¹ while $\eta = 4.7405 D$ is the conversion of PM in mas yr^{-1} to POS velocity in km s^{-1} given the distance D to the system, in kpc. The second equality of Eq. (4) assumes that the FS surface density is independent of position.

The GC contribution to the PPS density is the mean local GC velocity distribution function h , averaged along the LOS (Mamon et al. 2013; Read et al. 2020):

$$g_{GC}(R, \mathbf{v}) = 2 \int_R^\infty h(\mathbf{v}|R, r) v(r) \frac{r}{\sqrt{r^2 - R^2}} dr, \quad (5)$$

where r is the 3D distance to the system center, while $v(r)$ is the 3D tracer density (here stellar number density) profile.

MAMPOSST-PM assumes that the local velocity ellipsoid, in other words the local 3D velocity distribution function (VDF), is separable along the three spherical coordinates (r , θ , ϕ):

$$h(\mathbf{v}|R, r) = h(v_r|R, r) h(v_\theta|R, r) h(v_\phi|R, r). \quad (6)$$

MAMPOSST-PM further assumes that the three one-dimensional local VDFs are Gaussian:

$$h(v_i|R, r) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{v_i^2}{2\sigma_i^2}\right). \quad (7)$$

This Gaussian assumption for the local VDFs is much better than the very popular Gaussian assumption for the LOS-integrated

¹ The functional form of the field star PM distribution function, $f_{PM}(\mathbf{PM})$, is provided in Sect. 5.2.3.

VDF, $p(\mathbf{v}|R)$ because velocity anisotropy affects the shape of $p(\mathbf{v}|R)$, in particular that of $p(v_{LOS}|R)$ (Merritt 1987). By measuring the shape of the PPS, hence of $p(\mathbf{v}|R)$, MAMPOSST-PM gets a good handle on velocity anisotropy (see below), and therefore on the mass profile because of the mass-anisotropy degeneracy (Binney & Mamon 1982). MAMPOSST-PM considers the measurement errors by adding them in quadrature to the velocity dispersion for the GC component.

The velocity anisotropy profile β at radius r is defined as

$$\beta(r) = 1 - \frac{\sigma_\theta^2(r) + \sigma_\phi^2(r)}{2\sigma_r^2(r)}, \quad (8)$$

where θ and ϕ are the tangential components of the coordinate system and r is the radial component, while σ_i^2 stands for the velocity dispersion of the component i of the coordinate system.

When considering spherical symmetry in velocity space, the variances of the composite Gaussian VDF $h(\mathbf{v}|(R, r))$ can be written as

$$\sigma_{LOS}^2(R, r) = \left[1 - \beta(r) \left(\frac{R}{r}\right)^2\right] \sigma_r^2(r), \quad (9a)$$

$$\sigma_{POS_R}^2(R, r) = \left[1 - \beta(r) + \beta(r) \left(\frac{R}{r}\right)^2\right] \sigma_r^2(r), \quad (9b)$$

$$\sigma_{POS_t}^2(R, r) = [1 - \beta(r)] \sigma_r^2(r), \quad (9c)$$

where Eq. (9a) is from Binney & Mamon (1982), while Eqs. (9b) and (9c) are from Strigari et al. (2007). In Eqs. (9a)–(9c), the radial velocity variance $\sigma_r^2(r)$ is obtained by solving the spherical stationary Jeans equation with no streaming motions (in particular, rotation):

$$\frac{d(v\sigma_r^2)}{dr} + 2\frac{\beta(r)}{r} v(r)\sigma_r^2(r) = -v(r)\frac{GM(r)}{r^2}, \quad (10)$$

where $M(r)$ is the total mass profile.

The distribution of interlopers in PPS is straightforward: The spatial density is assumed to be uniform (hence Σ_{FS} does not depend on R in Eq. (4)). The distribution of LOS velocities, $f_{LOS}(v_{LOS})$ in Eq. (4), is assumed Gaussian. The two-dimensional distribution of PMs, $f_{PM}(\mathbf{PM})$ in Eq. (4), is found to have wider wings than a Gaussian (see Sect. 5.2.3 and Appendix B.1). The POS velocity errors are not added in quadrature to the velocity dispersions because the PM distribution of interlopers has wider tails (Appendix B.2). Instead, MAMPOSST-PM uses an approximation (Mamon & Vitral, in prep.) for the convolution of PM errors with the PM dispersions.

2.2. Dark component

MAMPOSST-PM allows for a diffuse dark component with a large choice of analytical density profiles. This dark component could represent dark matter or alternatively unseen stars.

Dark matter (DM) may dominate the outskirts of GCs, and perhaps also deeper inside. Contrary to distant GCs (Ibata et al. 2013), the presence of substantial DM in the outskirts of NGC 6397 is unlikely because this GC appears to have an orbit around the Milky Way that takes it from an apocenter of roughly 6.6 kpc to a pericenter of 2.9 kpc in 88 Myr (Gaia Collaboration 2018a), so it should have suffered from dozens of previous episodes of tidal stripping, leaving little DM (Mashchenko & Sills 2005), if there was any in the first place. In particular, the orbital parameters of Gaia Collaboration (2018a), $Z = -0.48 \pm 0.01$ kpc and $W = -127.9 \pm 3$ km s^{-1} , indicate that

NGC 6397 just passed through the Milky Way disk only ≈ 4 Myr ago².

One may ask whether there is substantial DM in the inner regions of GCs like NGC 6397. Using N -body simulations of isolated ultra-compact dwarf galaxies, which appear to be larger analogs of GCs, but with relaxation times longer than the age of the Universe, Baumgardt & Mieske (2008) found that the inner DM component is heated by dynamical friction from the stars, causing a shallower inner DM distribution than that of the stars. Shin et al. (2013) traced the dynamical evolution of NGC 6397 with Fokker-Planck methods, discarding those particles that reach the GC tidal radius. They found that the GC could have contained up to one-quarter of its initial mass in the form of DM and match the present-day surface density and LOS velocity dispersion profiles.

On the other hand, there may be unseen matter in the core of the GC that is not dark matter per se, but composed of unseen stars, for example an inner nuclear cluster of faint stars, possibly white dwarfs, neutron stars, or stellar-mass black holes (Zocchi et al. 2019; Mann et al. 2019) and binary stars (Mann et al. 2019). We thus performed MAMPOSST-PM runs including such a CUO instead of (or in addition to) an IMBH.

2.3. Velocity anisotropy profile

With LOS data, MAMPOSST is able (at least partially) to lift the degeneracy first pointed out by Binney & Mamon (1982) between the radial profiles of mass and velocity anisotropy (Mamon et al. 2013). With the three components of the velocity vector, MAMPOSST-PM is even much more able to disentangle mass and anisotropy profiles (Read et al. 2020; Mamon & Vitral in prep.). With the data used in this paper, we have thus a unique chance of constraining the mass profile of our sources, which is essential to restrain the estimates on the IMBH.

While GCs are often modeled with isotropic velocities, we performed many runs of MAMPOSST-PM with freedom in the anisotropy profile. Indeed, a central IMBH may modify the orbital shapes. Furthermore, the outer orbits of isolated GCs are often thought to be quasi radial (Takahashi 1995). Moreover, the frequent passages of NGC 6397 across the Galactic disk (Sect. 1) could alter the orbital shapes of GC stars.

The anisotropic runs of MAMPOSST-PM used the generalization (hereafter gOM) of the Osipkov-Merritt model (Osipkov 1979; Merritt 1985) for the velocity anisotropy profile:

$$\beta_{\text{gOM}}(r) = \beta_0 + (\beta_\infty - \beta_0) \frac{r^2}{r^2 + r_\beta^2}, \quad (11)$$

where r_β is the anisotropy radius, which can be fixed as the scale radius of the luminous tracer by MAMPOSST-PM³.

² NGC 6397 has just passed through the disk near its apocenter, and probably previously passed through the disk at pericenter, since most GCs highlighted in Fig. D.2 of Gaia Collaboration (2018a) have inclined orbits relative to the Galactic disk, passing through the disk near apocenter as well as near pericenter. Also, the important contributions of the thick disk, bulge/spheroid and dark matter halo to the gravitational potential imply that the tidal effect of the Milky Way is strongest at pericenter.

³ Mamon et al. (2019) found no significant change in models of galaxy clusters when using this model for $\beta(r)$ compared to one with a softer transition: $\beta(r) = \beta_0 + (\beta_\infty - \beta_0) r / (r + r_\beta)$, first used by Tiret et al. (2007).

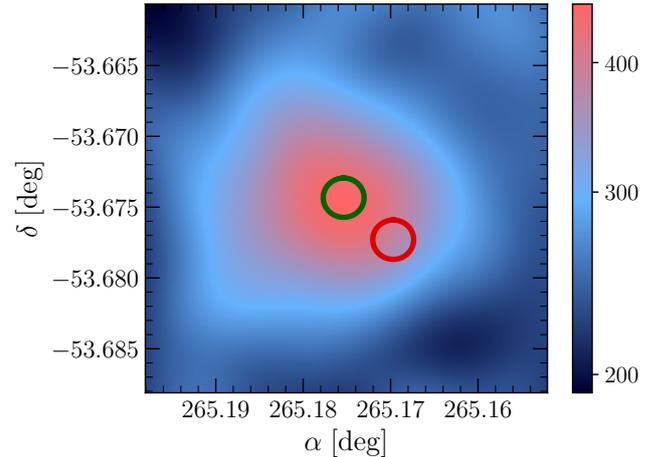


Fig. 1. Smoothed star counts map of NGC 6397 from *Gaia*. The image size is $100''$ on the side (half the HST field of view). The star counts are computed in cells of $0''.33$ and then smoothed by a Gaussian with $\sigma = 0''.10$ (using `scipy.stats.gaussian_kde` in Python, with a bandwidth of 0.2). The color bar provides the star counts per square degree. The green and red circles represent the GC center calculated by Goldsbury et al. (2010) and Gaia Collaboration (2018a), respectively.

3. Global structure and distance of NGC 6397

Our mass-orbit modeling of NGC 6397 (Sect. 2 below) assumes a distance, the knowledge of the GC center, spherical symmetry, and the absence of rotation. We investigate these assumptions for NGC 6397.

3.1. Spherical symmetry

NGC 6397 appears to be close to spherical symmetry. Its elongation on the sky is 0.07 (Harris 2010).

3.2. Center

Our analysis assumes that the IMBH is located at the GC center, so our choice of center is critical. We have considered two centers: (RA, Dec) = (265°175375, -53°674333) (Goldsbury et al. 2010, using HST data) and (RA, Dec) = (265°1697, -53°6773) (Gaia Collaboration 2018a, using *Gaia* data), both in epoch J2000.

We selected the center with the highest local stellar counts. It is dangerous to use our HST star counts for this because the positions are relative to the center, hence need to be rescaled assuming a given center (see Sect. 4.1). We used instead the *Gaia* star counts, which benefit from absolute positional calibration. Figure 1 shows a Gaussian-smoothed map of *Gaia* star counts in the inner $100'' \times 100''$ (half the HST field of view) of NGC 6397, where we overplotted the centers obtained by Goldsbury et al. (2010) and Gaia Collaboration (2018a).

Clearly, the center of Gaia Collaboration (2018a) is less well aligned with the density map than the center of Goldsbury et al. (2010). We therefore select the Goldsbury et al. (2010) center at (265°1754, -53°6743) in J2000 equatorial coordinates.

3.3. Distance

The adopted distance is important because the mass profile at a given angular radius (e.g., r in arcmin) deduced from the Jeans equation of local dynamical equilibrium (Eq. (10)) varies

Table 1. Distance estimates for NGC 6397.

Authors	Method	Observations	Distance (kpc)
Harris (2010)	CMD	Various	2.3
Reid & Gizis (1998)	CMD	HST	2.67 ± 0.25
Gratton et al. (2003)	CMD	VLT	2.53 ± 0.05
Hansen et al. (2007)	CMD	HST	2.55 ± 0.11
Dotter et al. (2010)	CMD	HST	3.0
Heyl et al. (2012)	Kinematics	HST	2.0 ± 0.2
Watkins et al. (2015b)	Kinematics	HST	$2.39^{+0.13}_{-0.11}$
Brown et al. (2018)	Parallax	HST	2.39 ± 0.07
Gaia Collaboration (2018a)	Parallax	<i>Gaia</i>	2.64 ± 0.005
Baumgardt et al. (2019)	Kinematics	<i>N</i> -body	2.44 ± 0.04
Shao & Li (2019)	Parallax	<i>Gaia</i>	2.62 ± 0.02
Valcin et al. (2020)	CMD	HST	$2.67^{+0.05}_{-0.04}$
This work	Kinematics	HST & MUSE	2.35 ± 0.10

Notes. The columns are: (1) authors; (2) method; (3) observations; (4) distance in kpc. Distances based on kinematics seek dynamical models that match the observed LOS and PM dispersion profiles, where Watkins et al. (2015b) used Jeans modeling, while Baumgardt et al. (2019) used *N*-body simulations. The distance of Baumgardt et al. (2019) is a weighted mean with the value given by Harris (1996), while that of Valcin et al. (2020) used the value of Dotter et al. (2010) as a prior.

roughly as $M \propto r\sigma_v^2$, thus as distance D for given LOS velocities and as D^3 for given PMs. Table 1 shows a list of distance estimates for NGC 6397. Recent estimates are bimodal around 2.39 kpc and 2.64 kpc. The CMD-based distance estimates favor the larger distance, the kinematics favor the smaller distance, while the parallax distances point to small (HST) or large (*Gaia*) distances.

We estimated a kinematical distance by equating the LOS velocity dispersion of stars measured with VLT/MUSE (see Sect. 4.3, below) with the HST PM dispersions of the same stars (see Sect. 4.1). To avoid Milky Way field stars, we used the cleaned MUSE and HST samples, for which we had 692 matches, among which 445 with separations smaller than $0''.14$. For these 445 stars, we measured a LOS velocity dispersion of $4.91 \pm 0.16 \text{ km s}^{-1}$, and an HST PM dispersion of $0.439 \pm 0.015 \text{ mas yr}^{-1}$ in the RA direction and $0.442 \pm 0.015 \text{ mas yr}^{-1}$ along the Dec direction (where the uncertainties are taken as the values divided by $\sqrt{2(n-1)}$). This yields a kinematic distance of $4.91/[c(0.439 + 0.442)/2] = 2.35 \pm 0.10 \text{ kpc}$, where $c = 4.7405$ is the POS velocity of a star of $\text{PM} = 1 \text{ mas yr}^{-1}$ located at $D = 1 \text{ kpc}$.

We adopted the lower distance of 2.39 kpc (i.e., a distance modulus of 11.89), for three reasons: (1) We trust more the HST parallax than the *Gaia*-DR2 parallax, given that the former is based on a much longer baseline; (2) Our study of NGC 6397 is based on kinematics, and thus the larger distance would lead to abnormally high POS velocity dispersions compared to the LOS velocity dispersions. (3) It is consistent with our kinematic estimate of the distance. We do not adopt our estimated kinematic distance of $2.35 \pm 0.10 \text{ kpc}$ because it is based on the perfect equality of LOS and POS velocity dispersions, which supposes velocity isotropy, which in turn is not certain. For our adopted distance of 2.39 kpc, 1 arcmin subtends 0.7 pc.

4. Data

4.1. HST data

The HST data were kindly provided by A. Bellini, who measured PMs for over 1.3 million stars in 22 GCs, including NGC 6397 (Bellini et al. 2014). The data for NGC 6397 has a 202 arcsec square field of view (Wide Field Camera) and reached down to less than 1 arcsec from the Goldsbury et al. (2010) center (see the right panel of Fig. 8). This minimum projected radius to the center is smaller than the BH radius of influence $r_{\text{BH}} \sim G M_{\text{BH}}/\sigma_0^2 \simeq 0.11 (M_{\text{BH}}/600 M_\odot) \text{ pc}$ (using the LOS velocity dispersion of Sect. 3.3), which corresponds to $9 (M_{\text{BH}}/600 M_\odot) \text{ arcsec}$, given our adopted distance of 2.39 kpc.

This dataset was provided in a particular master frame shape (for details, see Table 29 from Bellini et al. 2014 as well as Anderson et al. 2008), with both GC center and PM mean shifted to zero. The first step in our analysis was to convert the positions and PMs to the absolute frame.

4.1.1. HST absolute positions

We applied the Rodrigues (1840) rotation formula to shift the relative positions back to their original center to translate the GC stars to their true positions on the sky. We used the center of Goldsbury et al. (2010), which was the one considered by Bellini et al. (2014). We then rotated the subset with respect to its true center, so that the stars originally parallel to the dataset's increasing x axis remained parallel to the right ascension increasing direction. We then verified our method by matching the stars in sky position with *Gaia* (see Sect. 4.4 below).

4.1.2. HST absolute proper motions

The HST PMs were measured relative to the bulk PM of the GC. We corrected the relative PMs of Bellini et al. (2014) with their provided PM corrections. We just added Cols. 4 and 5 to Cols. 31 and 32 from Table 29 of Bellini et al. (2014), respectively. We then converted the relative PMs to absolute PMs by computing the bulk PM of NGC 6397 using the stellar PMs provided by *Gaia* DR2 as explained in Sect. 4.2 below.

The small field of view of HST and the few pointed observations do not allow the observation of sufficiently numerous background quasars to obtain an absolute calibration of HST PMs. On the other hand, the *Gaia* reference frame obtained with more than half a million quasars provides a median positional uncertainty of 0.12 mas for $G < 18$ stars (Gaia Collaboration 2018b) and therefore allows us, by combining its accuracy with HST's precision, to know NGC 6397 PMs with unprecedented accuracy. We will compare the PMs of stars measured both by HST and *Gaia* in Sect. 4.4.

Our HST data had 13 593 stars. The PM precision varies across the field of view because of the different time baselines (principally 1.9 and 5.6 years) in part due to the smaller size of the older HST cameras. Figure 2 shows that the PMs are much more accurate in a large square portion of the field of view, including the western part and extending to the central region, where the baseline is 5.6 years. At magnitude $F606W = 17$, the one-dimensional PM precision is 0.02 mas yr^{-1} in the more precise portion of field of view and 0.08 mas yr^{-1} outside.

4.2. Gaia data

Gaia DR2 presented an overall astrometric coverage of stellar velocities, positions and magnitudes of more than 10^9 stars in the

⁴ The distribution of log separations is strongly bimodal with peaks at $0''.06$ and $1''$.

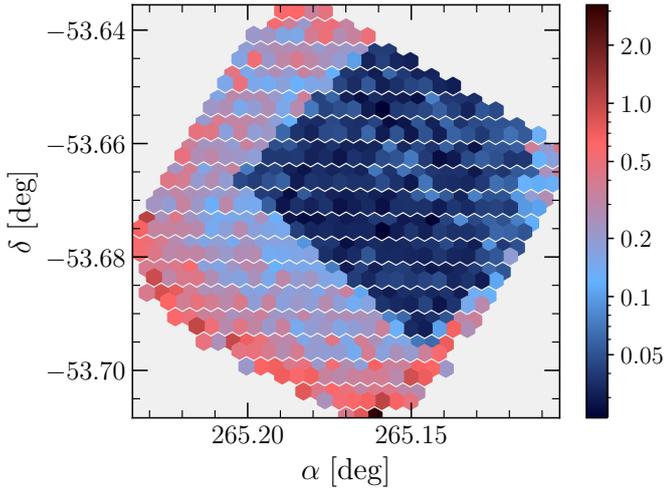


Fig. 2. Map of HST PM errors (semi-major axis of error ellipse, in mas yr^{-1}), color-coded according to the median of each hexagonal cell, using the entire set of HST data (i.e., before any cuts). The total HST field of view is $202''$ on the side.

Milky Way and beyond. *Gaia* measured PMs in over 40 000 stars of magnitude $G < 17$ within a 1° cone around NGC 6397. These PMs have a typical precision of 0.17 mas yr^{-1} at $G = 17$. Since the *Gaia* G and HST $F606W$ magnitudes differ by less than 0.1, one sees that *Gaia* PMs are less precise than those of HST by factors of two to eight depending on the region of the HST field of view. The *Gaia* data was also used to infer the number density profile out to large distances from the GC center.

4.3. VLT/MUSE data

We complemented the PMs using LOS velocities that Husser et al. (2016) acquired with the MUSE spectrograph on the VLT. A mosaic of 5×5 MUSE pointings led to an effective square field of view of 5 arcmin on the side. The bulk LOS velocity is $\langle v_{\text{LOS}} \rangle = 17.84 \pm 0.07 \text{ km s}^{-1}$ (Husser et al. 2016). In a companion article, Kamann et al. (2016) assigned membership probabilities to the stars according to their positions in the space of LOS velocity and metallicity, compared to the predictions for the field stars from the Besançon model of the Milky Way (Robin et al. 2003). The data, kindly provided by S. Kamann, contained 7130 LOS velocities, as well as the membership probabilities.

4.4. Consistency between datasets

4.4.1. Positional accuracy

We used Tool for OPERations on Catalogues And Tables (TOPCAT, Taylor 2005) to check the match positions between catalogs, which we did in our own Python routines that filtered the original datasets (Sect. 5 below). Performing symmetric matches among each pair of datasets, with a maximum allowed separation of 1 arcsec, we obtained median separations of $0''.38$ for the 4455 stars in common between *Gaia* and HST, as well as $0''.01$ for the 4440 stars in common between MUSE and HST.

4.4.2. Proper motions

Many of the 4455 stars in common between HST and *Gaia* have very high *Gaia* PM uncertainties (we saw in Sect. 4.2

that HST PM uncertainties were much smaller, but a few have large values). Limiting to 1179 stars in common with both *Gaia* and HST PM uncertainties below 0.4 mas yr^{-1} , we find that the HST PMs in the (RA, Dec) frame are $(0.006 \pm 0.812, 0.029 \pm 1.239) \text{ mas yr}^{-1}$ above those from *Gaia*, where the “errors” represent the standard deviation. The uncertainties on the means are $\sqrt{1179} = 34$ times lower: $(0.024, 0.036) \text{ mas yr}^{-1}$. The mean shifts in PMs are (0.3, 0.8) times the uncertainties on the mean, and thus not significant. This possible offset in PMs is of no concern as long as we only use HST data, since we analyze them relative to the bulk PM of the GC. Even in MAMPOSST-PM runs with the combined HST+*Gaia* dataset, this shift in PMs is not statistically significant and appears too small to affect the results.

5. Data cleaning

We now describe how we selected stars from each dataset for the mass modeling runs with MAMPOSST-PM. We also used more liberal criteria to select stars in our estimates of the surface density profile (see Sect. 7.1).

For each dataset, we discarded all stars with PM errors above half of $0.394 \text{ mas yr}^{-1}$, corresponding (for our adopted distance) to the one-dimensional PM dispersion of NGC 6397 measured by Baumgardt et al. (2019) for the innermost 2000 *Gaia* stars, whose mean projected radius was $99''.25$. The PM error is computed as the semi-major axis of the error ellipse (Eq. (B.2) of Lindegren et al. 2018):

$$\epsilon_\mu = \sqrt{\frac{1}{2}(C_{33} + C_{44}) + \frac{1}{2}\sqrt{(C_{44} - C_{33})^2 + 4C_{34}^2}} \quad (12)$$

$$C_{33} = \epsilon_{\mu_{\alpha^*}}^2, \quad (13)$$

$$C_{34} = \epsilon_{\mu_{\alpha^*}} \epsilon_{\mu_\delta} \rho, \quad (14)$$

$$C_{44} = \epsilon_{\mu_\delta}^2, \quad (15)$$

where ϵ denotes the error or uncertainty and where ρ is the correlation coefficient between μ_{α^*} and μ_δ ⁵. Our HST data does not provide ρ . We thus selected stars with $\epsilon_\mu < 0.197 \text{ mas yr}^{-1}$, or equivalently for which the LOS velocity error satisfies $\epsilon_{v_{\text{LOS}}} < 0.197 c D = 2.23 \text{ km s}^{-1}$, for $D = 2.39 \text{ kpc}$ and where $c = 4.7405$ (see Sect. 3.3).

Figure 3 shows the PM errors versus magnitude for HST (blue) and *Gaia* (red), as well as the equivalent PM error for the LOS velocity errors of *Gaia* (purple) and MUSE (green). The figure indicates that our maximum allowed PM errors are reached at magnitudes $m_{F606W} = 19.7$ for the worst HST data (those with the shortest baseline, see Fig. 2), and for virtually all magnitudes for the other HST stars. *Gaia* stars reach the maximum allowed PM error at typically $G = 17.5$. The equivalent LOS velocity error limit is reached at magnitude $m_{F606W} = 16.8$ for MUSE, but only at $G = 13.4$ for *Gaia*. The *Gaia* radial velocities are thus of little use for our modeling, as too few of them have sufficiently precise values.

5.1. HST data cleaning

After selecting the stars with low PM errors, we cleaned our HST data in three ways: we discarded stars with (1) PMs far from the bulk PM of the GC; (2) lying off the color-magnitude diagram; (3) associated with X-ray binaries.

⁵ We use the standard notation $\mu_{\alpha^*} = \cos \delta \, d\alpha/dt$, $\mu_\delta = d\delta/dt$.

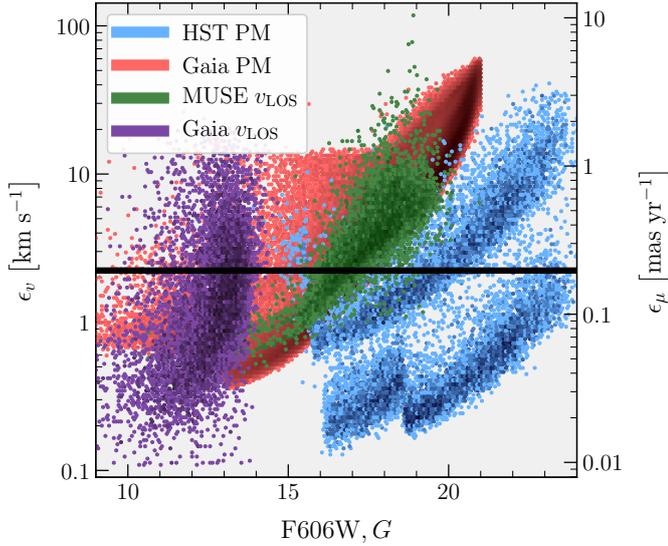


Fig. 3. Proper motion errors and LOS velocity errors (converted to PM errors for distance of 2.39 kpc) for the different datasets. The horizontal line displays $\epsilon_\mu = 0.197 \text{ mas yr}^{-1}$. The three different blue zones correspond to different baselines in the HST PM data. We note that magnitudes G and $F606W$ are close to equivalent (they match to ± 0.1).

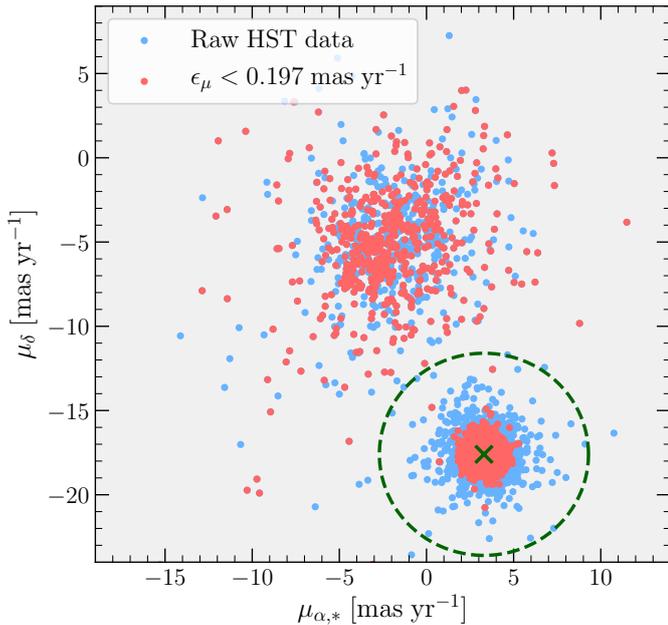


Fig. 4. HST proper motions, corrected to the mean GC PM shown in Table 2. The full set of HST stars is shown in blue, while the subset obtained after applying the low PM error cut (Sect. 5.1.1) are shown in red. The dashed green circle represents the very liberal hard cut on PMs to remove Milky Way field stars and the green cross highlights the mean PM presented in Table 2.

5.1.1. HST proper motion filtering

The higher surface number density of stars in the inner regions of NGC 6397 allows us to distinguish GC stars with field stars in PM space, as shown in Fig. 4. Moreover, some high velocity stars could in principle be caused by very tight (separations smaller than ~ 0.1 AU) GC binaries.

We made a very liberal cut to select GC stars in PM space, using a circle of radius 6 mas yr^{-1} (green circle in Fig. 4),

Table 2. Comparison of estimates on $\mu_{\alpha,*}$ and μ_δ with previous studies.

Method	$\mu_{\alpha,*}$ [mas yr $^{-1}$]	μ_δ [mas yr $^{-1}$]
Gaia Collaboration (2018a)	3.291 ± 0.0026	-17.591 ± 0.0025
Vasiliev (2019b)	3.285 ± 0.043	-17.621 ± 0.043
Baumgardt et al. (2019)	3.30 ± 0.01	-17.60 ± 0.01
This work (hybrid)	3.306 ± 0.013	-17.587 ± 0.024
Overall average	3.296 ± 0.019	-17.600 ± 0.024

corresponding to 15 times the PM dispersion measured by Baumgardt et al. (2019). The PM center of the GC was set at the mean of the values in Table 2. Our cut in PM space also corresponds to a velocity dispersion of 75 km s^{-1} , which is over 13 times the highest LOS velocity dispersion measured by Kamann et al. (2016). We chose such a liberal cut to ensure that we would not miss any high velocity GC members because otherwise our modeling would underestimate the mass. This left us with 9149 stars among the 9624 with low PM errors.

The filtering of field stars in PM space is not fully reliable because the cloud of field stars in PM space (upper part of Fig. 4) may extend into the (smaller) GC cloud (and past it). We thus proceed to another filtering in the color-magnitude diagram (CMD).

5.1.2. HST color-magnitude filtering

The left panel of Fig. 5 shows the locations in the CMD of the stars that survived the PM error cut and the filtering in PM space. While most stars follow a tight relation, a non-negligible fraction are outliers. We took a conservative cut on the stars on the CMD using kernel density estimation as displayed in the right panel of Fig. 5⁶. This graph displays the 1, 2 and 3 σ contours of the kernel density estimation, drawn as black lines. We selected stars inside the 2 σ region because the 3 σ region appears too wide to rule out binaries, while the 1 σ region was too conservative, extracting too few stars for our analysis.

The CMD filtering not only removes field stars whose PMs coincide by chance with those of GC stars, but also removes GC members that are unresolved binary stars and lie in the edges of the main-sequence, as well as particular Blue Stragglers in the $15 < F606W < 16.5$ range, which are believed to be the result of past mergers of GC members (Leonard 1989). Removing binaries and stars who have gone through mergers is important because their kinematics are dominated by two-body interactions, while our modeling assumes that stellar motions are dominated by the global gravitational potential of the GC. As discussed by Bianchini et al. (2016), three types of binaries need to be considered.

1. Resolved (i.e., wide) binaries will produce their own PMs that can be confused with the parallax. But, following Bianchini et al. (2016), the PMs of such resolved binaries, of order of a/T where a is the semi-major axis of the binary and T is the time baseline, are negligible in comparison to the GC velocity dispersion. Indeed, to be resolved at the distance of NGC 6397, the binary star should be separated by at least $0''.1$, which corresponds to 24 AU at the distance of NGC 6397; the most massive binaries, with mass below $2 M_\odot$, will have a period of 90 years, thus $a/T = 1.26 \text{ km s}^{-1}$,

⁶ We used `scipy.stats.gaussian_kde`, setting the bandwidth method to the Silverman's rule (Silverman 1986).

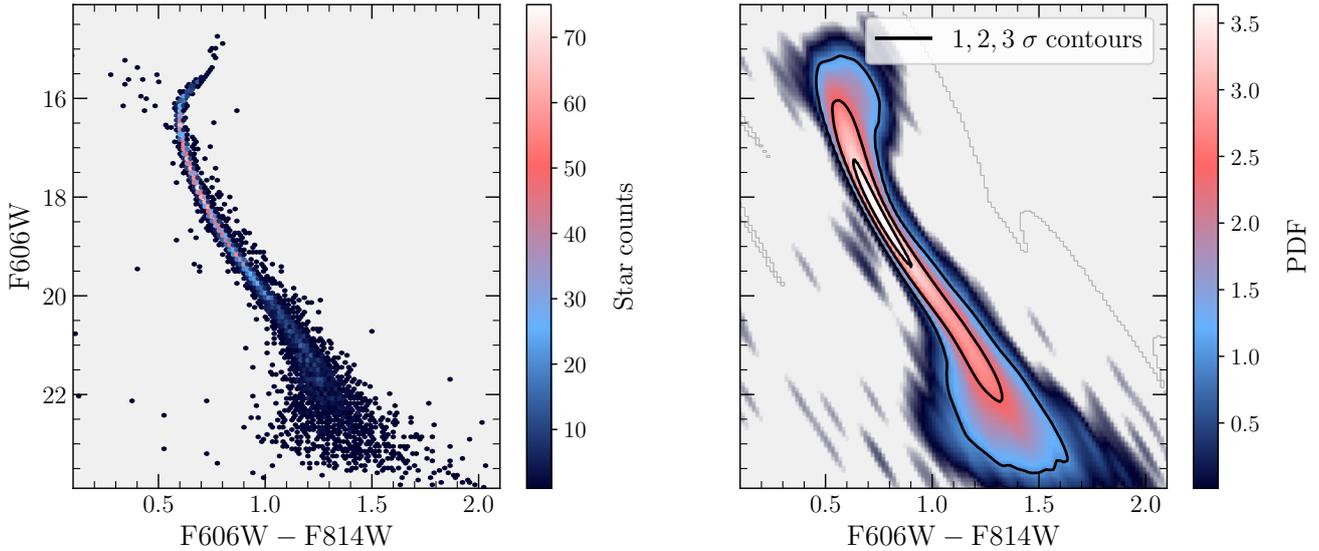


Fig. 5. HST color-magnitude diagrams. *Left:* CMD after the PM filtering process explained in Sect. 5.1.1. *Right:* Kernel density estimation of the HST isochrone displayed in the *left panel*. The 1, 2 and 3σ contours are displayed (from in to out).

while less massive binaries will have longer periods, hence lower a/T .

2. Nearly resolved binaries will not be deblended, and their distorted image (in comparison to the PSF) will lead to less precise astrometry, which will be flagged.
3. The orbits of unresolved binaries, considered as single entities, will be those of test particles in the gravitational potential. But their higher mass should lead them to have lower velocity dispersions than ordinary stars, in particular in the dense inner regions of GCs where the two-body relaxation time is sufficiently short for mass segregation.

Our CMD filtering left us with 7259 stars among the 9149 surviving the previous filters.

5.1.3. Removal of X-ray binaries

Bahramian et al. (2020) detected 194 X-ray sources within 200 arcsec from the center of NGC 6397, some of which could potentially be background AGN. The remaining sources are thought to be X-ray binaries, and as such will have motions perturbed by their invisible compact companion. We therefore deleted the 50 HST stars whose positions coincided within 1 arcsec with the “centroid” position of an X-ray source.

Bianchini et al. (2016) found that the effect of unresolved binaries on GC dispersion profiles is only important for high initial binary fractions (e.g., $\sim 50\%$), which can induce a difference of $0.1\text{--}0.3\text{ km s}^{-1}$ in the velocity dispersion profile, mainly in the GC inner regions, where the binary fraction should be highest. For lower initial binary fractions, Bianchini et al. (2016) found that unresolved binaries do not significantly affect the PM dispersion profile, but only the kinematics error budget. Therefore, the binary fractions calculated by Davis et al. (2008) and Milone et al. (2012a) for NGC 6397 are clearly insufficient (i.e., $\lesssim 5\%$ and $\lesssim 7\%$, respectively) to require a special treatment.

5.1.4. HST final numbers

After these cuts, we are left with 7209 stars from the HST observations (among the original 13 593). The combination of the CMD and PM filtering along with the removal of X-ray

binaries effectively cleans the data of most of the binaries that would affect our modeling, only leaving binaries of different luminosities that are at intermediate separations (≈ 0.15 to 5 AU, causing peculiar motions between 1 and 15 times the GC velocity dispersion). We checked that using a less liberal cut of 7.5σ instead of 15σ affects very little our results.

5.2. Gaia data cleaning

We followed similar steps in cleaning the *Gaia* data as we did for the HST data.

5.2.1. Quality flags

We filtered the *Gaia* stars with $\epsilon_\mu < 0.197\text{ mas yr}^{-1}$ using two data quality flags proposed by Lindegren et al. (2018). First, we only kept stars whose astrometric solution presented a sufficiently low goodness of fit:

$$\sqrt{\frac{\chi^2}{N-5}} < 1.2 \text{ Max} \{1, \exp[-0.2(G-19.5)]\}, \quad (16)$$

where N is the number of points (epochs) in the astrometric fit of a given star and 5 is the number of free parameters of the astrometric fit (2 for the position, 1 for the parallax and two for the PM). Equation (16) gives a sharper HR diagram, removing artifacts such as double stars, calibration problems, and astrometric effects from binaries. It is more optimized than the (`astrometric_excess_noise` < 1) criterion, used in Baumgardt et al. (2019) and Vasiliev (2019b), especially for brighter stars ($G \lesssim 15$), according to Lindegren et al. (2018).

Second, we only kept stars with good photometry.

$$1.0 + 0.015(G_{\text{BP}} - G_{\text{RP}})^2 < E < 1.3 + 0.06(G_{\text{BP}} - G_{\text{RP}})^2, \quad (17)$$

where $E = (I_{\text{BP}} + I_{\text{RP}})/I_G$ is the flux excess factor. Equation (17) performs an additional filter in the HR diagram, removing stars with considerable photometric errors in the BP and RP photometry, affecting mainly faint sources in crowded areas. This poor photometry broadens the CMD, leading to more confusion with field stars.

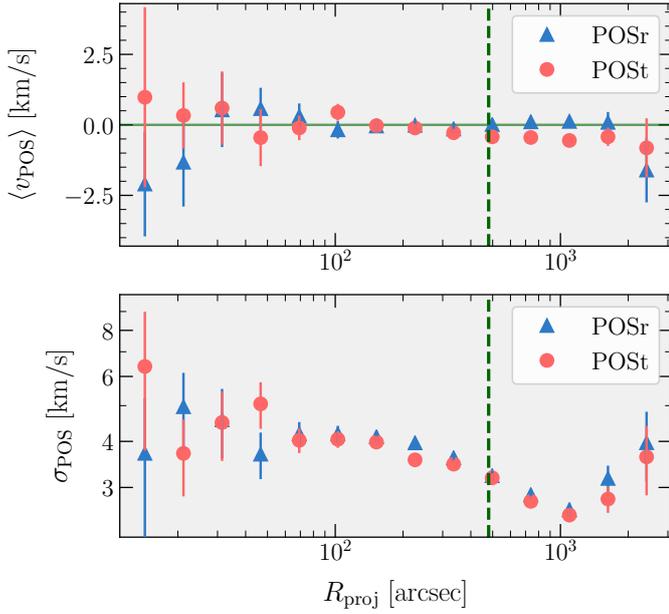


Fig. 6. Radial profiles of mean plane of sky velocity (*top*) and velocity dispersion (*bottom*) of NGC 6397 from cleaned *Gaia* DR2. The POS motions are split between radial (POSr, blue triangles) and tangential (POSt, red circles) components. The dashed green vertical line displays the 8' limit for use in MAMPOSS-PM.

Those variables correspond to the following quantities in the *Gaia* DR2 archive:

- χ^2 : `astrometric_chi2_al`
- ν' : `astrometric_n_good_obs_al`
- E : `phot_bp_rp_excess_factor`
- $G_{\text{BP}}-G_{\text{RP}}$: `bp_rp`

5.2.2. Maximum projected radius

NGC 6397 likely suffers from tidal heating every time it passes through the Milky Way's disk, in particular during its last passage less than 4 Myr ago (Sect. 2.2). The tidal forces felt by GC stars during passages through the disk will produce velocity impulses that are effective in perturbing the least bound orbits, which typically are those of the stars in the outer envelope of the GC. The *Gaia* data can trace the effects of such tidal disturbances on the GC kinematics.

We searched for anomalous kinematics in NGC 6397 using *Gaia* DR2 out to a maximum projected radius of 1° with respect to NGC 6397's center. The top panel of Fig. 6 indicates that the mean POS velocities in the radial and tangential directions differ beyond $8'7$. We also checked the concordance of the radial and tangential components of the radial profiles of POS velocity dispersion (or equivalently PM dispersion). The bottom panel of Fig. 6 shows excellent agreement between the two components of the velocity dispersion from $2'$ to $20'$, suggesting that the velocity ellipsoid is nearly isotropic in this range of projected radii. We conservatively adopted a maximum projected radius of $8'$ as set by the divergence of the mean velocity profiles beyond that radius.

5.2.3. Gaia proper motion filtering and the bulk proper motion of NGC 6397

As for HST, we filtered *Gaia* stars in PM space to later filter them in CMD space. Since *Gaia* data extends to much greater projected radii from the GC center, thus to lower GC surface densities, the GC stands out less prominently from the field stars (FS) in PM space. We therefore first estimated the bulk PM of the GC and we assigned a first-order probability of membership using a GC+FS mixture model.

We were tempted to assign two-dimensional (2D) Gaussian distributions for both GC stars and interlopers. However, the FS PM distribution has wider tails than a Gaussian. This means that stars on the other side of the GC, relative to the center of the field star component (i.e., its bulk motion) in PM space are more likely to be field stars than assumed by the Gaussian model. We found that the PM-modulus “surface density” profile (the velocity analog of the surface density profile) is well fit by a Pearson type VII distribution (Pearson 1916), as explained in detail in Appendix B.1. This distribution relies on two free parameters, a scale radius a and an outer slope γ , and can be written as:

$$f_{\mu}(\mu) = -\frac{\gamma + 2}{2\pi a^2} \left[1 + \left(\frac{\mu}{a}\right)^2 \right]^{\gamma/2}, \quad (18)$$

where $\mu = (\mu_{\alpha,*}, \mu_{\delta})$ and

$$\mu_i = \sqrt{(\mu_{\alpha,*i} - \overline{\mu_{\alpha,*i}})^2 + (\mu_{\delta,i} - \overline{\mu_{\delta,i}})^2}, \quad (19)$$

where the suffix i stands for the component analyzed, which in the case of Eq. (18) is the interlopers (i.e., field stars, hereafter FS). The reader can verify that, indeed, $\int f_{\mu}(\mu) d\mu = 1$, with $d\mu = 2\pi \mu d\mu$.

With the respective Gaussian and Pearson VII distributions of PMs of GC stars and FS, we performed a joint fit to the two-dimensional distribution of PMs, which provided us with a precise bulk PM of the GC. For this, we considered the convolved expressions of the PM distributions of both GC and interlopers with the errors provided by the *Gaia* archive ($\epsilon_{\mu_{\alpha,*}}, \epsilon_{\mu_{\delta}}$ and $\rho_{\mu_{\alpha,*}\mu_{\delta}}$). When passing onto polar coordinates, the uncertainty propagation of Eq. (19) produces

$$\begin{aligned} \epsilon_{\mu,i}^2 = & \left(\frac{\mu_{\alpha,*i} - \overline{\mu_{\alpha,*i}}}{\mu_i} \right)^2 \epsilon_{\mu_{\alpha,*i}}^2 + \left(\frac{\mu_{\delta,i} - \overline{\mu_{\delta,i}}}{\mu_i} \right)^2 \epsilon_{\mu_{\delta,i}}^2 \\ & + 2 \frac{(\mu_{\alpha,*i} - \overline{\mu_{\alpha,*i}})(\mu_{\delta,i} - \overline{\mu_{\delta,i}})}{\mu_i^2} \epsilon_{\mu_{\alpha,*}\mu_{\delta,i}}, \end{aligned} \quad (20)$$

where $\epsilon_{\mu_{\alpha,*}\mu_{\delta}} = \epsilon_{\mu_{\alpha,*}} \epsilon_{\mu_{\delta}} \rho_{\mu_{\alpha,*}\mu_{\delta}}$. The convolution with Gaussian errors was straightforward in the case of GC stars since their PM distribution was also modeled as a Gaussian, and thus we just added the errors to the dispersions in quadrature:

$$\sigma_{\text{GC,new}}^2 = \sigma_{\text{GC}}^2 + \epsilon_{\mu,\text{GC}}^2. \quad (21)$$

However, the convolution of the field star distribution with Gaussian errors cannot be reduced to an analytic function; numerical evaluation of the convolution integrals for each star would dramatically increase the calculation time. We therefore used the analytical approximation for the ratio of convolved to raw probability distribution functions of PM moduli, (which is also incorporated in MAMPOSS-PM), as briefly described in Appendix B.2 (details are given in Mamon & Vitral, in prep.). This allowed us to perform our mixture model fit to the PM

⁷ Drukier et al. (1998) had noticed a similar effect in the M 15 GC.

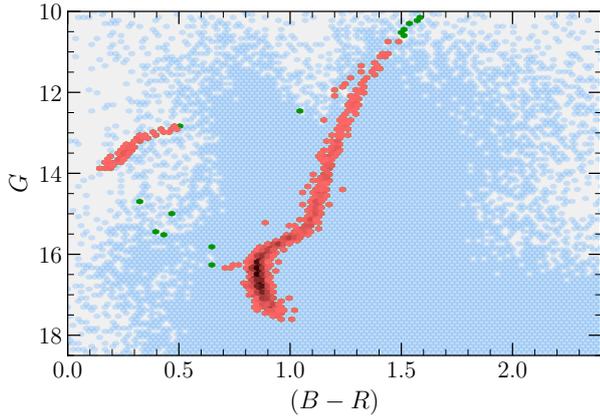


Fig. 7. Color magnitude diagram (CMD) of NGC 6397 *Gaia* DR2 data, after different filtering steps. The blue points indicate the stars that failed the PM error, astrometric and photometric flags, maximum projected radius and PM filters. The green points the stars that passed these 3 filters but are offset from the CMD. The red points show the *Gaia* sample of stars after the previous filters and subsequent CMD filtering, color-coded from red to dark red, according to increasing star counts.

data, using Markov chain Monte Carlo (MCMC)⁸ to estimate bulk motions of both the GC and the field stars and assign probabilities of GC membership for each star.

Table 2 displays our estimates of $\mu_{\alpha,*}$ and μ_{δ} , which show a good agreement with the literature values for the GCs mean PMs. In the same table, we display the overall average of the literature plus our estimates, with its respective uncertainty ϵ , calculated as $\epsilon^2 = \langle \epsilon_i \rangle^2 + \sigma^2$, where ϵ_i stands for the uncertainties on the estimated values, and σ stands for their standard deviation. This overall average and uncertainties were also later used as priors for the MAMPOSST-PM mass modeling analysis.

We only kept stars with GC membership probability higher than 0.9 according to our mixture model. This corresponds to a 1.43 mas yr^{-1} cut in the distribution of PMs of the GC, thus a 3.6σ cut.

5.2.4. *Gaia* color-magnitude filtering

We CMD-filtered the *Gaia* data roughly following the same KDE method as we used for the HST data. The only differences were that (1) we used the equivalent filters *G*, *B*, and *R*, and (2) we selected the 3σ region of the KDE, instead of 2σ , since the latter appeared to be too conservative a cut for *Gaia*. Figure 7 displays the final stars after this filter, along with the previously filtered subsets.

5.2.5. Removal of X-ray binaries

X-ray binaries are also in the *Gaia* sample. We therefore removed the five X-ray binaries that were matched to *Gaia* stars.

5.2.6. *Gaia* final numbers and comparison to HST

The cleaned *Gaia* sample contains 1905 stars. Figure 8 shows that the spatial coverage of *Gaia* DR2 is indeed poorer than HST when taking into account only stars from the clean samples. While the cleaned HST sample contains 221 stars within the $9''$ IMBH radius of influence (for an IMBH with a mass of

⁸ For all MCMC analyses except the one in MAMPOSST-PM, we used the Python package EMCEE (Foreman-Mackey et al. 2013).

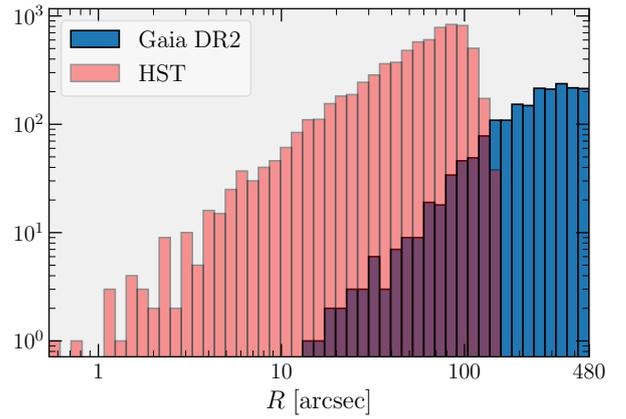


Fig. 8. Distribution of projected radii of stars from the GC center of Goldsbury et al. (2010), with cleaned HST data in red and cleaned *Gaia* DR2 data in blue. This plot indicates that HST is much better suited to probe a possible IMBH in the center.

$600 M_{\odot}$ as measured by Kamann et al. 2016, see Sect. 4.1), the cleaned *Gaia* presents none, which is clearly insufficient to provide reliable results on the presence of an IMBH.

5.3. MUSE data cleaning

We filtered the MUSE sample by removing stars whose probability of being a member of the GC, according to their position in velocity – metallicity space (Kamann et al. 2016, kindly provided by S. Kamann), was less than 0.9. This step left us with 6595 stars among the original 7130. We then removed stars with LOS velocity errors greater than 2.232 km s^{-1} (half of the GC velocity dispersion), as well as stars that did not match the HST stars in a symmetric ≤ 1 arcsec match. We were left with 532 stars, of which 4 were previously identified as X-ray binaries (see Sect. 5.1.3), yielding a final MUSE sample containing 528 stars, thus adding LOS velocities to $\sim 7\%$ of the HST sample.

Our Gaussian prior on v_{LOS} used the mean provided by Husser et al. (2016): $\langle v_{\text{LOS}} \rangle = 17.84 \text{ km s}^{-1}$. For the uncertainty on $\langle v_{\text{LOS}} \rangle$, we allowed a much wider dispersion of 2.5 km s^{-1} instead of the 0.07 km s^{-1} of Husser et al. (2016), given the relatively wide range of bulk LOS velocities reported in other studies (Milone et al. 2006; Lind et al. 2008; Carretta et al. 2009a; Lovisi et al. 2012).

5.4. Merging of the different datasets

Although we ran MAMPOSST-PM on HST and *Gaia* individually for data homogeneity, we preferred to combine the data from HST, MUSE and *Gaia* to probe the mass and velocity anisotropy parameters with more accuracy. We started with the HST filtered subset, which accounted for most of the data we would use, and then we merged it to the other datasets following the steps below:

1. We restricted the *Gaia* stars to those with *G* magnitudes within the limits of *F606W* magnitudes from the HST subset, (i.e., 16.11 and 22.14), given the quasi equivalence between these two filters. This step ensured that mass segregation effects would be the same for all data subsets.
2. We removed *Gaia* stars that were symmetrically matched to HST stars to better than 1 arcsec (see Sect. 4.4.1) since HST PMs presented smaller errors (Fig. 3).
3. We incorporated LOS velocities from MUSE, according to the approach described in Sect. 5.3.

After these steps, we were left with 8255 stars: 7209 of which had PMs from HST, with 583 of those presenting LOS velocities from MUSE, as well as 1046 additional stars with PMs from *Gaia* DR2.

6. Streaming motions in NGC 6397

6.1. Rotation

The presence of rotation in quasi-spherical systems makes the kinematical modeling more difficult. Any rotation will be interpreted by codes neglecting it, such as MAMPOSST-PM, as disordered motions, and should lead to different mass profiles and deduce velocity anisotropies that are more tangential than in reality.

Fortunately, NGC 6397 does not appear to have significant amount of rotation. Indeed, while Gebhardt et al. (1995) found that the integrated light of the core of NGC 6397 showed rotation with a projected amplitude of 2 km s^{-1} , Kamann et al. (2016) concluded with individual stars from VLT/MUSE observations (see Sect. 4.3 below) that rotation contributes negligibly to the second velocity moment in this GC. Vasiliev (2019c) reported a typical systematic uncertainty in *Gaia* DR2 PMs of up to $\sim 0.02 \text{ mas yr}^{-1}$ at any radius and considered rotation to be confirmed only when its peak amplitude exceeded $\sim 0.05 \text{ mas yr}^{-1}$, which was not the case for his NGC 6397 measurements. On the other hand, Bianchini et al. (2018) found POS rotation in NGC 6397 with *Gaia* DR2 at a 2σ level for this GC, but with $v/\sigma = 0.03$ only, the smallest value of the 51 GCs they analyzed. This is consistent with the quasi-null mean POS velocity profile seen in Fig. 6. Finally, Sollima et al. (2019) detected a rotation of 0.48 km s^{-1} from a 3D analysis, which is less than 10% of its velocity dispersion. We therefore neglect rotation in NGC 6397.

6.2. Radial motions

Similarly to rotation, any radial streaming motions will be interpreted by codes such as MAMPOSST-PM as radial dispersion. However, as seen in the top panel of Fig. 6, the *Gaia* DR2 PMs do not show strong signs of radial motions.

7. Practical considerations

We now discuss the practical implementation of MAMPOSST-PM to our cleaned sample of the stellar motions inside NGC 6397.

7.1. Surface density

7.1.1. Basic approach

Equation (4) requires the knowledge of the GC surface density (SD) profile, $\Sigma_{\text{GC}}(R)$. More importantly, Eq. (5) requires the knowledge of the 3D number density $\nu(r)$ when integrating the local VDF along the LOS. While MAMPOSST-PM has a mode where it jointly fits the parameters of $\nu(r)$, $M(r)$ and $\beta(r)$ to the distribution of stars in projected phase space (see Eq. (11) of Mamon et al. 2013, which is different from our Eq. (4)), this requires the data to have constant completeness as a function of projected radius. If the astrometric measures of PMs are not independent of projected radius, one first needs to estimate (and deproject) the SD profile based on a wider dataset than that with PM values to obtain priors on the parameters of $\nu(r)$. Such an analysis had been performed for mass-orbit modeling of

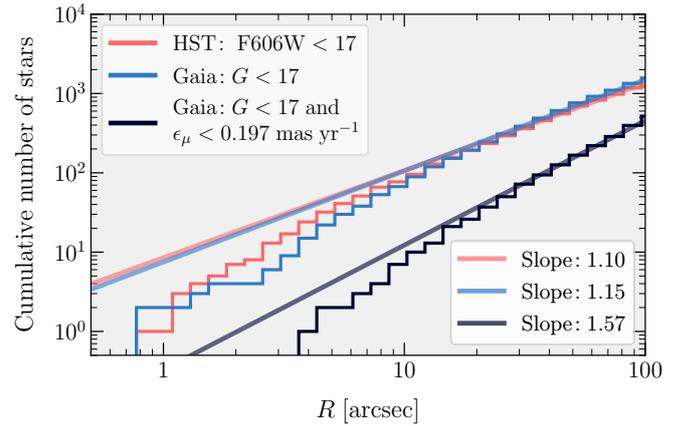


Fig. 9. Cumulative distribution functions of projected radii for $G < 17$ *Gaia* stars (blue), for the subset of $G < 17$ *Gaia* stars with precise PMs from Eq. (12) (black), and for the subset of HST stars (red). The straight lines are power laws to guide the eye.

galaxy clusters with spectroscopic information whose completeness depended on projected radius (e.g., Mamon et al. 2019).

The PM data are likely to be incomplete at small projected radii because of the increased crowding of stars as one moves inward, especially for *Gaia*, whose mirrors are smaller than that of HST. We noticed that *Gaia* data with PMs traced less well the inner cusp of the SD than the full *Gaia* data (requesting only positions and G magnitude). Figure 9 shows that the *Gaia* PM subsample with $G < 17$ is indeed incomplete in the inner regions relative to the full $G < 17$ *Gaia* sample since the distribution of projected radii of *Gaia* stars from the GC center is shallower for stars with precise PMs than for all stars (including those without PM measurements). A Kolmogorov-Smirnov test indicates that there is only 10^{-5} probability that the distributions of projected radii of the subsample with precise PMs and that of the full sample (including stars without PM measurements and those with imprecise ones) arise from the same parent population. Another difficulty is to measure the outer envelope given confusion with field stars. This is important because MAMPOSST-PM computes outward integrals.

We therefore needed to provide MAMPOSST-PM with a precomputed SD profile. In MAMPOSST-PM, this profile had to be a simple analytical one- or two-parameter model (where free parameters are scale and possibly shape). More precisely, MAMPOSST-PM assumes Gaussian priors on the precomputed SD profile parameters.

7.1.2. Choice of model and main parameters

The choice of a good model for SD is crucial because the mass of a possible IMBH is linked to the inner slope of the SD profile (van der Marel & Anderson 2010). NGC 6397 has long been known to have a cuspy (steep) inner SD profile (Auriere 1982; Lauzeral et al. 1992; Lugger et al. 1995; Noyola & Gebhardt 2006; Kamann et al. 2016), which prompted Djorgovski & King (1986) to classify it as core-collapsed. Unfortunately, no simple analytic model was ever fit to the SD profile of NGC 6397 extending to large projected radii⁹. Trager et al. (1995) estimated

⁹ Martinazzi et al. (2014) and Kamann et al. (2016) fitted the SD profile of NGC 6397 to large projected radii with Chebyshev polynomials (of $\log \Sigma$ vs. $\log R$) and multiple Gaussians, respectively. The former have no analytical deprojections, and while Gaussians are easily deprojected, the multiple Gaussians involve too many parameters for MAMPOSST-PM.

an effective (projected half-light) radius $R_e = 2'.90$ for NGC 6397 using ground-based data and analyzing the SD profile in different apertures. Using *Gaia* DR2 data, [de Boer et al. \(2019\)](#) recently inferred $R_e = 3'.35$ (which we infer from their 3D half-luminosity radius r_h , coupled with their assumed distance and with the ratio $R_e/r_h \simeq 0.732$ kindly estimated for us by M. Gieles). But their analysis was based on a fit to two dynamical models that both assumed a cored inner density profile, which clearly does not represent the inner SD profile of NGC 6397.

We thus had to estimate the SD profile ourselves. The cuspy profile led us to fit a Sérsic model ([Sérsic 1963](#); [Sérsic 1968](#)) to the distribution of projected radii since the Sérsic model has a shape parameter (Sérsic index) that makes it flexible to handle both cuspy and cored density profiles (e.g., Fig. A.1 of [Vitrail & Mamon 2020](#)). Furthermore, there are simple analytical forms for the precise (but not exact) deprojection of the Sérsic model. In Appendix A, we extend the analysis of [Vitrail & Mamon \(2020\)](#) of the precision of the different analytical approximations as a function of R_e and n , to even smaller values of R/R_e representative of our data. MAMPOSST-PM will then use the best-fit values of $\log R_e$ and n and their uncertainties as the respective mean and standard deviation of the Gaussian priors.

We required HST data to fit correctly the inner cusp, and we also needed *Gaia* data to fit correctly the outer envelope. We therefore fitted the GC surface density profile $\Sigma_{GC}(R)$ together with the field star surface density Σ_{FS} (assumed constant) from the *Gaia* DR2 photometric data, supplemented by the HST data to better probe the very inner regions. The two datasets are not cleaned in any way, except for cuts in magnitude and maximum (minimum) projected radius. The HST data may not be radially complete (we did not have a parent sample with all stars including those without PM measurements), but this is the best we can do. Our approach does not remove the Milky Way field stars, but instead incorporates them in our SD fit. It allows us to keep *Gaia* stars with no PM information, which would be filtered out in a PM cut, even though their contribution to the surface density shape is relevant.

7.1.3. Stitching HST and *Gaia*

We stitched the HST and *Gaia* datasets as follows:

1. We cut the *Gaia* stars (GC stars plus interlopers) at $G = 17$, where we expect negligible mass segregation. We explored different magnitude limits, but moderate changes affected little our fits. However, using very faint magnitude limits ($G = 19$ or 20.5), we noticed a lack of continuity in the transition from the HST to the *Gaia* radial range, which appears to be a consequence of overcrowding in the cluster's center ([Arenou et al. 2018](#)), where fainter magnitude stars are more difficultly detected.
2. We counted the number N_G of *Gaia* stars inside the annulus $2''.7 < R < 100''$, corresponding to the region where our HST data seems complete (see Fig. 8).
3. We removed the *Gaia* stars in that annulus and added the N_G HST brightest stars (a tiny fraction of which may be interlopers) in that annulus, so we could correctly probe the inner surface density profile.

7.1.4. Surface density results

We performed maximum likelihood estimation (MLE) to fit $\log R_e$ and n to the distribution of projected radii, using a model SD profile with a Sérsic GC plus a uniform (constant SD) field

star contribution, considering the data up to a maximum projected radius, R_{\max}^{10} . This maximum radius must be chosen large enough to capture the behavior of the SD profile in the outer envelope of the GC as well as the contribution of the field stars, but not too large to be overwhelmed by the field stars in the MLE fit.

We performed such MLE fits of the SD profile for 500 values of log-spaced R_{\max} between 0.2° and 1° . Figure 10 (left) shows a plateau for R_{\max} in the range $\sim 1300'' - 1800''$, where R_e , n and the field star SD are each indeed constant. At lower (resp. higher) R_{\max} the field star SD fluctuates at lower (higher) values with corresponding higher (lower) and fluctuating R_e and n .

For better precision, we then performed 30 MCMC runs for log-spaced R_{\max} between 1300 and 1800 arcsec and assigned the mean of best-fitted values and their uncertainties (green crosses and error bars of Fig. 10, respectively) as the global Sérsic radius and index means and uncertainties. The results are displayed in Fig. 10 (bottom). We found $R_e = 4'.51 \pm 0'.36$ and $n = 3.26 \pm 0.23$. Our effective radius is significantly higher than the estimates of [Trager et al. \(1995\)](#) and [de Boer et al. \(2019\)](#), which both fail to properly capture the cusp. Indeed, the Sérsic SD profile with the same index ($n = 3.26$) but with the [Trager et al. \(1995\)](#) effective radius (green curve in the right panel of Fig. 10) fails to reproduce the steep inner SD profile.

7.2. Multiple populations

Mass-orbit modeling codes such as MAMPOSST-PM assume a given population with a given SD profile, as well as given kinematics (i.e., $p(\mathbf{v}|R)$ for MAMPOSST-PM). Systems with multiple populations, each with their SD profiles and kinematics, should be analyzed jointly. MAMPOSST-PM can handle such multiple populations. In practice, we predetermine priors for the SD profile of each population, and then run MAMPOSST-PM with several tracer components.

There are two mechanisms that bring multiple populations in GCs. First, the formation of a GC may occur in several episodes, with different populations with their own chemistry and kinematics. Second, the short two-body relaxation time of GCs in general and especially of core-collapsed GCs such as NGC 6397, allows for energy exchanges between stars that are sufficient to drive them toward energy equipartition. This leads to mass segregation, where massive stars are confined to smaller radii and have lower typical velocities at a given radius. We examine these two mechanisms in turn.

7.2.1. Multiple chemical populations

Many GCs show signs of different chemical populations (e.g., [Carretta et al. 2009a](#)), in particular NGC 6397 ([Carretta et al. 2009b](#); [Milone et al. 2012b](#)). Unfortunately, the separation of stars into different chemical populations in NGC 6397 was only performed with HST data limited to small projected radii. Such a chemical separation of stars is impossible with the *Gaia* data, for lack of sufficient wavebands to provide a large enough dimensionality of colors to distinguish chemical populations. Moreover, this requires cleaning the data for differential reddening. Therefore, the splitting of the stars of NGC 6397 into chemical populations, while possible, is beyond the scope of the present article.

Besides, [Cordoni et al. \(2020\)](#), who analyzed the spatial distributions and kinematics of seven GCs (other than NGC 6397)

¹⁰ We performed MLE using `scipy.optimize.differential_evolution` in PYTHON.

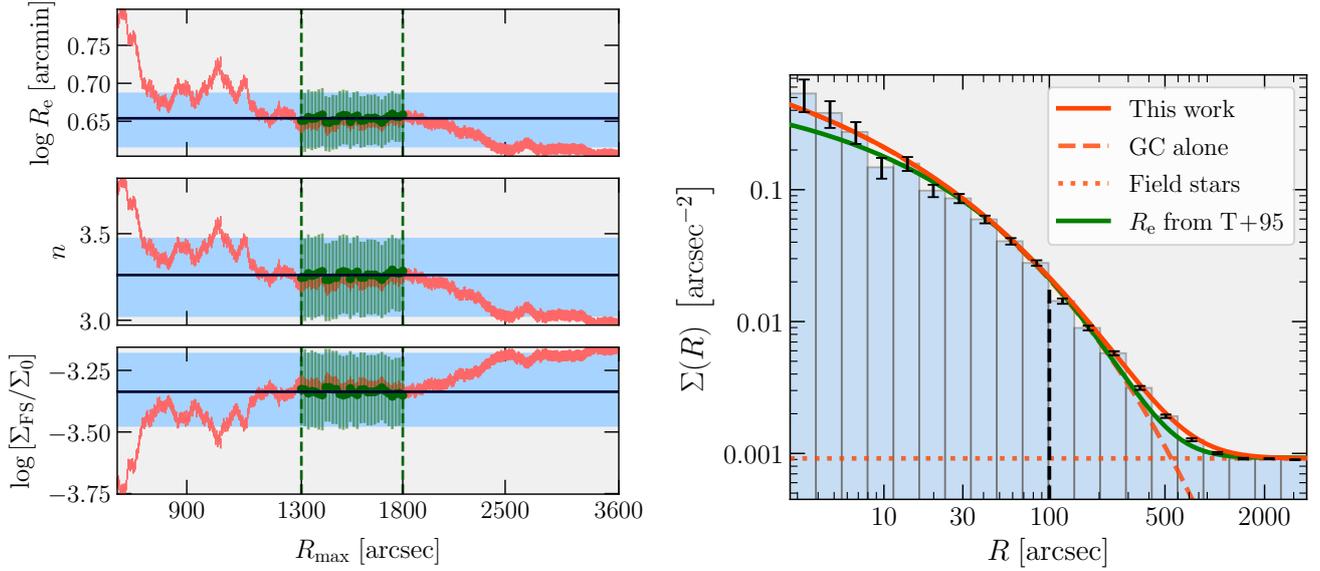


Fig. 10. Fits to the stellar surface density profile of NGC 6397. *Left:* Sérsic plus constant field star surface density fits, as a function of the maximum allowed radius, R_{\max} . The values of R_e , n , and Σ_{FS} obtained by MLE for 500 values of R_{\max} are shown in red. The black horizontal line and light blue shaded region respectively show the mean marginal values of 30 MCMC fits and their uncertainties, performed in the range of R_{\max} (delimited by the green vertical lines) where R_e , n and Σ_{FS} are roughly independent of R_{\max} . *Right:* goodness of fit of the surface density profile. The histogram shows the empirical profile, using 20 logarithmic radial bins extending from the innermost data point to 1° . The curves show different models: our MCMC fit (red) of a Sérsic model (dashed) plus constant field stars SD (dotted), as well as the total (solid) to compare with the data. We also show the Sérsic plus constant background prediction from MLE fits, but with the effective radius fixed to the value found by [Trager et al. \(1995\)](#) (green). The vertical black line highlights the separation between HST and *Gaia* data used for the fits.

split by their two detected chemical populations, found that only the two GCs with the highest two-body relaxation times showed signs of different kinematics between their two chemical populations. If the two populations have roughly the same mass functions, then two-body relaxation should wash out any differences in their positions in projected phase space ([Vesperini et al. 2013](#)), except in their outer regions where two-body relaxation is incomplete. The very low relaxation time of NGC 6397 (600 Myr at its half-mass radius according to [Baumgardt et al. 2019](#)) thus suggests that its chemical populations should have similar kinematics, at least within its effective radius of 4.5 as we found above.

7.2.2. Mass segregation

We explored the range of masses of the stars in our datasets by comparing synthetic CMDs to our observed ones. Following [Carballo-Bello et al. \(2012\)](#), we generated a CMD with a set of isochrones from the PARSEC ([Marigo et al. 2017](#); [Bressan et al. 2012](#); [Pastorelli et al. 2019](#) and references therein) software. We provided $n_{\text{inTPC}} = 15$ since [Marigo et al. \(2017\)](#) mention this suffices to recover the main details of thermal pulse cycle variations in the evolutionary tracks. We also assigned $\eta_{\text{Reimers}} = 0.477$, which is the median value of the Reimers coefficient ([Reimers 1975](#)) among clusters, provided by [McDonald & Zijlstra \(2015\)](#); in their Fig. 4, they showed that NGC 6397 presents a value of η_{Reimers} corresponding to this value within $\sim 1\sigma$. We adopted an age of 12.87 Gyr and a metallicity of $[\text{Fe}/\text{H}] = -1.54$ from [Marín-Franch et al. \(2009\)](#), while adopting our distance of 2.39 kpc (Sect. 3.3) and a Galactic extinction $A_V = 0.50$ obtained from $E(B-V) = 0.160$ ([Schlafly & Finkbeiner 2011](#))¹¹ and

¹¹ We used the Galactic Dust Reddening and Extinction tool, <https://irsa.ipac.caltech.edu/applications/DUST/> in the center of NGC 6397.

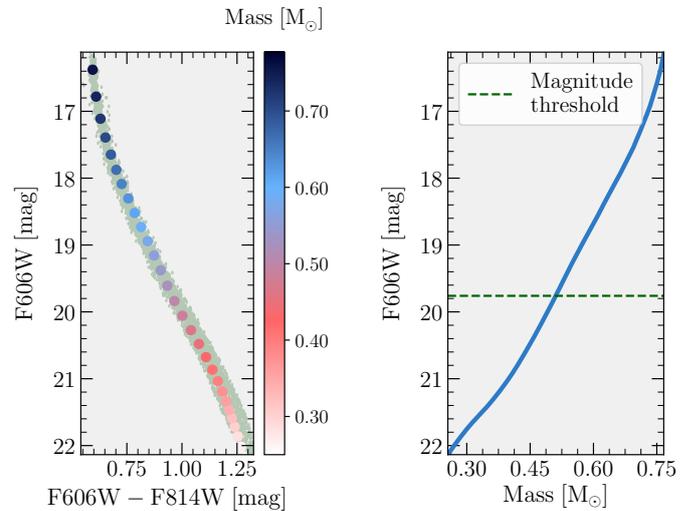


Fig. 11. Stellar masses. *Left:* comparison of the cleaned HST data and the fitted PARSEC isochrone, displaying the expected stellar mass relative to each main-sequence position. *Right:* magnitude – mass relation from the PARSEC model. The green dashed horizontal line shows the magnitude threshold that we used to split our sample into two mass populations.

$R_V = 3.1$. We note that this value of extinction is 14% lower than $E(B-V) = 0.188$ found in the older reddening maps of [Schlegel et al. \(1998\)](#), but consistent with $A_V = 0.52 \pm 0.02$ found by [Valcin et al. \(2020\)](#) despite their larger distance of 2.67 pc (Table 1).

The left panel of Fig. 11 shows that the PARSEC model fits well the CMD of the cleaned subsample of HST stars. The mass ratio among the heaviest and the lightest stars in our cleaned

HST sample is 2.57, sufficiently high to worry about the possibility of mass segregation in our cleaned HST sample.

In fact, mass segregation is visible in NGC 6397 (Heyl et al. 2012; Martinazzi et al. 2014). Analyzing HST stars in the range $3' < R < 7'.5$, Heyl et al. (2012) found mass segregation of main-sequence stars: brighter stars show 4% lower median projected radius and 12% lower median PM moduli than fainter stars. In a subsequent study by the same team, Goldsbury et al. (2013) showed that two characteristic radii scale as a M^γ , with $\gamma = -1.0 \pm 0.1$ for NGC 6397. Martinazzi et al. (2014) found that the mean mass of main-sequence stars drops with *physical* radius (after deprojection) from $>0.7 M_\odot$ for $r < 10''$ to $\approx 0.56 M_\odot$ for $r > 100''$ (after correcting for the completeness with magnitude, estimating it by adding artificial PSF-convolved stars to the images), thus a 20% effect.

MAMPOSST-PM is able to treat multiple stellar populations together. We therefore also performed SD fits with two populations of stars, as we explain below.

The magnitude threshold we chose to separate the two population of stars was based on the analysis of Heyl et al. (2012): Their Figs. 7 and 12 indicate that NGC 6397 main-sequence stars present a (small) variation of radial distribution and velocity dispersion profiles, respectively, at a magnitude of $F814W \sim 18.75$, which corresponds to $F606W = 19.76$. We thus used this limit to divide our cleaned subset of 8255 stars into two populations, one with brighter magnitudes (6527 heavier stars) and the other with fainter ones (1728 less massive stars).

Unfortunately, we could not perform a surface density fit as robust as for the single population case because (1) our data was considerably incomplete at higher magnitudes (the fainter subset) and (2) as mentioned in Sect. 7.1.3, the HST plus *Gaia* stacked subset presented discontinuous trends when allowing stars with fainter magnitudes ($G > 17$). We therefore let MAMPOSST-PM fit the SD profile of each population from the kinematics only, more precisely from the conditional probabilities $p(v|R)$. We adopted Gaussian priors for $\log R_e$, with mean equal to that found by our previous SD fit of the single population (Sect. 7.1) and a wide (0.2 dex) uncertainty. We also considered Gaussian priors for the Sérsic indices, but with lower uncertainties to avoid a degeneracy in its marginal distribution.

We helped MAMPOSST-PM by providing narrow Gaussian mass priors for each population. We derived mass fractions for each population using the power-law main-sequence stellar mass function of slope $\alpha = -0.52$ ¹² together with the mass limits of our subsets. Indeed, a power-law relation $dN/dm \propto m^\alpha$ implies a total stellar mass in the range of stellar masses (m_1, m_2) of

$$M_{\text{total}} \propto \int_{m_1}^{m_2} m \frac{dN}{dm} dm = \frac{m_2^{2+\alpha} - m_1^{2+\alpha}}{2 + \alpha}, \quad (22)$$

and thus derive

$$\frac{M_{\text{bright}}}{M_{\text{faint}}} = \frac{m_{\text{bright}}^{2+\alpha} - m_{\text{cut}}^{2+\alpha}}{m_{\text{cut}}^{2+\alpha} - m_{\text{faint}}^{2+\alpha}}, \quad (23)$$

where M_{bright} and M_{faint} are the masses of the brighter and fainter populations, respectively and m_{bright} , m_{faint} and m_{cut} are the respective highest, lowest and two-population threshold masses of the global subset. With $m_{\text{bright}} = 0.77 M_\odot$, $m_{\text{cut}} = 0.51 M_\odot$,

¹² The slope $\alpha = -0.52$ of the main-sequence stellar mass function of NGC 6397 is given in H. Baumgardt's very useful web site on GCs, <https://people.smp.uq.edu.au/HolgerBaumgardt/globular>

and $m_{\text{faint}} = 0.25 M_\odot$ (right panel of Fig. 11), Eq. (23) yields a bright mass fraction of 0.56. Since our main mass estimates of NGC 6397's mass with one single population were centered around $10^5 M_\odot$, we passed logarithmic Gaussian priors to each population, centered at $\log(M/M_\odot) = 4.7$, with standard deviation of 0.05.

7.2.3. MAMPOSST-PM assumptions and priors

To summarize, MAMPOSST-PM assumes that NGC 6397, lying at a distance of 2.39 kpc, with its center at the position found by Goldsbury et al. (2010), is a spherical self-gravitating system with no internal streaming motions (rotation, expansion/collapse, etc.). It also assumes that the *local* velocity distribution function is an ellipsoidal 3D Gaussian aligned with the spherical coordinates.

By relying on the Jeans equation, MAMPOSST-PM assumes that stars are test particles orbiting the gravitational potential, and that they do not interact with one another. While binary stars have their own motions and violate the test particle hypothesis, these are in great part filtered out of our data. However, the low two-body relaxation time of NGC 6397 means that star-star interactions are not negligible, which violates the test particle assumption.

We now summarize our adopted priors. The surface density profile of GC stars is assumed to follow the Sérsic law. We adopted Gaussian priors for $\log R_e$ and for n (see Sect. 7.1) for both single- and two-population analyses.

Following Sect. 5.2.3, we adopted Gaussian priors on the bulk LOS velocity and PM of the GC and Pearson VII priors on the modulus of the bulk PM of the field stars. The field star LOS distribution is Gaussian with a wide (1.7 dex) uncertainty. The log ratio of field stars to GC stars is flat between -5 and -2 .

Our standard runs assume isotropic velocities. We also ran cases with anisotropic velocities, and following Mamon et al. (2019) we used flat priors on $\beta_{\text{sym}} = \beta/(1 - \beta/2)$, which varies from -2 for circular orbits to 0 for isotropic orbits and $+2$ for radial orbits, and with $\beta_{\text{sym}} \rightarrow \beta$ for $|\beta| \ll 1$. Our flat priors were limited to $-1.9 < \beta_{\text{sym}} < 1.9$, where the two extremes correspond to $\beta = -38$ and $\beta = 0.97$, respectively (see Eq. (8)).

We also adopted flat priors on the GC stellar log mass (3 to 6) for the single-population runs and Gaussian priors with equal means and wide dispersions for the two-population runs (Sect. 7.2.2). The (sub-)cluster of unseen objects is modeled with different density models with a wide Gaussian (1 dex) prior on the scale radius uncertainty, assuming isotropic velocities.

8. Analysis

8.1. Marginal distributions and covariances

We explored parameter space to determine marginal distributions and parameter covariances using the same Markov chain Monte Carlo (MCMC) approach as Mamon et al. (2013, 2019), see Sect. 4.7 of Mamon et al. (2019). This makes use of the public Fortran COSMOMC code (Lewis & Bridle 2002)¹³. In particular, we used 6 MCMC chains run in parallel and stopped the exploration of parameter space after one of the chains reached a number of steps $N_{\text{steps}} = 10\,000 N_{\text{free}}$, where N_{free} is the number of free parameters of the model. The only difference is that

¹³ <https://cosmologist.info/cosmomc/>

we now discard the 3000 N_{free} first steps of each MCMC chain (burn-in phase) instead of the first 2000 N_{free} as Mamon et al. (2019) previously did.

8.2. Bayesian information

There are several ways to compare the different results of MAMPOSS-T-PM using different priors. The simplest would be to compare the log-likelihoods. But there is then the risk of over-fitting (under-fitting) the data when using too many (few) free parameters. We considered two model selection criteria to distinguish our different models and priors. We first considered the corrected Akaike Information Criterion (Sugiura 1978)

$$\text{AICc} = \text{AIC} + 2 \frac{N_{\text{free}} (1 + N_{\text{free}})}{N_{\text{data}} - N_{\text{free}} - 1}, \quad (24)$$

where AIC is the original Akaike Information Criterion (Akaike 1973)

$$\text{AIC} = -2 \ln \mathcal{L}_{\text{MLE}} + 2 N_{\text{free}}, \quad (25)$$

and where \mathcal{L}_{MLE} is highest likelihood found when exploring the parameter space, N_{free} is the number of free parameters, and N_{data} the number of data points¹⁴. We also considered the Bayes Information Criterion (BIC, Schwarz 1978):

$$\text{BIC} = -2 \ln \mathcal{L}_{\text{MLE}} + N_{\text{free}} \ln N_{\text{data}}. \quad (26)$$

Given that the relative likelihood (given the data) of one model relative to a reference one is

$$\exp\left(-\frac{\text{AIC} - \text{AIC}_{\text{ref}}}{2}\right) \quad (27)$$

(Akaike 1983), a model with a higher AICc (or AIC) relative to a reference model can be ruled out with 95% confidence if $\text{AICc} > \text{AICc}_{\text{ref}} + 6$. BIC follows the analogous to Eq. (27) for the relative likelihood (Kass & Rafferty 1995):

$$\exp\left(-\frac{\text{BIC} - \text{BIC}_{\text{ref}}}{2}\right), \quad (28)$$

so a model with a higher BIC relative to a reference model can be ruled out with 95% confidence if $\text{BIC} > \text{BIC}_{\text{ref}} + 6$.

Since BIC penalizes extra parameters more (factor $\ln N_{\text{data}} \approx 9.0$) than AICc (i.e., by a factor two for our dataset), BIC effectively prefers simpler models than does AICc. BIC is well suited for situations where the true model is among the tested ones, while AIC(c) is more robust in the opposite case (Burnham & Anderson 2002). In practice, we do not expect any of our models to be true: for example, there is no expectation that the surface density of stars in a GC should precisely follow a Sérsic model. To paraphrase the statistician George Box: “All models are wrong, but some are useful” (Box 1979)¹⁵.

We therefore present both AICc and BIC in our results. However, we give preference to AICc whenever having to decide between different models.

9. Mass-orbit modeling results

9.1. MAMPOSS-T-PM runs

Table 3 displays our standard runs, along with some robustness tests on velocity isotropy, mass segregation and presence of central components. Our runs have come with 4 variations for the inner mass distribution: (1) no IMBH nor (sub)-cluster of unresolved objects (CUO, see Sect. 2.2), (2) IMBH with no CUO, (3) CUO without an IMBH, and (4) both IMBH and CUO.

The left panel of Fig. 12 shows the marginal distributions and covariances from the MCMC output of model 1 (IMBH without CUO and with isotropic stars), for a selection of missing parameters (missing are the bulk GC motion and the log ratio of field stars to GC members). The MCMC ran well in the sense that the MLE values (red arrows) are close to the peaks of the marginal distributions (blue shaded histograms). One notices that higher IMBH masses are obtained with higher GC scale radii, independently of the Sérsic index. The reason is that, at given total stellar GC mass, more concentrated stars will lead to more mass in the inner regions, meaning less IMBH mass. This is the opposite of what van der Marel & Anderson (2010) proposed, but consistent with what Baumgardt (2017) found for NGC 6397.

The right panel of Fig. 12 shows the same for model 10 (CUO without IMBH and with isotropic stars). Again, the MLE values are coincident with the peaks of the marginal distributions. But surprisingly, the MCMC favors a GC scale radius at the upper edge of the 3σ prior obtained from the SD profile fit of Sect. 7.1.4. We will discuss this below. The mass of the CUO increases with the GC and CUO scale radii, and decreases with the GC Sérsic index (because low index leads to flatter SD profiles, which are somewhat analog to wider SD profiles). Interestingly, the scale radius of the CUO population is small ($7''$), as we will discuss in Sect. 9.4.1.

9.2. Velocity anisotropy

The combination of HST data properly probing the inner regions of NGC 6397 with the *Gaia* data probing the outer regions allows us to estimate the velocity anisotropy across the cluster. We ran MAMPOSS-T-PM using different priors on the anisotropy. Our standard prior has isotropic velocities throughout the GC. Our other priors assume the gOM anisotropy model (we found that the softer varying Turet et al. 2007 model performs almost as well, but not better).

The highest likelihoods (lowest $-\Delta \ln \mathcal{L}_{\text{max}}$) between the first five models are achieved by model 2 (everything free), but it did not converge ($R^{-1} = 0.07 > 0.02$). In particular, model 2 leads to quasi-isotropic inner and outer velocities, with a narrow constraint for β_0 and a quite wide one for β_∞ . The transition radius, r_β , is very poorly constrained. Compared to model 2, the isotropic model 1 is very strongly preferred by BIC (according to Eq. (28) for $\Delta \text{BIC} = 26$) and marginally so (92% confidence according to Eq. (27) for $\Delta \text{AIC} = 4.87$) by AICc. Model 3 with the anisotropy radius tied to the effective radius of the SD profile (TAND) is strongly preferred by BIC Bayesian evidence (with respect to model 2), but not with AICc. Interestingly, model 3 leads to negligible anisotropy in the center with narrow uncertainty. But it points to mildly tangential anisotropy at $r = r_{\text{out}} = 8'$, which corresponds to our maximum projected radius R_{max} . Forcing central isotropy with TAND (model 4) leads to less tangential outer velocities. Conversely, forcing outer isotropy with TAND (model 5) leads to strongly well constrained isotropic inner velocities.

¹⁴ The main results of this paper use $N_{\text{data}} = 8255$ and $N_{\text{free}} \sim 10$.

¹⁵ The angular fluctuation spectrum of the cosmic microwave background appears to be a counter-example.

Table 3. Main results and priors for MAMPOSST-PM runs, using the merged sample from the cleaned HST, *Gaia* and MUSE datasets.

Model	r_β [pc]	$\beta(r)$	Σ	N_{free}	R^{-1}	β_0	β_{out}	r_{scale} [arcmin]	n	M_{BH} [M_\odot]	M_{CUO} [$10^3 M_\odot$]	M_{GC} [$10^4 M_\odot$]	$-\Delta \ln \mathcal{L}_{\text{max}}$	ΔAICc	ΔBIC
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1	–	iso	S	8	0.003	0	0	$4.38^{+0.41}_{-0.38}$	$3.29^{+0.21}_{-0.06}$	658^{+70}_{-338}	–	$9.75^{+0.66}_{-0.70}$	18.82	29.62	8.90
2	$1.06^{+52.66}_{-0.99}$	gOM	S	11	0.070	$0.06^{+0.04}_{-0.13}$	$-0.05^{+0.66}_{-0.50}$	$4.15^{+0.68}_{-0.26}$	$3.28^{+0.16}_{-0.06}$	506^{+229}_{-200}	–	$9.50^{+0.98}_{-0.43}$	18.25	34.49	34.81
3	TAND	gOM	S	10	0.007	$0.05^{+0.03}_{-0.06}$	$-0.17^{+0.17}_{-0.08}$	$4.24^{+0.51}_{-0.41}$	$3.29^{+0.18}_{-0.06}$	453^{+221}_{-207}	–	$9.76^{+0.68}_{-0.72}$	18.45	32.89	26.20
4	TAND	gOM	S	9	0.005	0	$-0.08^{+0.10}_{-0.12}$	$4.21^{+0.58}_{-0.34}$	$3.27^{+0.22}_{-0.04}$	530^{+190}_{-224}	–	$9.61^{+0.85}_{-0.56}$	18.50	30.98	17.28
5	TAND	gOM	S	9	0.003	$-0.01^{+0.06}_{-0.02}$	0	$4.24^{+0.56}_{-0.34}$	$3.30^{+0.19}_{-0.08}$	576^{+143}_{-274}	–	$9.57^{+0.86}_{-0.52}$	18.85	31.68	17.98
6	–	iso	S	11	0.017	0	0	$3.27^{+0.39}_{-0.42}$	$3.31^{+0.08}_{-0.09}$	511^{+158}_{-207}	–	$5.45^{+0.74}_{-0.95}$	7.09	12.17	12.49
7	–	iso	S	–	–	0	0	$8.35^{+0.94}_{-2.07}$	$3.23^{+0.09}_{-0.07}$	–	–	$5.61^{+0.34}_{-0.87}$	–	–	–
8	–	iso	S	7	0.002	0	0	$4.02^{+0.36}_{-0.34}$	$3.79^{+0.11}_{-0.24}$	–	–	$9.18^{+0.58}_{-0.44}$	25.98	41.93	14.20
9	–	iso	S	10	0.009	0	0	$2.46^{+0.34}_{-0.14}$	$3.33^{+0.19}_{-0.06}$	–	–	$4.32^{+0.83}_{-0.41}$	13.68	23.35	16.66
9	–	iso	S	10	0.007	0	0	$6.46^{+1.61}_{-0.77}$	$3.28^{+0.07}_{-0.10}$	–	–	$5.39^{+0.51}_{-0.64}$	–	–	–
9	–	iso	S	10	0.007	0	0	$5.58^{+0.08}_{-0.79}$	$3.27^{+0.05}_{-0.09}$	42^{+92}_{-26}	–	$10.94^{+0.25}_{-0.93}$	10.09	16.17	9.48
9	–	iso	P	–	–	0	0	$0.13^{+0.03}_{-0.05}$	–	–	$2.08^{+0.08}_{-0.90}$	–	–	–	–
10	–	iso	S	9	0.002	0	0	$5.31^{+0.36}_{-0.52}$	$3.27^{+0.05}_{-0.09}$	–	–	$10.68^{+0.52}_{-0.68}$	9.86	13.70	0.00
10	–	iso	P	–	–	0	0	$0.12^{+0.02}_{-0.04}$	–	–	$2.01^{+0.16}_{-0.76}$	–	–	–	–
11	–	iso	S	9	0.004	0	0	$5.28^{+0.38}_{-0.55}$	$3.25^{+0.08}_{-0.06}$	–	–	$10.41^{+0.56}_{-0.64}$	11.14	16.26	2.56
11	–	iso	H	–	–	0	0	$0.08^{+0.02}_{-0.04}$	–	–	$3.09^{+0.26}_{-1.57}$	–	–	–	–
12	–	iso	S	10	0.012	0	0	$5.30^{+0.36}_{-0.51}$	$3.24^{+0.08}_{-0.06}$	–	–	$10.62^{+0.54}_{-0.66}$	9.69	15.37	8.68
12	–	iso	S	–	–	0	0	$0.16^{+0.05}_{-0.06}$	$0.92^{+1.68}_{-0.11}$	–	$1.92^{+0.35}_{-0.69}$	–	–	–	–
13	TAND	gOM	S	11	0.031	$-0.02^{+0.06}_{-0.04}$	$-0.03^{+0.11}_{-0.14}$	$5.31^{+0.36}_{-0.52}$	$3.27^{+0.05}_{-0.08}$	–	–	$10.69^{+0.53}_{-0.67}$	9.92	17.83	18.15
13	–	gOM	P	–	–	0	0	$0.12^{+0.03}_{-0.04}$	–	–	$1.91^{+0.30}_{-0.67}$	–	–	–	–
14	–	iso	S	12	0.014	0	0	$4.42^{+0.85}_{-0.85}$	$3.27^{+0.09}_{-0.07}$	–	–	$6.69^{+0.08}_{-1.43}$	0.00	0.00	7.33
14	–	iso	S	–	–	0	0	$9.38^{+1.02}_{-2.07}$	$3.25^{+0.05}_{-0.11}$	–	–	$4.96^{+1.12}_{-0.10}$	–	–	–
14	–	iso	P	–	–	0	0	$0.11^{+0.02}_{-0.04}$	–	–	$1.72^{+0.28}_{-0.67}$	–	–	–	–
15	TAND	gOM	S	16	0.009	$0.03^{+0.06}_{-0.05}$	$0.00^{+0.08}_{-0.15}$	$3.99^{+0.57}_{-0.56}$	$3.27^{+0.10}_{-0.07}$	–	–	$6.16^{+0.63}_{-0.89}$	–0.56	6.91	42.29
15	TAND	gOM	S	–	–	$-0.11^{+0.10}_{-0.08}$	$0.13^{+0.15}_{-0.68}$	$9.85^{+1.00}_{-2.47}$	$3.24^{+0.07}_{-0.10}$	–	–	$5.42^{+0.66}_{-0.55}$	–	–	–
15	–	iso	P	–	–	0	0	$0.11^{+0.03}_{-0.04}$	–	–	$1.84^{+0.19}_{-0.90}$	–	–	–	–

Notes. Columns are (1): run number; (2): anisotropy radius (where TAND is tied anisotropy number density, $r_\beta = r_{\text{scale}}$); (3): velocity anisotropy model (“iso” for isotropic); (4): surface density model (“S” for Sérsic, “P” for Plummer, “H” for Hernquist); (5): number of free parameters; (6): MCMC convergence criterion ($R^{-1} \leq 0.02$ is considered as properly converged, worse convergence runs are shown in bold red); (7): anisotropy value at $r = 0$; (8): anisotropy value at $r_{\text{out}} = 8'$; (9): scale radius (effective – half projected number – radius R_c for Sérsic, otherwise radius of 3D density slope -2 , with $r_{-2} = 1.63 R_c$ for Plummer and $r_{-2} = 0.28 R_c$ for Hernquist); (10): Sérsic index; (11): black hole mass; (12): mass of inner subcluster of unresolved objects (CUO); (13): mass of the stellar population considered; (14): difference in minus natural logarithm of the maximum likelihood relative to model 14; (15): difference in AIC (Eq. (25)) relative to best value; (16): difference in BIC (Eq. (26)) relative to best value. Blue bold zeros for $-\Delta \ln \mathcal{L}_{\text{max}}$, ΔAICc and ΔBIC represent the reference values among all runs, for each column, respectively. Values in bold gray were fixed parameters. The values of Cols. (9)–(13) are at maximum likelihood (black) or medians (orange and italics, when the MLE is outside the 16–84 percentiles). The uncertainties are from the 16th and 84th percentiles of the marginal distributions. The number of stars in each subset is 8255, the maximum allowed PM error (Eq. (12)) for those runs is $0.197 \text{ mas yr}^{-1}$, the distance was considered as 2.39 kpc , and the GC center was that of Goldsbury et al. (2010). For models considering two main-sequence populations, the parameters of the brighter (fainter) population is displayed first (next).

We also ran a CUO case, where the stars had free inner and outer anisotropy, but with TAND. Compared to its isotropic analog (model 10), model 13 with a single population has a slightly worse likelihood (but did not fully converge), and the inner and outer anisotropies are consistent with isotropic orbits ($\beta_0 = -0.02 \pm 0.05$, $\beta_{\text{out}} = -0.03 \pm 0.05$). An analogous run for the two-population CUO model also yields close to isotropic orbits for the bright population ($\beta_0 = 0.03 \pm 0.05$, $\beta_{\text{out}} = 0.00 \pm 0.12$). The faint population has inner velocities consistent with isotropy ($\beta_0 = -0.1 \pm 0.1$). On the other hand, its outer anisotropy is poorly constrained ($\beta_{\text{out}} = 0.1^{+0.2}_{-0.7}$) because of the near complete lack of faint *Gaia* stars with decent PM errors, hence we are limited to the inner regions. Still, the outer velocities are again consistent with isotropy.

In summary, there is strong evidence for isotropy with BIC, but weaker evidence with AICc. However, even the anisotropic runs produce anisotropy profiles that are very close to isotropic throughout. We thus conclude that the visible stars in NGC 6397

have quasi-isotropic orbits, at least for the stars brighter than $F606W = 19.76$ and up to $R = 8'$. We therefore adopted isotropic orbits as our standard when investigating other quantities.

9.3. Intermediate mass black hole

We tested the scenario with an IMBH and no CUO (models 1 to 6). Among models 1 to 6, the most likely one and very strongly favored by AICc (model 6, which is the analog of model 1, but with 2 populations for the GC stars) yields an IMBH mass of $511^{+158}_{-207} M_\odot$, while model 1, which is weakly favored by BIC over model 6, yields an IMBH mass of $658^{+70}_{-338} M_\odot$. Both models 1 and 6 indicate an IMBH mass above $200 M_\odot$ at 95% confidence. Furthermore, AICc (resp. BIC) indicates very strong (resp. quite strong) evidence for the presence of an IMBH in the absence of a central diffuse component (i.e., comparing IMBH models 1 and 6 to models 7 and 8, respectively).

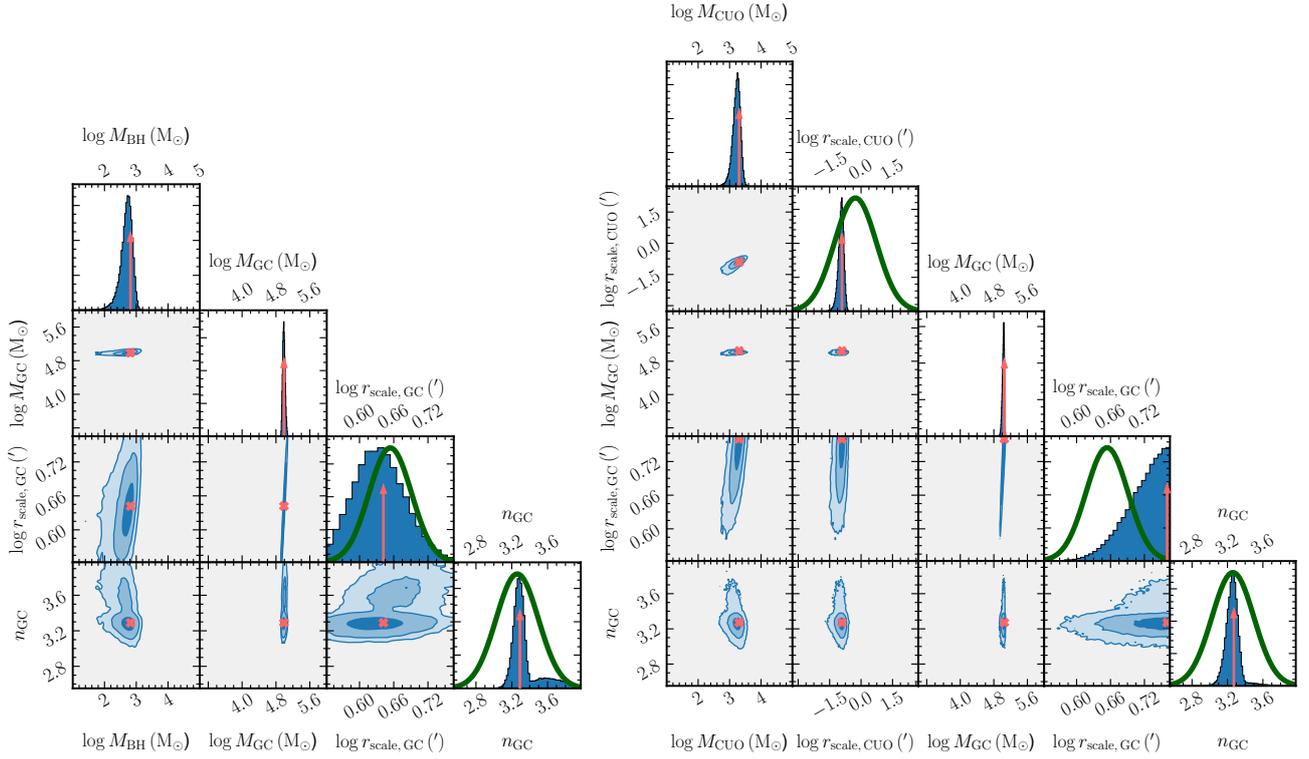


Fig. 12. Selected marginal distributions and covariances for models 1 (IMBH, *left*) and 10 (subcluster of unresolved objects, CUO, *right*), both for single main-sequence population and isotropic velocities. Only the elements of the MCMC chains past the burn-in phase are considered. The green curve shows the Gaussian prior, while its absence indicates flat priors in the provided range. The red arrows and crosses indicate maximum likelihoods. The contours are the equivalent of 1, 2, and 3σ . The two panels only show four out of eight (*left*) and five out of nine (*right*) of the free parameters.

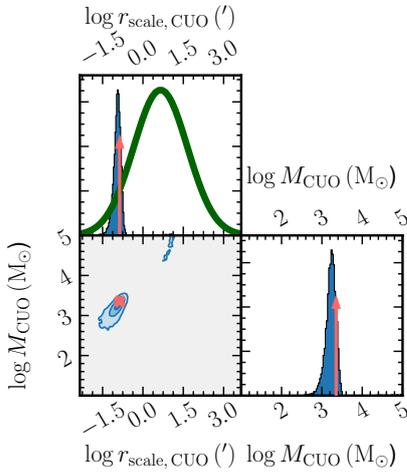


Fig. 13. Selected marginal distributions of the CUO effective radius and mass, and their covariance, for a preliminary MAMPOSST-PM run for an isotropic, single-population plus Plummer CUO SD profile, similar to model 10, but with a prior on the log CUO scale radius centered at $r_{-2, \text{CUO}} = R_{e, \text{GC}} = 4.51$. The notation is the same as in Fig. 12.

However, both AICc and BIC indicate strong evidence against the IMBH hypothesis in comparison with the presence of a CUO, with differences of 15.9 in AICc and 8.9 in BIC between isotropic, single-population models 1 and 10. With two-population GC stars, we have differences of 12.2 in AICc between isotropic models 6 and 14. But with BIC, the difference is only 5.2, leading to only moderately strong evidence (92% confidence) of favoring the CUO over the IMBH. But given our preference for AICc for the

complex physics of GC kinematics with imperfect models for the surface density profile for example (see Sect. 8), we conclude that the evidence is strong in favor of a dark component that is diffuse for the two-population model. Hence, the unseen inner matter of NGC 6397 is very likely to be diffuse. We therefore now investigate the CUO model in more detail.

9.4. Inner subcluster of unresolved objects (CUO)

9.4.1. CUO density profile

If the dark component is diffuse as a CUO instead of a singular IMBH, we first need to measure its extent. MAMPOSST-PM can provide constraints on the shape and scale radius of the density profile of the unseen population, through the sole use of the conditional probabilities of velocity at given projected radius, $p(\mathbf{v}|R)$, without directly fitting the distribution of projected radii.

We first assumed a [Plummer \(1911\)](#) model for the CUO with the same effective radius as that of the GC stars (4.5), but with a wide standard deviation (1 dex) for the Gaussian prior on log scale radius. Interestingly, as seen in Fig. 13, MAMPOSST-PM converged to $r_{-2} = 7''$, thus an effective (half-projected number) radius of only 4.3 for the CUO¹⁶, which is 60 times lower than the median of the prior, thus confirming our suspicion that the CUO might indeed be significantly more concentrated than the main-sequence stars.

We then ran MAMPOSST-PM for three models with a much smaller scale radius prior for the CUO (30''), to allow for more

¹⁶ We used the Plummer model relations $v(r) \propto (r^2 + a^2)^{-5/2}$, $\Sigma(R) \propto (R^2 + a^2)^{-2}$, $M(r)/M_\infty = r^3/(r^2 + a^2)^{3/2}$, with $R_e/a = 1/2$ and $r_{-2}/a = \sqrt{2/3}$.

accurate fits. We used three different CUO density models: a Sérsic model, a [Plummer \(1911\)](#) model, and a [Hernquist \(1990\)](#) model. All three led to very small scale radii, which convert to effective (half projected number) radii $R_e = 9''.6$, $17''.4$ and $4''.4$ for the Sérsic, Hernquist and Plummer models, respectively¹⁷.

The Sérsic model (12) is the most likely one, closely followed by the Plummer model (10). AICc has a weak preference for the Plummer model, while BIC prefers the Plummer model, with strong evidence against the more complex Sérsic model, but weak evidence against the Hernquist model (11). Model 12 with the Sérsic CUO density profile produced a low Sérsic index of $n = 0.92$, which leads to a shallow inner slope that is not too different from the zero slope of the inner Plummer density profile (see Fig. A.1 from [Vital & Mamon 2020](#)). The steeper inner profile of the Hernquist density profile makes it less similar to the $n = 0.92$ Sérsic model than is the Plummer profile, as confirmed by the Hernquist model 11 producing the lowest likelihood (highest $-\ln \mathcal{L}$) among the three models.

In summary, it is hard to distinguish which is the best density model for the CUO scenario. There is weak evidence for a shallow slope. We adopted the Plummer model given that it is the preferred of the three density models for both AICc and BIC, albeit with weak evidence for both.

9.4.2. Presence of an IMBH in addition to the CUO

One may ask whether the center of NGC 6397 can host both an IMBH and a diffuse dark component. Model 9 contains both, but it is somewhat less likely than model 10, and in comparison it is strongly disfavored by BIC, although only weakly disfavored by AICc. We note that the Plummer model used for the CUO is the one that best distinguishes the CUO from a possible additional IMBH. Moreover, the recovered mass of the additional IMBH is so small $42_{-26}^{+92} M_\odot$ that it can no longer be called an IMBH.

9.5. Two-mass populations

Table 3 also displays the MAMPOSST-PM results of 4 models with two populations, split by apparent magnitude, hence by mass (Sect. 7.2.2). Model 14, with CUO but no IMBH, has the highest likelihood of the three models considering isotropy. It is very strongly favored with AICc over models 6 (with IMBH but no CUO) and 8 (no IMBH nor CUO). It is also preferred by BIC, which strongly favors it over model 8, but only marginally favors it over model 6 ($\Delta\text{BIC} = 5.16$, leading to 92% confidence in preferring model 14 over model 6, according to Eq. (28)). This preference of the CUO model over the IMBH model resembles that found for the single population (Sect. 9.3).

But there are differences in the MAMPOSST-PM results between single-population and two-population models, both in their Bayesian evidence and in their best-fit parameters. Indeed, the two-population model 14 shows the highest likelihood and AICc evidence of the first fourteen models listed in Table 3. In particular, comparing two-population models to their single-population equivalents, there is strong AICc evidence ($\Delta\text{AICc} = 13.7$) favoring two-population model 14 than 1-population model 10. But there is strong BIC evidence ($\Delta\text{BIC} = 7.3$) the other way, with model 10 displaying the best BIC evidence of all fourteen models listed in Table 3.

The reader may note that the GC stellar masses summed over the one or two populations are 9% greater in the two-population

runs for CUO without IMBH scenarios (model 14 versus model 10) and 13% higher for IMBH without CUO scenarios (model 6 versus model 1). Moreover, the IMBH mass is 22% lower in the two-population model 6 relative to the single-population model 1. Given that the standard deviations in the log GC mass (single or sum of Bright and Faint) are less than 0.03 dex, the difference in the means of $\log M_{\text{GC}}$ of 0.086 dex is highly significant for samples of order of 10^5 points, as we checked with a Student t test. Similarly, the standard deviations in $\log M_{\text{IMBH}}$ are of order 0.2 dex; thus the difference of 0.037 dex in their means is again highly significant, as we also checked with the Student t test.

MAMPOSST-PM yields interesting results on the differences between the bright and faint populations. First, the two-population runs can be tested for the respective masses in each. Despite our prior of equal masses for each, MAMPOSST-PM returns best-fit bright population mass fractions of 0.49, 0.44, and 0.57 for models 6 (IMBH without CUO), 8 (no IMBH nor CUO), and 14 (CUO without IMBH). Only model 14 has a bright fraction close to the expected value of 0.56 (Sect. 7.2.2).

Secondly, in all three two-population models, the brighter population has a much lower scale radius than its fainter counterpart, by factors of 2.5, 2.6 and 2.1 for models 6, 8, and 14, respectively. These lower scale radii for the bright population are highly statistically significant. Indeed, the fractions of MCMC chain elements leading to higher scale radius of the brighter population are less than 0.04% for models 6, 8, and 14. Therefore, MAMPOSST-PM is able to find very strong kinematic signatures of luminosity (hence mass) segregation, by fitting $p(v|R)$ with the same priors on the scale radii of the two populations, without directly fitting the distribution of projected radii.

Recall that we used the same Gaussian priors on scale radius for both populations. As can be seen in Fig. 14, these priors may have been too narrow, and MAMPOSST-PM may have thus under-estimated the differences in the scale radii of the bright versus faint populations.

Figure 15 illustrates the quality of the 3 classes of models in reproducing the observed velocity dispersion profiles. Model 8, with no additional dark component, clearly underestimates the velocity dispersions below $10''$. Model 6, with an IMBH, overestimates it below $4''$. Finally, model 14, with a CUO, does best (the non-perfect match appears to be caused by sampling noise).

10. Conclusions and discussion

10.1. Main results

We ran the MAMPOSST-PM code on a combined set of PMs from HST and *Gaia*, as well as with LOS velocities from MUSE. Our results indicate that the GC star velocities are close to isotropic out to 2 effective radii. Our models with inner central components (IMBH or diffuse subcluster of unresolved objects, CUO) are strongly preferred over models without any. Models where the central mass is in the form of an IMBH find favored masses of 500–650 M_\odot , with a 5th lower percentile of 200 M_\odot . But we found better likelihoods and Bayesian evidence for a diffuse central unseen mass component (CUO), with effective (projected) radius of order of 2.5 to 5'' and mass of order 1000 to 2000 M_\odot .

10.2. Robustness

We tested the robustness of our results for several variations in our assumptions and our cuts to the data. These are: Single versus two stellar populations, the adopted SD profile, the minimum allowed PM error, and the HST quality flags.

First, the preference for the CUO component is relatively robust to our analysis with a single bright population or two

¹⁷ We used the Hernquist model relations $v(r) \propto a^4/[r(r+a)^3]$, $M(r)/M_\infty = r^2/(r+a)^2$, with $R_e/a = 1.8153$ ([Hernquist 1990](#)), with his expression for $\Sigma(R)$, and also $r_{-2}/a = 1/2$.

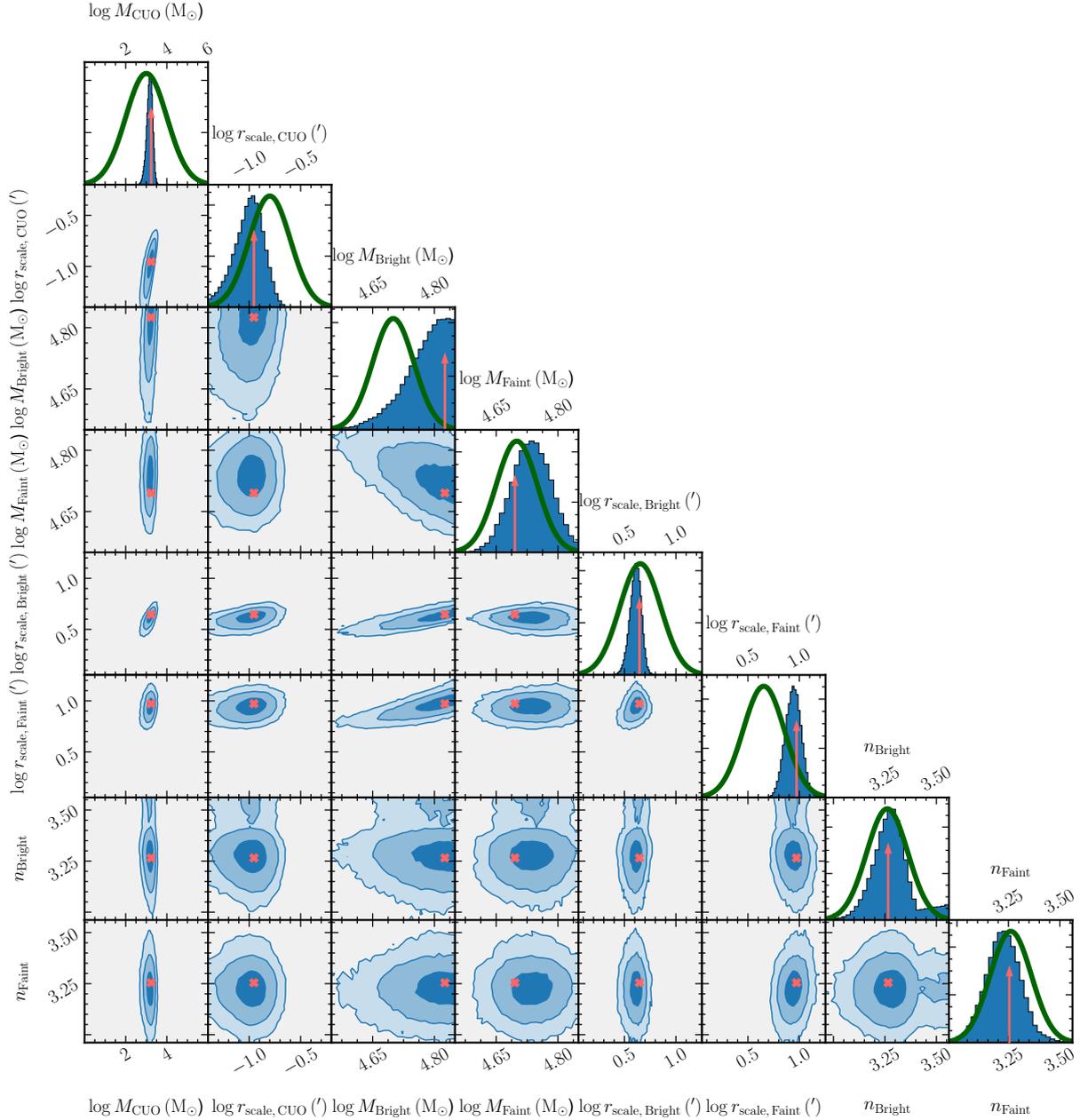


Fig. 14. Same as Fig. 12, but for model 14 with a inner subcluster of unresolved objects (CUO) with two populations of stars, a brighter (heavier) and a fainter (lighter) one, all with isotropic velocities. Only eight of twelve free parameters are shown.

populations. It is also quite robust to our assumptions on velocity anisotropy from free inner and outer values to isotropic. The differences in AICc and BIC of IMBH relative to CUO are +16 and +9 for single-population isotropic, +15 and +8 for single-population, free inner and outer anisotropy but fixing the anisotropy transition radius to the density scale radius (TAND), and +12 and +5 for double-population isotropic. In other words, CUO is very strongly preferred to IMBH by both AICc and BIC. We also ran an isotropic model restricting the sample to stars with $F606W$ and G magnitudes brighter than 17.5, to limit the effects of possible mass segregation of these bright stars with the fainter, lower-mass stars. We found, again, that the CUO scenario is strongly preferred over the IMBH one ($\Delta\text{AICc} = 7.0$).

The preference for CUO versus IMBH is also robust to the choice of the density profile. Indeed, adopting a Sérsic profile, but fixing the effective radius to 2.9 as found by Trager et al.

(1995), produces a worse match to the kinematic data. However, it strongly prefers the CUO model to the IMBH one, but provides much weaker constraints on the central masses. Similarly, forcing an inner cored profile to the luminous mass, with a Plummer model, keeping the same effective radius, the kinematics indicate much worse AICc compared to Sérsic. The AICc then strongly prefers anisotropic models, for which it still strongly prefers the CUO option to the IMBH one.

In addition, the preference for a diffuse central mass (CUO) versus an IMBH is also robust to the choice of maximum allowed PM error. We obtain the same $\Delta\text{AICc} = 12$ for both $\epsilon_\mu < 0.276 \text{ mas yr}^{-1}$ and our standard choice of $\epsilon_\mu < 0.197 \text{ mas yr}^{-1}$, as well as the same $\Delta\text{BIC} = 5$. On the other hand, to $\epsilon_\mu < 0.141 \text{ mas yr}^{-1}$, yields $\Delta\text{AICc} = 6.5$ instead, presumably because of the small dataset satisfying this low maximum allowed PM error.

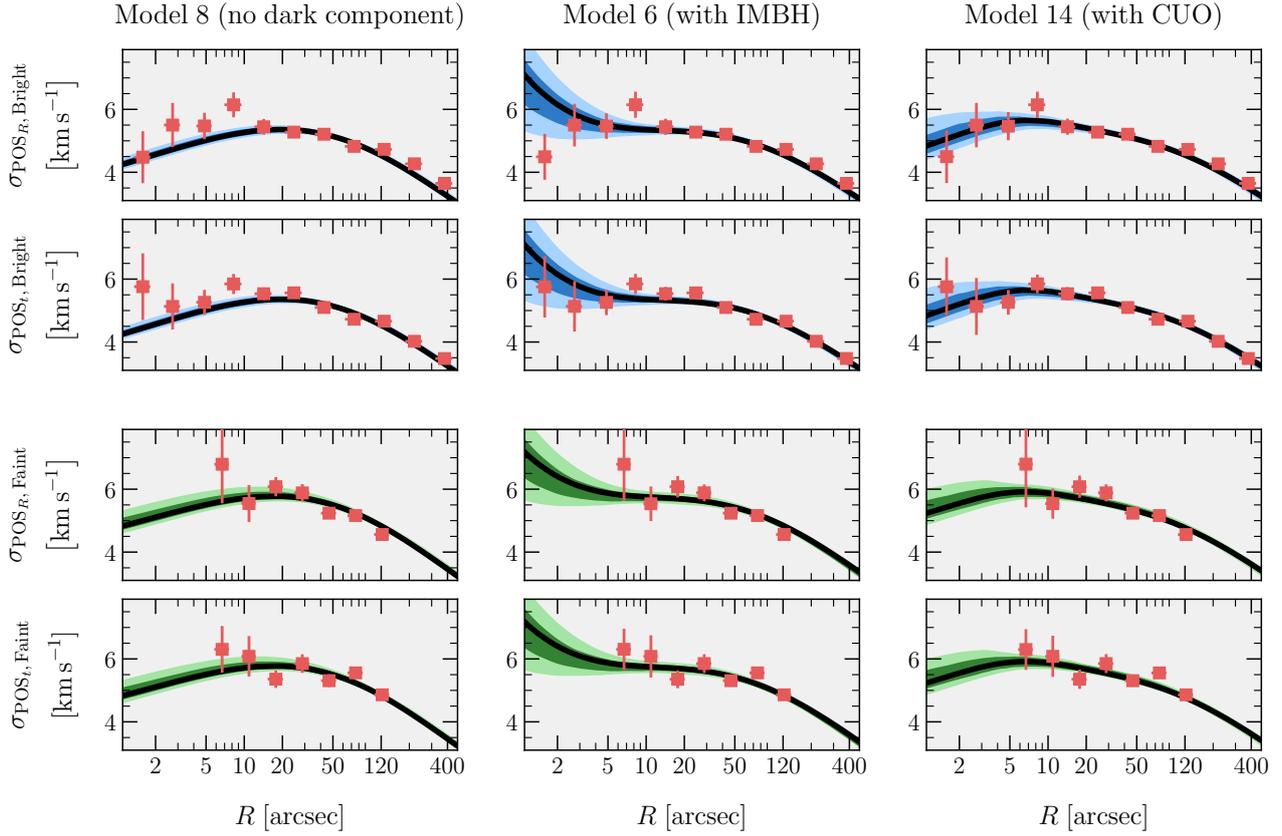


Fig. 15. Goodness of fit plots of plane of sky (POS) velocity dispersions as a function of the projected radius for models 8 (no dark component, *left*), 6 (IMBH, *middle*) and 14 (central unresolved objects, *right*), which all consider two stellar populations, one of brighter (blue) stars and another of fainter (green) ones. The black curves display the maximum likelihood solutions, while the darker and lighter shaded regions show the [16, 84] and [2.5, 97.5] percentiles, respectively. The red squares show the data in logarithmic spaced bins. The vertical error bars were calculated using a bootstrap method, while the horizontal error bars considered the radial quantization noise.

Finally, we also tested the robustness of our conclusions when also applying HST PM quality flags, such as the reduced χ^2 flag. Limiting our sample to stars satisfying $(\chi^2_{\mu_\alpha \cos \delta} + \chi^2_{\mu_\delta})/2 < 1.25$, which is the threshold chosen by Bellini et al. (2014) for NGC 7078, we noticed that the fitted parameters agreed within 1σ to the ones in Table 3. Furthermore, the CUO scenario was again favored over the IMBH one with moderately strong AICc evidence ($\Delta\text{AICc} = 4.91$, i.e., 91% confidence) instead of the strong AICc evidence observed for the subset without this χ^2 filter. Moreover, Bellini et al. (2014) argue that no quality parameter filters are required for nearby GCs with multiple observations, such as NGC 6397, when the PM dispersion depends little on them, as we indeed checked for both the reduced χ^2 and for their QFIT parameter. Our conclusions are thus robust with respect to the HST data quality flags.

10.3. Effects of the dataset

It is useful to check how efficient are our different datasets in reaching our conclusions. Such a comparison will be useful for future mass-orbit modeling of other GCs, for example in the absence of HST and LOS data.

Table 4 displays a summary of the masses and CUO size from single-component isotropic runs using different datasets. We used the same SD profile for all samples.

First, with MUSE data alone, we cannot constrain the mass of the IMBH. In contrast, Kamann et al. (2016) found $M = 600 \pm$

Table 4. Masses and CUO scale radius obtained using different datasets.

Dataset	N_{data}	M_{GC} [$10^4 M_\odot$]	M_{IMBH} [M_\odot]	M_{CUO} [$10^3 M_\odot$]	r_{CUO} [arcmin]
MUSE	528	$8.81^{+1.66}_{-1.15}$	22^{+398}_{-2}	–	–
MUSE	528	$8.77^{+1.69}_{-1.16}$	–	$2.73^{+3.29}_{-2.70}$	$0.32^{+98.86}_{-0.30}$
HST	7209	$8.80^{+0.74}_{-0.40}$	472^{+172}_{-238}	–	–
HST	7209	$10.20^{+0.89}_{-1.09}$	–	$1.87^{+0.25}_{-0.79}$	$0.11^{+0.03}_{-0.04}$
<i>Gaia</i>	1905	$10.79^{+0.08}_{-1.32}$	165^{+912}_{-141}	–	–
<i>Gaia</i>	1905	$10.11^{+0.74}_{-0.76}$	–	$0.35^{+5.25}_{-0.32}$	$15.02^{+76.26}_{-13.01}$
All 3 (mod. 1)	8255	$9.75^{0.66}_{-0.70}$	658^{+70}_{-338}	–	–
All 3 (mod. 10)	8255	$10.68^{+0.52}_{-0.68}$	–	$2.01^{+0.16}_{-0.76}$	$0.12^{+0.02}_{-0.04}$

Notes. The models are all isotropic single-component. The last two refer to our models 1 and 10, respectively.

$200 M_\odot$ for the IMBH. The difference can be explained by the much more liberal cut in velocity errors used by Kamann et al. (2016), 5 km s^{-1} , compared to our cut at 2.2 km s^{-1} , which led them to have 4608 stars compared to our 528 stars. Our small MUSE sample prevents us from constraining the CUO mass and size.

This is also true using *Gaia* data alone. It is not surprising, given that we have only 152 *Gaia* stars within the inner $100''$ and none within the inner $10''$. This low number is caused by our cut on PM errors. Later *Gaia* releases will have lower PM

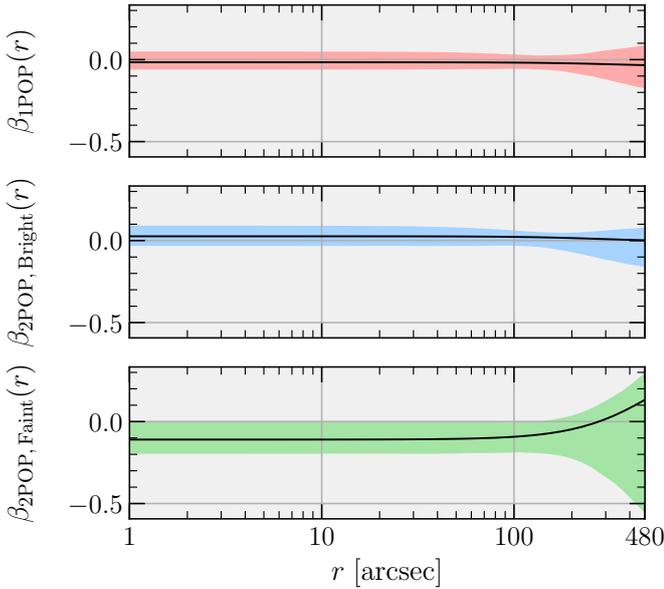


Fig. 16. Velocity anisotropy profiles of NGC 6397 for models 13 (*top*) and 15 (bright in *middle* and faint at *bottom*).

errors, and will thus have larger numbers of inner stars to use for mass-orbit modeling.

Moreover, a subset with just HST and MUSE, extending up to $2'.5$ is not able to constrain the outer isotropy indicated by *Gaia* DR2. Indeed, a run similar to model 13 with this dataset yielded $\beta(r = 8') = -1.0^{+0.5}_{-0.3}$, which is a significantly tangential anisotropy, in contrast with $\beta(8') = -0.03^{+0.11}_{-0.14}$ when *Gaia* data is also considered.

In summary, *Gaia* offers much improvement to HST(+MUSE) data, by (1) allowing a better estimation of the surface density profile, hence of the 3D density profile, and (2) allow measuring velocity anisotropy well beyond the effective radius. With the prior knowledge of the surface density profile and velocity anisotropy profile, the addition of *Gaia* provides somewhat better constrained masses for the inner component (IMBH or CUO). On the other hand, *Gaia* data by itself, has much too few stars with sufficiently accurate PMs at low projected radii to probe, the IMBH or CUO mass, and even less the CUO size. This highlights the requirement of combining HST and *Gaia* for proper mass-orbit modeling of the PM of stars in GCs.

10.4. Velocity anisotropy

Our combination of HST data probing the inner regions of NGC 6397 and *Gaia* data probing the outer regions, allowed us to obtain a wide-range view of the variation of the orbital anisotropy with radius. Bayesian evidence favors isotropy throughout the cluster. This confirms the *projected* isotropy previously determined by Heyl et al. (2012) from $3'.5$ to $7'$ using HST (with 10% individual errors), as well as Watkins et al. (2015a): $\sigma_{\text{POSr}} = 0.98 \sigma_{\text{POSr}}$, up to $R_e/10$ using HST, and by Jindal et al. (2019) at larger radii up to $12' \approx 3R_e$ using *Gaia*, with signs of radial motions at $R > 8'$.

But a constant projected isotropy can hide radial variations of the 3D anisotropy $\beta(r)$. Figure 16 displays the first known constraints on the velocity anisotropy profile of NGC 6397, which we obtained from our MAMPOSST-PM analyses from $1''$ to $8'$. Indeed, it is highly consistent with isotropic orbits throughout,

as seen in Fig. 16. Our parametric representation of $\beta(r)$ prevents us to see a possible sharp up- or down-turn in this radial range. Nevertheless, MAMPOSST-PM, like other mass-orbit modeling algorithms, probes physical radii beyond the maximum projected radius. This suggests that the orbits of stars at two effective radii are not much disturbed by the Milky Way, despite the differences in the two components of the mean POS velocity beyond $R = 8'$.

Only two previous studies have measured velocity anisotropy profiles in GCs. Cluster ω Cen was modeled by van der Marel & Anderson (2010), who fit the PM dispersions measured in bins of projected radius. They found a gOM anisotropy profile with $\beta_0 = 0.13 \pm 0.02$ and $\beta_\infty = -0.52 \pm 0.22$, thus slightly radial in the center and slightly tangential in the outer regions. Messier 15 was analyzed by den Brok et al. (2014), who used an updated version of JAM (Cappellari 2008). They found a gOM anisotropy profile with $\beta_0 = -0.21 \pm 0.30$ and $\beta_\infty = 0.015 \pm 0.12$, hence consistent with isotropic at all radii.

In comparison, the present study provides the first Bayesian mass-orbit modeling of discrete GC data. Our model 13, as well as the bright population of model 15, both yield inner anisotropies consistent with zero with $\approx \pm 0.05$ uncertainty (Table 3). Both yield anisotropies at $8'$ of $|\beta| \leq 0.03$ with uncertainties ≈ 0.14 (Table 3).

The velocity isotropy of NGC 6397 at fairly large radii appears to contradict models where core-collapsed GCs have isotropic inner motions and very radial outer motions (e.g., Takahashi 1995). But such models are for isolated GCs, whereas NGC 6397 has crossed the Galactic disk many times and as recently as 4 Myr ago (Sect. 2.2). GC stars are perturbed by tidal forces from the Milky Way (disk, bulge and halo, which sum up to a total gravitational potential not far from spherical, see Fig. 2.19 of Binney & Tremaine 2008). Indeed, we noticed a difference in the mean POS velocities between the radial and tangential components beyond $8'$ (top panel of Fig. 6), suggesting that the outer parts of NGC 6397 are out of equilibrium.

While the tidal field on an embedded GC inside a galaxy should be compressive and radial (Dekel et al. 2003), it is now understood that this does not lead to radial orbits in the outer envelopes of GCs, where the stars are most susceptible to tidal perturbations. Indeed, comparing stars of same binding energy, those that are on radial orbits spend more time at larger radii and feel stronger tidal forces, while those that are on tangential orbits are less tidally perturbed (see Giersz & Heggie 1997). This reduces the radial outer anisotropy caused by two-body relaxation. This is confirmed by the isotropy at the half-mass radius of GCs with short half-mass relaxation times determined with mass-orbit modeling by Watkins et al. (2015a), who also found increasingly radial orbits at r_h for GCs with increasingly longer half-mass relaxation times.

This is also consistent with N -body simulations of the internal motions of a GC subject to the tidal field of a point-mass galaxy (Tiongco et al. 2016; Zocchi et al. 2016) or of a realistic galactic potential by itself or within an infalling dwarf galaxy (Bianchini et al. 2017). These simulations indicate that, after the establishment of radial outer anisotropy at core collapse, the outer velocity anisotropy becomes less radial thanks to the tidal field of the Milky Way, especially when the tidal field is strong, as measured by small values of the ratio of Jacobi to half-mass radii, r_J/r_h , also called the filling factor. One can compare these simulations in detail to NGC 6397. We estimate the filling factor for NGC 6397 using its estimated pericentric and apocentric radii of 2.9 and 6.6 kpc (Gaia Collaboration 2018a) and

the Milky Way mass profile of [Cautun et al. \(2020\)](#), which produces $r_J = 31$ to 51 pc, hence $r_h/r_J = 0.08$ to 0.14 , assuming $R_e/r_h = 0.74$ ¹⁸. The single simulation of [Zocchi et al. \(2016\)](#) reaches this range of r_h/r_J for snapshots 3 to 11, for which $\beta(r_h)$ decreases from 0.13 to -0.02 , as the tides preferentially remove the radial orbits, leading to $\beta(r_h) = 0.00$ to 0.08 for $r_h/r_J = 0.12$, decreasing to more tangential values for higher filling factors. Similarly, among the MW-only simulations of [Bianchini et al. \(2017\)](#), who also included a realistic initial mass function and stellar evolution, only one of them has a filling factor at 10 Gyr that matches the one deduced above for NGC 6397, and this simulation produced $\beta(r_h) = 0.075$. So both studies would predict that a core-collapsed cluster like NGC 6397, subject to the tidal field of the Milky Way, should recover quasi-isotropic orbits at r_h , consistent with what we found. This needs to be checked with more refined N -body simulations that consider elongated orbits of a GC around its galaxy and possibly the incorporation of binaries at the start of the simulation.

There are several additional possible mechanisms for producing outer isotropic orbits. First, the passage through the differentially rotating Galactic disk could lead to exchanges of angular momentum leading to tangential orbits, in particular for the outer less bound stars. But on the scale of a GC, the differences in disk rotation velocity must be minute. Second, the stars in the outer parts of the GC must suffer violent relaxation during passages through the Galactic disk, but the passages are so short (a few Myr) that there may not be enough time for violent relaxation to effectively perturb the GC stars. Third, solenoidal modes of turbulence on Galactic gas clouds generate angular momentum within these clouds and within the stars formed therein, which in turn can transfer this angular momentum to the passing GC. But again, the short duration of these passages limits the amount of angular momentum that the GC stars can acquire. Therefore, the outer isotropy appears to be produced by the tidal field of the Milky Way preferentially removing those stars that were on radial orbits.

10.5. Intermediate mass black hole

Our IMBH mass estimates between 500 and $600 M_\odot$, with uncertainties of $\approx 200 M_\odot$, are consistent with those of [Kamann et al. \(2016\)](#) ($M_{\text{IMBH}} = 600 \pm 200 M_\odot$) and [Tremou et al. \(2018\)](#) ($M_{\text{IMBH}} < 610 M_\odot$). Our 95% confidence lower limits (i.e., 5th percentiles) are $201 M_\odot$ for both models 1 (single population) and 6 (two-population).

But if IMBHs grow by BH mergers, one expects that the gravitational radiation emitted during these mergers are anisotropic and lead to substantial recoil of the BH remnant of the merger ([Peres 1962](#)). Such recoils are strong enough for 70% of IMBHs of mass $< 1000 M_\odot$ to escape a GC of escape velocity 50 km s^{-1} over time ([Holley-Bockelmann et al. 2008](#)). We can estimate the escape velocity from the GC center, $v_{\text{esc}} = \sqrt{-2\Phi(0)}$, for our best IMBH model 6, given that the central potential is $\Phi_0 = -(2/\pi) b^n \Gamma(n) / \Gamma(2n) G M / R_e$ for the Sérsic model ([Ciotti 1991](#)), where $b(n)$ follows the relation given in [Ciotti & Bertin \(1999\)](#). Model 6 then has $v_{\text{esc}} = 20.9 \text{ km s}^{-1}$ and 13.1 km s^{-1} for the Bright and Faint components, respectively. Summing these two terms quadratically leads to a global escape velocity of $v_{\text{esc}} = 24.7 \text{ km s}^{-1}$ for model 6. This escape velocity is lower than the 50 km s^{-1} assumed by [Holley-Bockelmann et al. \(2008\)](#), which should lead to an even much higher fraction of

¹⁸ We expect this R_e/r_h ratio from the deprojected Sérsic of index 3.3 ([Simonneau & Prada 2004](#) approximation).

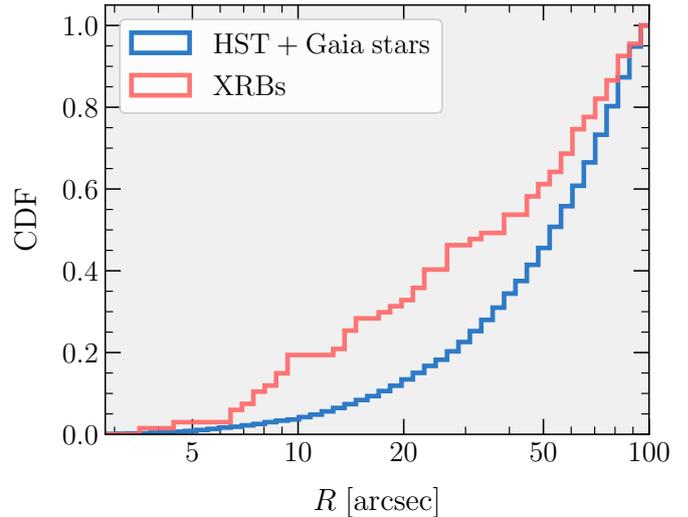


Fig. 17. Cumulative distribution functions of projected radii for our HST+*Gaia* subset in blue and for the X-ray binaries from [Bahramian et al. \(2020\)](#) in red. We considered the subsets in the range of $2.7'' < R_{\text{proj}} < 100''$, where the HST sample seemed complete, according to Fig. 8.

escaping IMBHs of mass $1000 M_\odot$. And our 95% confidence upper limits on the IMBH mass (without a CUO) are lower, again implying that a yet higher fraction of IMBHs escape their GC. A simpler model by [Holley-Bockelmann et al. \(2008\)](#) indicates that 25 IMBH+BH mergers should lead to 85% of $500 M_\odot$ IMBHs escaping their GC of escape velocity 17 km s^{-1} . This weakens our confidence in the central IMBH scenario for NGC 6397.

10.6. Cluster of unresolved objects

The scale radius of the CUO in NGC 6397 is consistent with a feature in the SD profile of NGC 6397. Indeed, as seen in the right panel of Fig. 10, the SD profile suggests a separation between two populations at $R = 10''$, consistent with the CUO effective radius of a 2.5 to $5''$.

We now discuss the nature of the subcluster of unresolved objects that we found in the center of NGC 6397. The much smaller scale radius for the CUO indicates that the objects must be more massive than the stars that we studied, thus more massive than $m_{\text{bright}} = 0.77 M_\odot$ (Sect. 7.2.2). Indeed, such unresolved massive objects would sink to the center by dynamical friction ([Chandrasekhar 1943](#)). The core-collapsed state of NGC 6397 is a strong argument in favor of a population of heavier stars in the cluster's inner regions, as it suggests a short two-body relaxation time, hence mass segregation. These heavier objects constituting the CUO could be white dwarfs, neutron stars, stellar black holes, or unresolved binaries.

We compared the radial distribution from our cleaned merged sample and from X-ray binaries from [Bahramian et al. \(2020\)](#). Figure 17 shows the cumulative distribution function (CDF) of these two datasets in the range of 2.7 – $100''$, where our HST data was complete. It is clear that these two populations do not follow the same radial distribution (we find a KS p -value of 1.7×10^{-4}). Furthermore, the bulk of the X-ray binaries seems to be located within $6''$ – $50''$ arcsec (Fig. 17). This is consistent with the CUO effective radius of 2.5 to $5''$ (Table 3).

One may ask which among white dwarfs, neutron stars, BHs and massive binaries dominates the mass of the CUO. We can

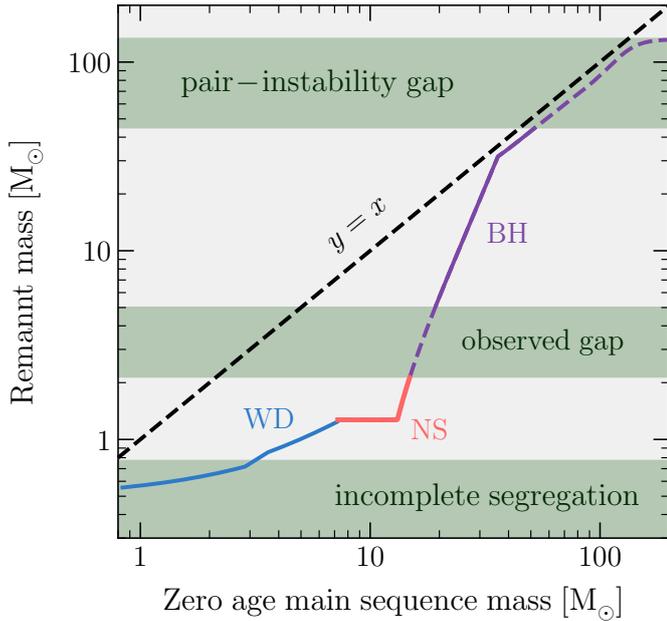


Fig. 18. Initial – final mass relation of white dwarfs (WD, MIST model from Cummings et al. 2018, blue), neutron stars (NS, red) and black holes (BH, purple) (PARSEC/SEVN model with $Z = 0.0002$ from Spera et al. 2015). Factors of two changes in Z are barely visible in the red (purple) line and do not affect the blue line. The *lower green band* indicates incomplete mass segregation because some of the main-sequence stars are more massive than the white dwarfs. The *middle green band* indicates the gap where no BHs have (yet) been detected. The *upper green bands* highlight the gap where pair-instability supernovae fully explode the progenitor star without leaving a black hole. The black line shows equality as a reference.

first discard unresolved binaries of main-sequence stars. Indeed, if they are unobserved, their total magnitude must be fainter than $F606W = 22$, which is the rough magnitude limit of our HST sample (*Gaia* is not relevant given the very low effective radius of the CUO). There is no way that their total mass can exceed that of the bulk of our sample. We could then have unresolved binaries of a main-sequence star with a compact star (white dwarf or neutron star) or possibly a BH. But the main-sequence star will have a mass of $m_{\text{faint}} = 0.25 M_{\odot}$ (Sect. 7.2.2), thus at least three times lower than our cut at $m_{\text{bright}} = 0.77 M_{\odot}$ and at least 20 times lower than that of a BH. Therefore, the main-sequence mass can be neglected and we are left with the compact star or BH.

We can compare the total mass in white dwarfs, neutron stars and BHs, by integrating over the zero-age (stellar mass function of the) main sequence (ZAMS):

$$\int_{m_{\text{min}}}^{m_{\text{max}}} m_{\text{remnant}}(m) n(m) dm, \quad (29)$$

where $n(m)$ is the ZAMS. We adopted the initial – final (remnant) mass relations from Eqs. (4)–(6) of Cummings et al. (2018) for white dwarfs and from Eqs. (C1), (C2), (C11), and (C15) of Spera et al. (2015) for neutron stars and BHs. Figure 18 displays these initial - final mass relations. We took a minimum remnant mass of $0.77 M_{\odot}$, corresponding to the maximum main-sequence mass (Sect. 7.2.2), below which mass segregation is not effective against the most massive main-sequence stars. The maximum possible neutron star mass is $M \approx 2.15 M_{\odot}$ (Rezzolla et al. 2018). We conservatively assumed a mass gap between 2.15 and $5 M_{\odot}$ from the lack of LIGO detections in this mass range. We

also considered a maximum stellar BH mass above which pair-instability supernovae fully explode the progenitor star leaving no remnant, adopting $M_{\text{BH,max}} = 45 M_{\odot}$ (Farmer et al. 2019) or $52 M_{\odot}$ (Woosley 2017). We ignored the formation of very massive BHs above the pair-instability gap (i.e., $M_{\text{BH}} > 133 M_{\odot}$ Woosley 2017).

Summing the masses of each component by integrating Eq. (29) over the ZAMS mass function (i.e., the initial mass function, which in this mass range always has the Salpeter 1955 slope of -2.3) leads to BHs accounting for $\approx 58\%$ of the CUO, with only $\approx 30\%$ from white dwarfs and $\approx 12\%$ from neutron stars. These fractions vary little with the maximum allowed BH mass: with 55% of the CUO mass in BHs with $M_{\text{BH}} < 45 M_{\odot}$ (Farmer et al. 2019) or 60% with $M_{\text{BH}} < 52 M_{\odot}$ (Woosley 2017). These fractions are insensitive to the metallicity of NGC 6397 from $Z = 0.00013$ to 0.0004 (i.e., from 1% to 3% solar, consistent with the metallicities derived for NGC 6397 by Marín-Franch et al. 2009 and Jain et al. 2020, respectively).

However, we must take this CUO mass dominance by BHs with caution because BHs may merge, losing of order of 5% of their mass to gravitational waves (e.g., Abbott et al. 2016), leading to kicks, some of which are strong enough to drive them out of the GC (Sect. 10.5). Still, one part of the BH population may merge and end up escaping the GC, while another part has not merged, but would nevertheless be located at low radii thanks to orbital decay by dynamical friction. If the CUO mass fraction in BHs were at least $f_0 = 0.55$ or 0.6 before merging and escapes, and if a mass fraction f_d of this BH component disappeared through mergers or escapes, then BHs would still dominate the CUO if $f_0(1 - f_d) > 1 - f_0$, i.e. $f_d < 2 - 1/f_0 = 0.19$ or 0.33 considering the maximum BH mass of $45 M_{\odot}$ (Farmer et al. 2019) or $52 M_{\odot}$ (Woosley 2017). Put another way, BHs could contribute to half the CUO mass, if the maximum surviving BH mass is $40 M_{\odot}$ according to our integrations of Eq. (29). The more massive BHs would sink faster to the center by dynamical friction and preferentially merge, leaving these lower mass ones. This discussion is a simplification because orbital decay by dynamical friction is stochastic, and one needs to test this with N -body simulations.

Our preference for the CUO model agrees with the suggestions by Zocchi et al. (2019) and Mann et al. (2019) that such a CUO can mimic an IMBH. Both studies obtained constraints on the mass of the stellar BH population in ω Cen and 47 Tuc, respectively. While those GCs are not core-collapsed, contrary to NGC 6397, it is interesting to compare our results with theirs. We determine that the CUO accounts for 1 to 2% of the GC mass, and if BHs do not merge or escape, the bulk of the CUO should be in BHs. Zocchi et al. (2019) tried different BH mass fractions, and their best fits were obtained with 1 to 5% of the GC mass in BHs. Mann et al. (2019) used as input a BH component whose mass is 1.4% of the GC mass, but they show that only 8.5% of this mass can be retained to avoid negative IMBH mass. Zocchi et al. (2019) did not provide a value for the scale radius of their stellar BHs, while Mann et al. (2019) provided a radius as an input parameter, which amounts to 6% of their fitted GC scale radius. In comparison, we are able to strongly constrain the CUO scale effective to $1.7 \pm 0.5\%$ of the GC effective radius, making our BH system much more concentrated.

A more robust conclusion of our integration of Eq. (29) is that white dwarfs always dominate the neutron stars, by a factor of ≈ 4 , regardless of the possible dominance of BHs in the mass of the CUO component. Moreover, neutron stars can also merge together (or with black holes), and such an event has been detected by LIGO (Abbott et al. 2017). Presumably, the

gravitational waves emitted will be much weaker and the lower momentum vector of the waves should lead to smaller kicks on the remnant (despite the lower masses of neutron stars compared to BHs). In summary, the CUO mass should be dominated by BHs, but must also contain white dwarfs, contributing four times more mass to the CUO than neutron stars.

Interestingly, stellar black holes in a dense inner sub-cluster, such as our CUO, could represent a major channel for the gravitational wave detections by LIGO/VIRGO (Portegies Zwart & McMillan 2000). These detections involve black holes of mass $\gtrsim 20 M_{\odot}$ (see Abbott et al. 2019) where electromagnetic wave detections are still lacking (e.g., Casares 2007). Black hole mergers should be even more frequent, presently, among the core-collapsed GCs such as NGC 6397, whose high inner densities allow for faster dynamical processes, such as dynamical friction, dynamical creation of black hole binaries (Heggie & Hut 2003), and subsequent hardening in three-body encounters (Heggie 1975). Our dynamical analysis provides a promising route to determine the locations of these stellar mass black holes.

10.7. Final thoughts

The future of GC and IMBH science is very exciting, thanks to *Gaia* now supplementing HST PMs at large projected radii. Our discovery of a diffuse central mass, composed in large part of stellar mass black holes, enriches the physics of the inner regions of GCs, and renders the search of IMBHs in Milky Way GCs even more delicate. Continued pointings of GCs with HST and soon *James Webb* Space Telescope will lead to longer baselines and more accurate PMs. The third data release of the *Gaia* mission will double the PM precision, thus enabling more accurate mass-orbit modeling, not only of nearby GCs such as NGC 6397, but also of more distant ones, in conjunction with HST data. Future gravitational wave missions such as the Laser Interferometer Space Antenna (LISA) will probe IMBHs as well as stellar BHs, including in our Milky Way (Sesana et al. 2020). The physics of the inner parts of GCs with possible IMBHs as well as subclusters of stellar-mass black holes is truly enticing.

Acknowledgements. We are most grateful to Andrea Bellini and Sebastian Kamann for respectively providing PM and LOS velocity data for NGC 6397, without which this paper would not have been possible. We also thank Thomas de Boer and Mark Gieles for providing their unpublished estimate of the projected half-light radius of NGC 6397. We acknowledge the anonymous referee who provided useful comments and references. We also thank Gwenaél Boué and Radek Wojtak for useful comments and suggestions in our mass modeling approach, Lukas Furtak for his help with the PARSEC code, Marta Volonteri for insightful discussions, and Pierre Boldrini, Laura Watkins and Natalie Webb for useful references. Eduardo Vitral was funded by a grant from the Centre National d'Études Spatiales (CNES) and an AMX doctoral grant from École Polytechnique. This work greatly benefited from public software packages SCIPY, ASTROPY (Astropy Collaboration 2013), NUMPY (van der Walt et al. 2011), COSMOMC (Lewis & Bridle 2002), EMCEE (Goodman & Weare 2010), MATPLOTLIB (Hunter 2007), as well as Pierre Raybaut for developing the SPYDER Integrated Development Environment. We also intensively used TOPCAT (Taylor 2005).

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *Phys. Rev. Lett.*, **116**, 061102
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, *Phys. Rev. Lett.*, **119**, 161101
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, *ApJ*, **882**, L24
 Akaike, H. 1973, *Information Theory and an Extension of the Maximum Likelihood Principle* (New York: Springer), 199
 Akaike, H. 1983, *Int. Stat. Inst.*, **44**, 277
 Anderson, J., Sarajedini, A., Bedin, L. R., et al. 2008, *AJ*, **135**, 2055
 Arenou, F., Luri, X., Babusiaux, C., et al. 2018, *A&A*, **616**, A17

Aros, F. I., Sippel, A. C., Mastrobuono-Battisti, A. R., et al. 2020, *MNRAS*, **499**, 4646
 Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33
 Auriere, M. 1982, *A&A*, **109**, 301
 Bahramian, A., Strader, J., Miller-Jones, J. C. A., et al. 2020, *ApJ*, **901**, 57
 Baumgardt, H. 2017, *MNRAS*, **464**, 2174
 Baumgardt, H., & Mieske, S. 2008, *MNRAS*, **391**, 942
 Baumgardt, H., Hilker, M., Sollima, A., & Bellini, A. 2019, *MNRAS*, **482**, 5138
 Bellini, A., Anderson, J., van der Marel, R. P., et al. 2014, *ApJ*, **797**, 115
 Bianchini, P., Norris, M. A., van de Ven, G., et al. 2016, *ApJ*, **820**, L22
 Bianchini, P., Sills, A., & Miholics, M. 2017, *MNRAS*, **471**, 1181
 Bianchini, P., van der Marel, R. P., del Pino, A., et al. 2018, *MNRAS*, **481**, 2125
 Binney, J., & Mamon, G. A. 1982, *MNRAS*, **200**, 361
 Binney, J., & Tremaine, S. 2008, *Galactic Dynamics: Second Edition* (Princeton: Princeton University Press)
 Box, G. E. P. 1979, in *Robustness in Statistics*, eds. R. L. Launer, & G. N. Wilkinson (New York: Academic Press), 201
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, **427**, 127
 Brown, T. M., Casertano, S., Strader, J., et al. 2018, *ApJ*, **856**, L6
 Burnham, K. P., & Anderson, D. R. 2002, *A Practical Information-Theoretic Approach*, 2nd edn. (New York: Springer)
 Cappellari, M. 2008, *MNRAS*, **390**, 71
 Carballo-Bello, J. A., Gieles, M., Sollima, A., et al. 2012, *MNRAS*, **419**, 14
 Carretta, E., Bragaglia, A., Gratton, R., & Lucatello, S. 2009a, *A&A*, **505**, 139
 Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009b, *A&A*, **505**, 117
 Casares, J. 2007, in *Black Holes from Stars to Galaxies – Across the Range of Masses*, eds. V. Karas, & G. Matt, 238, 3
 Cautun, M., Benítez-Llambay, A. R., Deason, A. J., et al. 2020, *MNRAS*, **494**, 4291
 Chandrasekhar, S. 1942, *Principles of Stellar Dynamics* (Chicago: University of Chicago Press)
 Chandrasekhar, S. 1943, *ApJ*, **97**, 255
 Chilingarian, I. V., Katkov, I. Y., Zolotukhin, I. Y., et al. 2018, *ApJ*, **863**, 1
 Ciotti, L. 1991, *A&A*, **249**, 99
 Ciotti, L., & Bertin, G. 1999, *A&A*, **352**, 447
 Cordoni, G., Milone, A. P., Mastrobuono-Battisti, A., et al. 2020, *ApJ*, **889**, 18
 Courteau, S., Cappellari, M., de Jong, R. S., et al. 2014, *Rev. Mod. Phys.*, **86**, 47
 Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *MNRAS*, **365**, 11
 Cummings, J. D., Kalirai, J. S., Tremblay, P. E., Ramirez-Ruiz, E., & Choi, J. 2018, *ApJ*, **866**, 21
 Davis, D. S., Richer, H. B., Anderson, J., et al. 2008, *AJ*, **135**, 2155
 de Boer, T. J. L., Gieles, M., Balbinot, E., et al. 2019, *MNRAS*, **485**, 4906
 Dekel, A., Devor, J., & Hatzroni, G. 2003, *MNRAS*, **341**, 326
 den Brok, M., van de Ven, G., van den Bosch, R., & Watkins, L. 2014, *MNRAS*, **438**, 487
 Djorgovski, S., & King, I. R. 1986, *ApJ*, **305**, L61
 Dotter, A., Sarajedini, A., Anderson, J., et al. 2010, *ApJ*, **708**, 698
 Drukier, G. A., Slavin, S. D., Cohn, H. N., et al. 1998, *AJ*, **115**, 708
 Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019, *ApJ*, **875**, L1
 Farmer, R., Renzo, M., de Mink, S. E., Marchant, P., & Justham, S. 2019, *ApJ*, **887**, 53
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
 Gaia Collaboration (Helmi, A., et al.) 2018a, *A&A*, **616**, A12
 Gaia Collaboration (Mignard, F., et al.) 2018b, *A&A*, **616**, A14
 Gebhardt, K., Pryor, C., Williams, T. B., & Hesser, J. E. 1995, *AJ*, **110**, 1699
 Giersz, M., & Heggie, D. C. 1997, *MNRAS*, **286**, 709
 Giersz, M., Leigh, N., Hypki, A., Lützgendorf, N., & Askar, A. 2015, *MNRAS*, **454**, 3150
 Goldsbury, R., Richer, H. B., Anderson, J., et al. 2010, *AJ*, **140**, 1830
 Goldsbury, R., Heyl, J., & Richer, H. 2013, *ApJ*, **778**, 57
 Goodman, J., & Weare, J. 2010, *Commun. Appl. Math. Comput. Sci.*, **5**, 65
 Gratton, R. G., Bragaglia, A., Carretta, E., et al. 2003, *A&A*, **408**, 529
 Greene, J. E., Strader, J., & Ho, L. C. 2020, *ARA&A*, **58**, 257
 Haehnelt, M. G., & Rees, M. J. 1993, *MNRAS*, **263**, 168
 Haiman, Z. 2013, in *The Formation of the First Massive Black Holes*, eds. T. Wiklund, B. Mobasher, & V. Bromm, *Astrophys. Space Sci. Lib.*, **396**, 293
 Hansen, B. M. S., Anderson, J., Brewer, J., et al. 2007, *ApJ*, **671**, 380
 Harris, W. E. 1996, *AJ*, **112**, 1487
 Harris, W. E. 2010, ArXiv e-prints [arXiv:1012.3224]
 Hawking, S. 1971, *MNRAS*, **152**, 75
 Heggie, D. C. 1975, *MNRAS*, **173**, 729
 Heggie, D. C., & Hut, P. 1996, in *Dynamical Evolution of Star Clusters: Confrontation of Theory and Observations*, eds. P. Hut, & J. Makino, *IAU Symp.*, **174**, 303
 Heggie, D., & Hut, P. 2003, *The Gravitational Million-body Problem: A Multidisciplinary Approach to Star Cluster Dynamics* (Cambridge: Cambridge University Press)

- Hernquist, L. 1990, *ApJ*, **356**, 359
- Heyl, J. S., Richer, H., Anderson, J., et al. 2012, *ApJ*, **761**, 51
- Holley-Bockelmann, K., Gültekin, K., Shoemaker, D., & Yunes, N. 2008, *ApJ*, **686**, 829
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, **163**, 1
- Hoyle, F., & Fowler, W. A. 1963, *Nature*, **197**, 533
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Husser, T.-O., Kamann, S., Dreizler, S., et al. 2016, *A&A*, **588**, A148
- Ibata, R., Nipoti, C., Sollima, A., et al. 2013, *MNRAS*, **428**, 3648
- Jain, R., Prugniel, P., Martins, L., & Lançon, A. 2020, *A&A*, **635**, A161
- Jindal, A., Webb, J. J., & Bovy, J. 2019, *MNRAS*, **487**, 3693
- Kaaret, P., Prestwich, A. H., Zezas, A., et al. 2001, *MNRAS*, **321**, L29
- Kamann, S., Husser, T. O., Brinchmann, J., et al. 2016, *A&A*, **588**, A149
- Kass, R. E., & Rafferty, A. E. 1995, *J. Am. Stat. Assoc.*, **90**, 773
- Lauzeral, C., Ortolani, S., Auriere, M., & Melnick, J. 1992, *A&A*, **262**, 63
- Leigh, N. W. C., Böker, T., Maccarone, T. J., & Perets, H. B. 2013, *MNRAS*, **429**, 2997
- Leonard, P. J. T. 1989, *AJ*, **98**, 217
- Lewis, A., & Bridle, S. 2002, *Phys. Rev. D*, **66**, 103511
- Lima Neto, G. B., Gerbal, D., & Márquez, I. 1999, *MNRAS*, **309**, 481
- Lin, D., Strader, J., Romanowsky, A. J., et al. 2020, *ApJ*, **892**, L25
- Lind, K., Korn, A. J., Barklem, P. S., & Grundahl, F. 2008, *A&A*, **490**, 777
- Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, A2
- Loeb, A., & Rasio, F. A. 1994, *ApJ*, **432**, 52
- Lovisi, L., Mucciarelli, A., Lanzoni, B., et al. 2012, *ApJ*, **754**, 91
- Lugger, P. M., Cohn, H. N., & Grindlay, J. E. 1995, *ApJ*, **439**, 191
- Madau, P., & Rees, M. J. 2001, *ApJ*, **551**, L27
- Mamon, G. A., Biviano, A., & Boué, G. 2013, *MNRAS*, **429**, 3079
- Mamon, G. A., Cava, A., Biviano, A., et al. 2019, *A&A*, **631**, A131
- Mann, C. R., Richer, H., Heyl, J., et al. 2019, *ApJ*, **875**, 1
- Marigo, P., Girardi, L., Bressan, A., et al. 2017, *ApJ*, **835**, 77
- Marín-Franch, A., Aparicio, A., Piotto, G., et al. 2009, *ApJ*, **694**, 1498
- Martinazzi, E., Pieres, A., Kepler, S. O., et al. 2014, *MNRAS*, **442**, 3105
- Mashchenko, S., & Sills, A. 2005, *ApJ*, **619**, 258
- McDonald, I., & Zijlstra, A. A. 2015, *MNRAS*, **448**, 502
- Merritt, D. 1985, *AJ*, **90**, 1027
- Merritt, D. 1987, *ApJ*, **313**, 121
- Miller, M. C., & Hamilton, D. P. 2002, *MNRAS*, **330**, 232
- Milone, A. P., Villanova, S., Bedin, L. R., et al. 2006, *A&A*, **456**, 517
- Milone, A. P., Piotto, G., Bedin, L. R., et al. 2012a, *A&A*, **540**, A16
- Milone, A. P., Marino, A. F., Piotto, G., et al. 2012b, *ApJ*, **745**, 27
- Noyola, E., & Gebhardt, K. 2006, *AJ*, **132**, 447
- Oppenheimer, J. R., & Snyder, H. 1939, *Phys. Rev.*, **56**, 455
- Osipkov, L. P. 1979, *Sov. Astron. Lett.*, **5**, 42
- Pastorelli, G., Marigo, P., Girardi, L., et al. 2019, *MNRAS*, **485**, 5666
- Pearson, K. 1916, *Philos. Trans. R. Soc. London Ser. A*, **216**, 429
- Peres, A. 1962, *Phys. Rev.*, **128**, 2471
- Plummer, H. C. 1911, *MNRAS*, **71**, 460
- Portegies Zwart, S. F., & McMillan, S. L. W. 2000, *ApJ*, **528**, L17
- Portegies Zwart, S. F., & McMillan, S. L. W. 2002, *ApJ*, **576**, 899
- Read, J. I., & Steger, P. 2017, *MNRAS*, **471**, 4541
- Read, J., Mamon, G. A., Vasiliev, E., et al. 2020, *MNRAS*, **501**, 978
- Reid, I. N., & Gizis, J. E. 1998, *AJ*, **116**, 2929
- Reimers, D. 1975, *Mem. Soc. R. Sci. Liege*, **8**, 369
- Rezzolla, L., Most, E. R., & Weih, L. R. 2018, *ApJ*, **852**, L25
- Richardson, T., & Fairbairn, M. 2014, *MNRAS*, **441**, 1584
- Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, *A&A*, **409**, 523
- Rodrigues, O. 1840, *J. Math. Pures Appl.*, **5**, 380
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, **737**, 103
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schmidt, M. 1963, *Nature*, **197**, 1040
- Schwarz, G. 1978, *Annal. Stat.*, **6**, 461
- Sérsic, J. L. 1963, *Bull. Assoc. Argent. Astron.*, **6**, 41
- Sersic, J. L. 1968, *Atlas de Galaxias Australes* (Cordoba, Argentina: Observatorio Astronomico)
- Sesana, A., Lamberts, A., & Petiteau, A. 2020, *MNRAS*, **494**, L75
- Shao, Z., & Li, L. 2019, *MNRAS*, **489**, 3093
- Shin, J., Kim, S. S., & Lee, Y.-W. 2013, *J. Korean Astron. Soc.*, **46**, 173
- Silverman, B. W. 1986, *Density Estimation for Statistics and Data Analysis* (New York: Chapman and Hall)
- Simonneau, E., & Prada, F. 2004, *Rev. Mex. Astron. Astrofis.*, **40**, 69
- Sollima, A., Baumgardt, H., & Hilker, M. 2019, *MNRAS*, **485**, 1460
- Spera, M., Mapelli, M., Bressan, A. R., et al. 2015, *MNRAS*, **451**, 4086
- Strigari, L. E., Bullock, J. S., & Kaplinghat, M. 2007, *ApJ*, **657**, L1
- Sugiura, N. 1978, *Commun. Stat. Theor. Meth.*, **7**, 13
- Takahashi, K. 1995, *PASJ*, **47**, 561
- Taylor, M. B. 2005, in TOPCAT & STIL: Starlink Table/VOTable Processing Software, eds. P. Shopbell, M. Britton, & R. Ebert, *ASP Conf. Ser.*, **347**, 29
- The LIGO Scientific Collaboration & the Virgo Collaboration (Abbott, R., et al.) 2020, *Phys. Rev. Lett.*, **125**, 101102
- Thompson, T. A., Kochanek, C. S., Stanek, K. Z., et al. 2020, *Science*, **368**, eaba4356
- Tiongco, M. A., Vesperini, E., & Varri, A. L. 2016, *MNRAS*, **455**, 3693
- Tiret, O., Combes, F., Angus, G. W., Famaey, B., & Zhao, H. S. 2007, *A&A*, **476**, L1
- Trager, S. C., King, I. R., & Djorgovski, S. 1995, *AJ*, **109**, 218
- Tremou, E., Strader, J., Chomiuk, L., et al. 2018, *ApJ*, **862**, 16
- Valcin, D., Bernal, J. L., Jimenez, R., Verde, L., & Wandelt, B. D. 2020, *JCAP*, **2020**, 002
- van der Marel, R. P., & Anderson, J. 2010, *ApJ*, **710**, 1063
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *Comput. Sci. Eng.*, **13**, 22
- Vasiliev, E. 2019a, *MNRAS*, **482**, 1525
- Vasiliev, E. 2019b, *MNRAS*, **484**, 2832
- Vasiliev, E. 2019c, *MNRAS*, **489**, 623
- Vesperini, E., McMillan, S. L. W., D'Antona, F., & D'Ercole, A. 2013, *MNRAS*, **429**, 1913
- Vitral, E., & Mamon, G. A. 2020, *A&A*, **635**, A20
- Volonteri, M. 2010, *A&ARv*, **18**, 279
- Watkins, L. L., van de Ven, G., den Brok, M., & van den Bosch, R. C. E. 2013, *MNRAS*, **436**, 2598
- Watkins, L. L., van der Marel, R. P., Bellini, A., & Anderson, J. 2015a, *ApJ*, **803**, 29
- Watkins, L. L., van der Marel, R. P., Bellini, A., & Anderson, J. 2015b, *ApJ*, **812**, 149
- Woosley, S. E. 2017, *ApJ*, **836**, 244
- Zel'dovich, Y. B., & Novikov, I. D. 1966, *AZh*, **43**, 758
- Zocchi, A., Gieles, M., Hénault-Brunet, V., & Varri, A. L. 2016, *MNRAS*, **462**, 696
- Zocchi, A., Gieles, M., & Hénault-Brunet, V. 2019, *MNRAS*, **482**, 4713

Appendix A: Deprojection of the Sérsic surface density profile

The Sérsic surface density model used in this work is deprojected in the same way as Vitral & Mamon (2020, hereafter VM20), but we extend the deprojection to lower radii: $10^{-4} R_e$ instead of $10^{-3} R_e$. We consider again the approximations of Lima Neto et al. (1999, hereafter LGM99), Simonneau & Prada (2004, hereafter SP04), as well as the formula of VM20, to find which is better suited for each region of the $[n \times r/R_e]$ domain.

A.1. New coefficients for VM20 approximation extending to very low radii

The coefficients of the VM20 approximation were originally calculated for a logarithmic grid of $[n \times r/R_e]$ with $-3 \leq \log(r/R_e) \leq 3$ (100 steps) and $0.5 \leq n \leq 10$ (50 steps). However, our HST data could extend to even lower radii than $0.001 R_e$, if $R_e > 7'$ (the lowest projected radius is $0'.42$). Indeed, such large effective radii were found for the faint stars of the two-population models 14 and 15. We therefore recomputed the VM20 approximation, using a different region of the $[n \times r/R_e]$ domain, and an even finer grid.

We used the same approach as we did in VM20, but this time we performed the numerical deprojection of the Sérsic profile with MATHEMATICA 12, instead of PYTHON. Given Fig. 4 of

VM20 and the lower limit of r/R_e we needed to attain, we calculated the best VM20 parameters for a new logarithmic spaced region $[n \times r/R_e]$ limited to $-4 \leq \log(r/R_e) \leq 3$ (150 steps) and $0.5 \leq n \leq 3.5$ (100 steps). The resulting coefficients $a_{i,j}$, presented in the same way as VM20 (see their Eq. (28)), can be found online¹⁹ and we hereafter refer to them as VM20bis.

A.2. Choice of deprojection approximation

The two left panels of Fig. A.1 display the best fitting approximations in 100×150 grid in $[\log n \times \log(r/R_e)]$. We included in MAMPOSST-PM a simplified choice of best approximations to the deprojected Sérsic mass and density profiles. For the mass profile we used

- $n \times r/R_e \in [0.5; 1.5] \times [1; 10^3]$: LGM99
 - $n \times r/R_e \in [0.5; 3.4] \times [10^{-4}; 1]$: VM20bis
 - $n \times r/R_e \in (3.4; 10] \times [10^{-4}; 1] \cup (1.5; 10] \times [1; 10^3]$: SP04
- For the density profile, the division was even simpler:

- $n \times r/R_e \in (3.4; 10] \times [10^{-4}; 10^3]$: SP04
- $n \times r/R_e \in [0.5; 3.4] \times [10^{-4}; 10^3]$: VM20bis

The outcome of this approximation is highly accurate, as seen in the right panels of Fig. A.1 (which are the equivalent of Fig. 3 in VM20), with the grid lines indicating the divisions adopted for the 3 approximations. The reader can verify that the relative precision of this hybrid model is on the order of $\sim 0.1\%$.

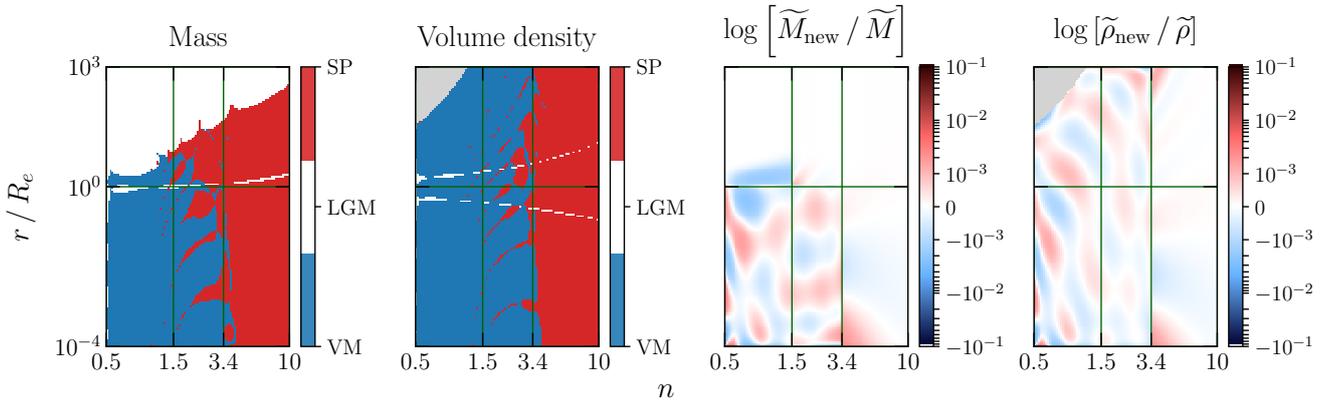


Fig. A.1. Characteristics of approximations to the mass and density profiles of the deprojected Sérsic model. *Left two panels:* most precise approximation. SP stands for Simonneau & Prada (2004), LGM stands for Lima Neto et al. (1999) and VM stands for the new VM20bis coefficients applied to the Vitral & Mamon (2020) method. The white curves indicate a thin region preferred by the LGM approximation. *Right two panels:* accuracy of deprojected mass (*left*) and density (*right*) of the hybrid model using VM20bis coefficients, LGM99 and SP04, with respect to the numerical integration done with MATHEMATICA. This is the analog of Fig. 3 of VM20: the color scale is linear for log ratios between -0.001 and 0.001 and logarithmic beyond. Both sets of figures employ a $[100 \times 150]$ grid of $[\log n \times \log(r/R_e)]$, which is shown in all four panels. The gray region in the upper left of each of the density panels is for regions where the numerical integration of MATHEMATICA reached the underflow limit because of the very rapid decline of density at large radii for low n .

¹⁹ https://gitlab.com/eduardo-vital/vital_mamon_2020b

Appendix B: Handling of field stars

B.1. Proper motion distribution function

In this appendix, we justify the need for using wider tails than a Gaussian for the distribution of PMs in the field stars, as mentioned in Sect. 5.2.3.

We downloaded *Gaia* DR2 PM data in four regions around NGC 6397 (and also for two other GCs, M4 and NGC 6752, in order to check for generality). These four regions were chosen by doing a cone search, with a 30' aperture, for positions 5° distant from the GC centers (α_{GC} , δ_{GC}), north, south, east, and west. We applied the same quality flags that we had applied to the NGC 6397 data (Sect. 5.2.1): the inequalities of Eqs. (12), (16) and (17).

We then fitted the distribution of PM moduli using both a Gaussian and the form of Eq. (18). We estimated the mean $\mu_{\alpha,*}$ and mean μ_{δ} for both distribution functions, a dispersion σ for the Gaussian assumption and for Eq. (18), we estimated the scale radius and outer slope. Finally, we took into account the convolution of both distributions with Gaussian errors.

Figure B.1 shows the distribution of PM moduli, for the four regions NGC 6397. One can easily verify that the new expression fits extremely better than a Gaussian in Fig. B.1. Moreover, the calculated kurtosis excess of both $\mu_{\alpha,*}$ and μ_{δ} always gave huge values (from 16 to over 400, compared to 0 expected for a Gaussian), which clearly implies a non-Gaussian behavior.

To check for robustness of our fits, we also verified that whenever limiting the fitting range to exclude the wider wings of the FS distribution (which cannot be done when considering a GC relatively separated from the FS in PM space), a Gaussian distribution is well fitted. Thus, we decided to employ Eq. (18) throughout our study in order to account for the entire PM range.

B.2. Convolution of field stars distribution

For the GC component, the local Gaussian-like velocity distribution function (VDF) must be convolved by the Gaussian LOS and POS velocity errors, as in Eq. (7). Similarly, since the field star LOS VDF is a Gaussian, the LOS velocity errors are added in quadrature to the LOS velocity dispersion.

On the other hand, since the interloper PM distribution function (Eq. (18)) is not Gaussian-like shaped, we need to perform the integral of its convolution. Calling R and R_0 the respective true and observed PMs and ϵ the PM error, it is straightforward to follow the recipe of Binney & Mamon (1982, Appendix C) for the convolution of two-dimensional data with circular symmetry:

$$p_{\text{conv}}(R_0) = \frac{R_0}{\epsilon^2} \int_0^{\infty} dR p(R) \exp\left(-\frac{R^2 + R_0^2}{2\epsilon^2}\right) I_0\left(\frac{RR_0}{\epsilon^2}\right). \quad (\text{B.1})$$

Mamon & Vitral (in prep.) provide an approximation for this integral.

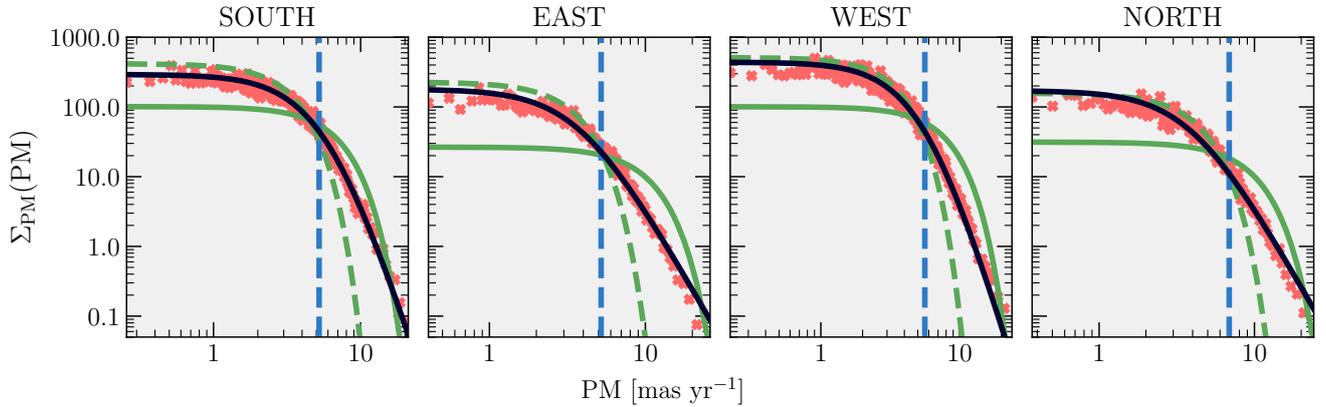


Fig. B.1. Surface density of proper motion moduli (defined in Eq. (19)) for the four 5° distant regions around NGC 6397 (for simplicity, we call them SOUTH, EAST, WEST and NORTH), represented by red crosses. Solid green curves display the best MLE fit for a Gaussian distribution, while dashed green curves display the best Gaussian MLE fit when only considering the regions with proper motions inside the limit of the dashed blue vertical line. The best MLE fit using Eq. (18) is displayed as the black curves.