

The universal distribution of halo interlopers in projected phase space. Bias in galaxy cluster concentration and velocity anisotropy?

G. A. Mamon, A. Biviano, G. Murante

▶ **To cite this version:**

G. A. Mamon, A. Biviano, G. Murante. The universal distribution of halo interlopers in projected phase space. Bias in galaxy cluster concentration and velocity anisotropy?. Astronomy and Astrophysics - A&A, 2010, 520, 10.1051/0004-6361/200913948 . hal-03646081

HAL Id: hal-03646081
https://hal.science/hal-03646081

Submitted on 18 Jul 2022

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

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

A&A 520, A30 (2010)
DOI: 10.1051/0004-6361/200913948
© ESO 2010

Astronomy & Astrophysics

The universal distribution of halo interlopers in projected phase space

Bias in galaxy cluster concentration and velocity anisotropy?⋆

G. A. Mamon[1,2], A. Biviano[3], and G. Murante[4]

[1] Institut d'Astrophysique de Paris (UMR 7095: CNRS & UPMC), 98 bis Bd. Arago, 75014 Paris, France
e-mail: gam@iap.fr
[2] Astrophysics & BIPAC, University of Oxford, Keble Rd, Oxford OX13RH, UK
[3] INAF, Osservatorio Astronomico di Trieste, Trieste, Italy
[4] INAF, Osservatorio Astronomico di Torino, Torino, Italy

Received 22 December 2009 / Accepted 1 July 2010

ABSTRACT

When clusters of galaxies are viewed in projection, one cannot avoid picking up a fraction of foreground/background interlopers, that lie within the virial cone, but outside the virial sphere. Structural and kinematic deprojection equations are known for the academic case of a static Universe, but not for the real case of an expanding Universe, where the Hubble flow (HF) stretches the line-of-sight distribution of velocities. Using 93 mock relaxed clusters, built from the dark matter (DM) particles of a hydrodynamical cosmological simulation, we quantify the distribution of interlopers in projected phase space (PPS), as well as the biases in the radial and kinematical structure of clusters produced by the HF. The stacked mock clusters are well fit by an $m = 5$ Einasto DM density profile (but only out to 1.5 virial radii), with velocity anisotropy (VA) close to the Mamon-Łokas model with characteristic radius equal to that of density slope −2. The surface density of interlopers is nearly flat out to the virial radius, while their velocity distribution shows a dominant Gaussian cluster-outskirts component and a flat field component. This distribution of interlopers in PPS is nearly universal, showing only small trends with cluster mass, and is quantified. A local $\kappa = 2.7$ sigma velocity cut is found to return the line-of-sight velocity dispersion profile (LOSVDP) expected from the NFW density and VA profiles measured in three dimensions. The HF causes a shallower outer LOSVDP that cannot be well matched by the Einasto model for any value of κ. After this velocity cut, which removes 1 interloper out of 6, interlopers still account for 23 ± 1% of all DM particles with projected radii within the virial radius (surprisingly very similar to the observed fraction of cluster galaxies lying off the Red Sequence) and over 60% between 0.8 and 1 virial radius. The HF causes the best-fit projected NFW or $m = 5$ Einasto model to the stacked cluster to underestimate the true concentration measured in 3D by 6 ± 6% (16 ± 7%) after (before) the velocity cut. These biases in concentration are reduced by over a factor two once a constant background is included in the fit. The VA profile recovered from the measured LOSVDP by assuming the correct mass profile recovers fairly well the VA measured in 3D, with a slight, marginally significant, bias towards more radial orbits in the outer regions. These small biases in the concentration and VA of the galaxy system are overshadowed by important cluster-to-cluster fluctuations caused by cosmic variance and by the strong inefficiency caused by the limited numbers of observed galaxies in clusters. An appendix provides an analytical approximation to the surface density, projected mass and tangential shear profiles of the Einasto model. Another derives the expressions for the surface density and mass profiles of the NFW model projected on the sphere (for future kinematic modeling).

Key words. galaxies: clusters: general – cosmology: miscellaneous – dark matter – galaxies: halos – gravitational lensing: weak – methods: numerical

1. Introduction

The galaxy number density profiles of groups and clusters of galaxies falls off slowly enough at large radii that material beyond the virial radius (within which these structures are thought to be in dynamical equilibrium) contribute non-negligibly to the projected view of cluster, i.e. to the radial profiles of surface density, line-of-sight velocity dispersion and higher velocity moments.

In principle, this contamination of observables by *interlopers*, defined here as particles that lie within the virial cone but outside the virial sphere, is not a problem, since we know how to express deprojection equations when interlopers extend to infinity along the line-of-sight. Consider the projection equation

$$\Sigma(R) = \int_{-\infty}^{+\infty} \nu(r)\,\mathrm{d}s = 2\int_R^\infty \nu(r)\frac{r\,\mathrm{d}r}{\sqrt{r^2-R^2}}, \quad (1)$$

where Σ and ν are the projected and space number densities, respectively, while R and r are the projected and space radial distances (hereafter, radii), respectively. Equation (1) can be deprojected through Abel inversion[1] to yield

$$\nu(r) = -\frac{1}{\pi}\int_r^\infty \frac{\mathrm{d}\Sigma/\mathrm{d}R}{\sqrt{R^2-r^2}}\,\mathrm{d}R. \quad (2)$$

The projection to infinity is explicit in Eqs. (1) and (2).

⋆ Appendices are only available in electronic form at http://www.aanda.org

[1] Alternatively, the projection Eq. (1) corresponds to a convolution and can therefore be deprojected with Fourier methods (see Discussion in Mamon & Boué 2010, and references therein).

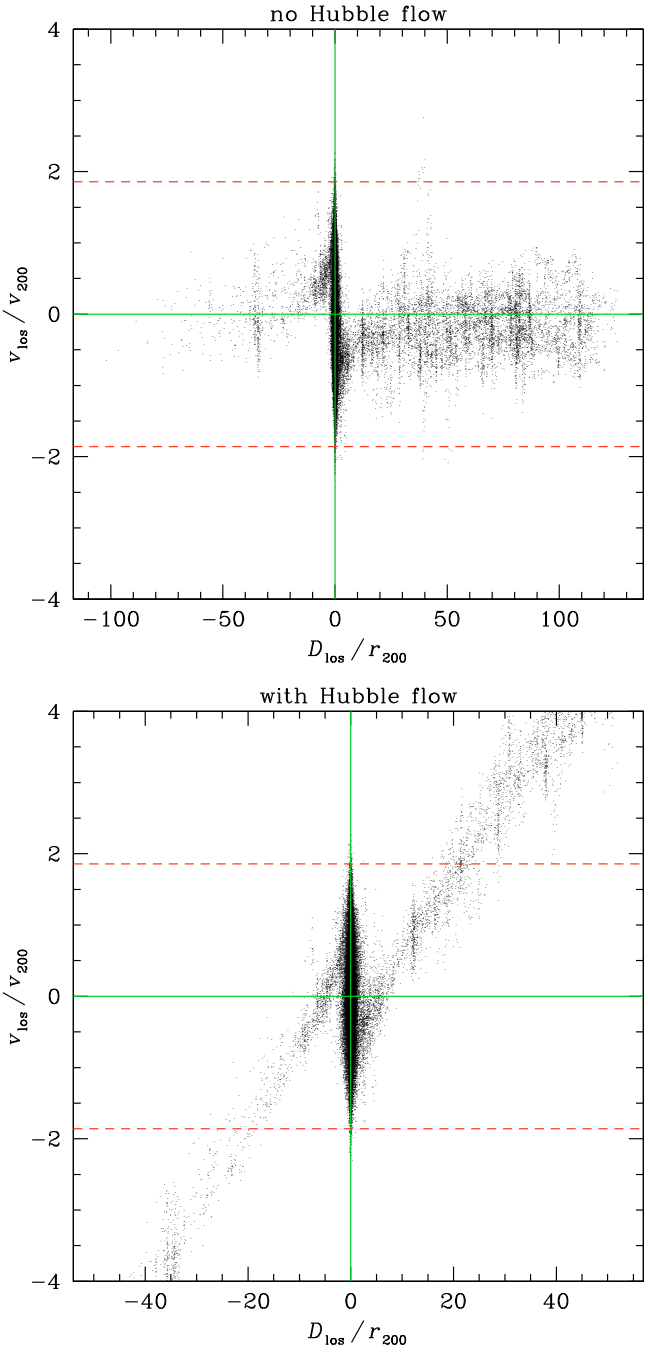


Fig. 1. Line-of-sight velocity as a function of real-space line-of-sight distance (see Fig. 2) for particles inside the virial cone obtained by stacking 93 cluster-mass halos in the cosmological simulation described in Sect. 2 without (*top*) and with (*bottom*) the Hubble flow (1 particle in 5 is shown for clarity). The *red dashed horizontal lines* roughly indicate the effects of a radius-independent 3σ clipping. Note that the velocity-distance relation without Hubble flow shown here is not entirely realistic, because the simulation was run in the context of an expanding Universe (and cannot be run in a static Universe, for lack of knowledge of realistic initial conditions), but should be accurate enough to illustrate our point.

However, the Hubble expansion complicates the picture, as the Hubble flow moves background (foreground) objects to high positive (negative) line-of-sight velocities. This is illustrated in Fig. 1 which shows how the line-of-sight velocity vs. real-space distance relation is affected by the Hubble flow. The line-of-sight

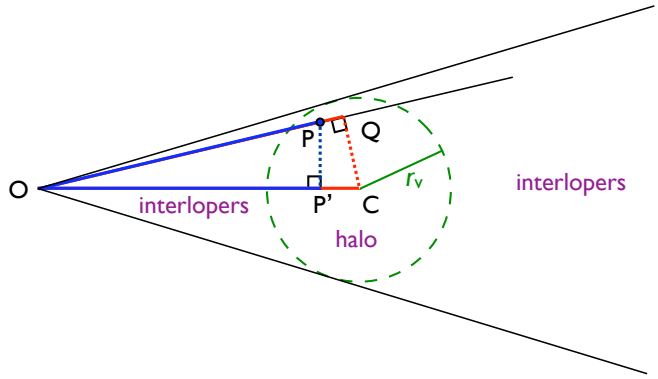


Fig. 2. Representation of the virial cone with halo particles inside the inscribed virial sphere and interlopers outside. Also shown is our definition of projected radius (CQ) and line-of-sight distance (OP and QP respectively in the observer and halo reference frames) for a random point P. For illustrative purposes, the distance to the cone is taken to be very small, so that the cone opening angle is much larger than in reality.

distances are computed as the segment length QP in Fig. 2. Now, clipping the velocity differences to, say, $\kappa = 3$ times the cluster velocity dispersion (averaged over a circular aperture, hereafter aperture velocity dispersion), σ_v, gets rid of all the distant interlopers. More precisely, the radius, $r_{\rm max}$, where the Hubble flow matches $\kappa\sigma_v$ is found by solving $H_0\, r_{\rm max} = \kappa\sigma_v$ (where H_0 is the Hubble constant) yielding

$$\frac{r_{\rm max}}{r_v} = \kappa \sqrt{\frac{\Delta}{2}} \left(\frac{\sigma_v}{v_v}\right), \tag{3}$$

where r_v is the virial radius where the mean density is Δ times the critical density of the Universe, $\rho_{\rm c} = 3H_0^2/(8\pi G)$ (where G is the gravitational constant), and where $v_v = \sqrt{\Delta/2}\, H_0\, r_v$ is the circular velocity at the virial radius. Clusters are thought to have density profiles consistent with the Navarro et al. profile (1996, hereafter NFW) profiles,

$$\nu(r) = \frac{1/\left(\ln 2 - 1/2\right)}{(r/r_{-2})\,(r/r_{-2}+1)^2} \left[\frac{M\,(r_{-2})}{4\,\pi\, r_{-2}^3}\right], \tag{4}$$

where r_{-2} is the radius of density slope -2, in number (Lin et al. 2004), luminosity (Łokas & Mamon 2003) and mass (Łokas & Mamon; Biviano & Girardi 2003; Katgert et al. 2004), with a concentration, $c = r_v/r_{-2}$, of 3 to 5. For isotropic NFW models, the aperture velocity dispersion is (Appendix A of Mauduit & Mamon 2007) $\sigma_v = \eta\, v_v$, where $\eta \simeq 0.62$ (weakly dependent on concentration), and Eq. (3) then becomes

$$\frac{r_{\rm max}}{r_v} \simeq 13.2 \left(\frac{\kappa}{3}\right) \sqrt{\frac{\Delta}{100}}\,. \tag{5}$$

Equation (5) indicates that a 3-sigma clipping will remove all material beyond 13 ($\Delta = 100$) or 19 ($\Delta = 200$) virial radii[2]. So, in the deprojection Eq. (2), the upper integration limit must be set to this value of $r_{\rm max}$. Although this effective cutoff in line-of-sight distances is quite far removed from the cluster, it is not

[2] With our chosen cosmology, the overdensity at the virial radius is $\Delta \simeq 100$, but many authors prefer to work with $\Delta = 200$, and we will do so too.

clear whether there may still be a measurable bias in the concentration of clusters that one measures by comparing the surface density distribution of galaxies in clusters with NFW models projected out to infinity. Moreover, it is not clear how accurate are such measures given the finite number of galaxies observed within clusters.

Finally, it is not clear how the stretching of the velocities affects the kinematic analyses of clusters, especially in the case of nearby clusters where the opening angle of the cone is non-negligible, leading to an asymmetry between the foreground and background absolute velocity distributions. For example, is the anisotropy of the 3D velocity distribution (hereafter velocity anisotropy or simply anisotropy)

$$\beta(r) = 1 - \frac{1}{2}\,\frac{\left\langle v_\theta^2(r) + v_\phi^2(r)\right\rangle}{\langle v_r^2(r)\rangle} \tag{6}$$

or equivalently[3]

$$\mathcal{A}(r) = [1-\beta(r)]^{-1/2} = \left[\frac{2\left\langle v_r^2(r)\right\rangle}{\left\langle v_\theta^2(r) + v_\phi^2(r)\right\rangle}\right]^{1/2} \tag{7}$$

affected by the Hubble flow? One can add interlopers beyond the virial radius as a separate component to the kinematical modeling (van der Marel et al. 2000; Wojtak et al. 2007). Unfortunately, we have no knowledge of the distribution of interlopers in projected phase space (projected distance to the halo center and line-of-sight velocity).

This paper provides the distribution of interlopers in projected phase space (projected distance to the halo center and line-of-sight velocity) as measured on nearly 100 stacked halos from a well-resolved cosmological simulation. We additionally measure the bias in the measured surface density and line-of-sight velocity dispersion and kurtosis profiles compared to those obtained in a Universe with no Hubble flow, and estimate how this bias affects the recovered concentration and velocity anisotropy of the cluster. In this paper, we use interchangeably the terms "clusters" and "halos".

We present in Sect. 2 the cosmological simulations we use and how the individual halos were built. In Sect. 3 we explain how we stack these halos. In Sect. 4, we present the statistics on the halo members and the interlopers in projected phase space. Then in Sect. 5, we explain how we remove the outer interlopers and show analogous statistics on the cleaned stacked halo in Sect. 6. We proceed in Sect. 7 to measure the biases induced by the Hubble flow and the imperfect interloper removal on the estimated concentration parameter and anisotropy profile. We discuss our results in Sect. 8.

2. Data from cosmological *N*-body simulations

The halos analyzed in this paper were extracted by Borgani et al. (2004) from their large cosmological hydrodynamical simulation performed using the parallel Tree+SPH GADGET-2 code (Springel 2005). The simulation assumes a cosmological model with present day parameters $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_b = 0.039$, $h = H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}) = 0.7$, and $\sigma_8 = 0.8$. The box size is $L = 192\,h^{-1}$ Mpc. The simulation used 480^3 dark matter particles and (initially) as many gas particles, for a dark matter particle mass of $4.62 \times 10^9\,h^{-1}\,M_\odot$. The softening length was set to $22.5\,h^{-1}$ comoving kpc until $z = 2$ and fixed afterwards (i.e., $7.5\,h^{-1}$ kpc). The simulation code includes explicit energy and entropy conservation, radiative cooling, a uniform time-dependent UV background (Haardt & Madau 1996), the self-regulated hybrid multi-phase model for star formation (Springel & Hernquist 2003), and a phenomenological model for galactic winds powered by type-II supernovae.

Dark matter halos were identified by Borgani et al. at redshift $z = 0$ by applying a standard Friends-of-friends (FoF) analysis to the dark matter particle set, with linking length 0.15 times the mean inter-particle distance. After the FoF identification, the center of the halo was set to the position of its most bound particle. A spherical overdensity criterion was then applied to determine the virial radius, $r_v = r_{200}$ of each halo. In this manner, 117 halos were identified within the simulated volume, among which 105 form a complete subsample with virial mass M_{200} larger than $10^{14}\,h^{-1}\,M_\odot$, thus representing a sample of mock galaxy clusters. Their mean and maximum masses are respectively $2.0 \times 10^{14}\,h^{-1}\,M_\odot$ and $1.1 \times 10^{15}\,h^{-1}\,M_\odot$.

To save computing time, we worked on a random subsample of roughly 2 million particles among the 480^3. Although the simulation also produced galaxies, we chose to use the dark matter particles as tracers of the galaxy distribution for two reasons: 1) simulated galaxy properties in cosmological simulations depend on details of the baryon physics implemented in the code, and can show some mismatch with observed properties (e.g. Saro et al. 2006); 2) only a handful of simulated clusters had over 50 galaxies (Saro et al.), so we would have strongly suffered from small-number statistics. There is some debate on whether the velocity distribution of galaxies is biased relative to the dark matter. On one hand, the galaxy velocity distribution, although close to the dark matter one, shows a preference for lower velocities (Biviano et al. 2006), perhaps as a consequence of dynamical friction. On the other hand, the velocity distribution of subhalos, selected with a minimum mass before entering their parent cluster-mass halos, is similar to that of the dark matter (Faltenbacher & Diemand 2006).

We visually inspected each of the 105 clusters in redshift space along three orthogonal viewing axes, and removed 12 clusters that appeared, within r_{200}, to be composed of two or three sub-clusters of similar mass (where the secondary had at least 40% of the mass of the primary). Most observers would omit such clusters when analyzing their radial structure or internal kinematics. This leaves us with 93 final mock clusters[4]. The median values (interquartile uncertainties) of their virial radii, virial masses, virial circular velocities, and velocity dispersions (within their virial spheres) are respectively $864 \pm 81\,h^{-1}$ kpc, $1.50 \pm 0.44\,10^{14}\,h^{-1}\,M_\odot$, 865 ± 81 km s^{-1}, and 584 ± 60 km s^{-1}.

3. Stacking the virial cones

For each cluster, we projected the coordinates along the virial cone (circumscribing the virial sphere, see Fig. 2), as follows. We first renormalized the 6 coordinates of phase space of the entire simulation box to be relative to the cluster. So the cluster most bound particle should be at the origin and its mean peculiar (bulk) velocity should be zero[5]. To take into account the periodic boundaries of the simulation box, we added or subtracted a box

[3] $\beta(r)$ enters the Jeans equation of local hydrostatic equilibrium, while $\mathcal{A}(r)$ is a more physical definition of velocity anisotropy.

[4] Including all 105 halos makes virtually no difference for the results of this article.

[5] Observers usually adopt the position of the brightest cluster galaxy as the center, and this corresponds to the most bound galaxy, so they should not suffer from important centering errors, although admittedly some clusters like Coma have two brightest galaxies.

length to those particles situated at over a half-box length from the cluster center. In this fashion, each cluster now effectively sits at the center of the simulation box. We then placed an observer at coordinates $(-D, 0, 0)$, $(0, -D, 0)$, or $(0, 0, -D)$, with $0 < D < L/2$. We present here the results for $D = 90\,h^{-1}$ Mpc, corresponding to a typical distance of observed clusters in the local Universe. At this distance, the median virial angular radius of our 93 clusters is 33 ± 3 arcmin. We do not expect that the results of this paper should depend on the adopted value of D. We assume that the observer's peculiar velocity is equal to the cluster's bulk velocity, so that the observer's velocity is zero in the renormalized coordinate system[6].

We then measured, for each cluster and for each of these three observers, the coordinates of all 2 million particles in both the observer frame and the cluster frame. Given the distance r_o of the particle to the observer and the projected coordinate of the particle in cylindrical coordinates R_c, we determined the projected distances in the observer frame (measured in a plane perpendicular to the line-of-sight passing through the cluster center) as $R_o = D\,R_c/r_o$ (see Fig. 2, where D = OC, r_O = OP, R_c = PP′, and R_o = CQ). This projection ensures that particles along the surface of the virial cone have $R = r_v$. We were then able to select all particles within a cone circumscribing the sphere of radius r_v, where we chose $r_v = r_{200}$, as well as $r_v = 1.35\,r_{200} \simeq r_{100}$. In practice, we extracted data from a wider cone, circumscribing the sphere of radius $3\,r_{200}$.

We next added the Hubble flow (for both the observer and cluster frames), using $H_0 = 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [7]. We then limited in depth to line-of-sight velocities within $4\,v_v$ from the cluster. As mentioned in Sect. 1, the Hubble flow effectively limits the depth of the cones to a half-length of $\hat{\kappa}\,\sqrt{\Delta/2} = 40$ virial radii (see Eq. (3)), where the cut in velocity space in units of virial velocity is $\hat{\kappa} = 4$ and the virial overdensity is $\Delta = 200$.[8] Our results should not depend much on the distance of our observer ($90/0.864 = 104$ virial radii), as our first cut at ± 4 virial velocities limits the line of sight to 40% of the observer's distance.

We finally normalized the particle relative positions and velocities to the virial radius, r_v and circular velocity at the virial radius, v_v, respectively, and finally computed the projections of the velocities in spherical coordinates (to later measure the 3D radial profiles of density and velocity anisotropy). For clarity, we will sometimes use the notation r_{200} for the virial radius and v_{200} for the virial velocity.

In the end, we have roughly 84 thousand particles for each of the three *cartesian* stacked virial cones, which we then stack together into our *global* stacked virial cone, with a grand total of roughly a quarter-million particles, among which nearly three-quarters lie within the virial radius. Hence, roughly 1/30th of all the particles in the simulation box lie within the virial radius, r_{200}, of our 93 clusters. Note that with our stacking method, some of the particles inside virial cones but outside virial radii can end up being selected in more than one of the three cartesian stacked virial cones. However, this fraction is small (27% of the interlopers, i.e. less than 8% of all particles in virial cones), so the three cartesian stacked virial cones are virtually independent (except that their halos are common), hence can be stacked into the global virial cone.

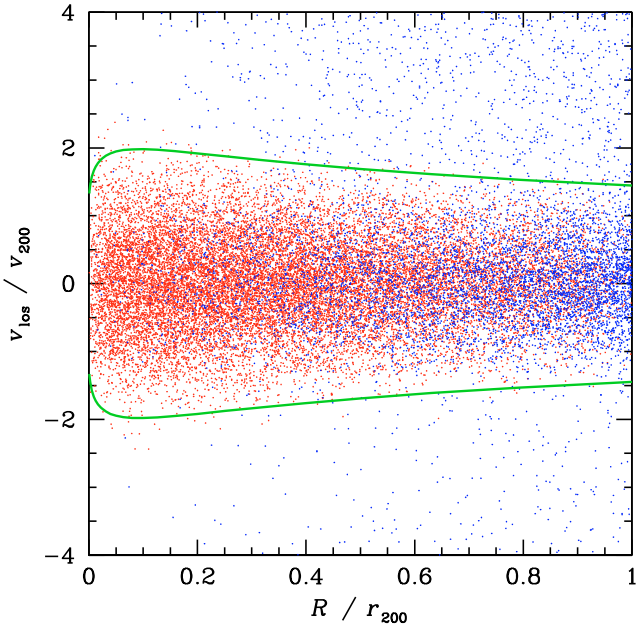


Fig. 3. Projected phase space diagram of stacked virial cone (built from 3×93 halos). Only 1 particle in 5 is shown for clarity. The *red* (*lighter*) and *blue* (*darker*) points refer to the particles within and outside the virial sphere, respectively. The *green curves* illustrate the $\pm 2.7\sigma_{los}(R)$ velocity cut (from Eqs. (21) and (22)) for the $c \equiv c_{200} = 4$ NFW model with $r_a = r_{-2}$ ML anisotropy (Eq. (20)).

Figure 3 shows the projected phase space (line-of-sight velocity, v_{los}, versus projected radius, R) distribution of particles of the stacked cluster, highlighting the *halo particles* (lying within the virial sphere) and *interlopers* (lying in the virial cone but outside of the virial sphere). Note that the halo particles are confined to fairly small velocities (the largest absolute halo particle velocity is $2.7\,v_v$). One notices an excess of positive velocity outliers in comparison with negative velocity ones, as expected from the conical projection used here.

4. Interloper statistics before the velocity cut

We now measure the distribution of interlopers in projected phase space and study its dependence on halo mass.

4.1. Global statistics

Figure 4 displays contours of the density in $(R, |v_{los}|)$ projected phase space. Note that the projected phase space density of Fig. 4 is proportional to $\mathrm{d}^2N/(R\,\mathrm{d}R\,\mathrm{d}|v_{los}|)$, hence the different shapes than seen in Fig. 3. The interlopers have a very different projected phase space density than the halo particles. In particular, their horizontal contours mean that the interloper projected phase space density is fairly independent of projected radius. Moreover, the interloper contours do not extend beyond $v_{los} = 2\,v_v$, except for a few islands (caused by cosmic variance), indicating that the velocity distribution of interlopers is close to flat at large velocities (the islands thus represent small, probably statistical, fluctuations in a flat background).

These issues can be looked in more detail through slices of the projected phase space density in velocity and radial space.

[6] In our simulation, the one-dimensional cluster bulk velocity dispersion is $228\,\mathrm{km\,s^{-1}}$, so given our adopted distance of $D = 90\,h^{-1}$ Mpc, the typical cluster bulk velocity is only $228/9000 = 2.5\%$. Therefore, our neglect of the observer's peculiar motion relative to the cluster is an adequate assumption.

[7] All our positions and cluster virial radii were expressed in h^{-1} kpc; the choice of H_0 does not matter as long as we normalize to the virial quantities r_v and v_v, which we computed with the same value of H_0.

[8] Throughout this paper, we use $\hat{x}$ to express quantity x in virial units.

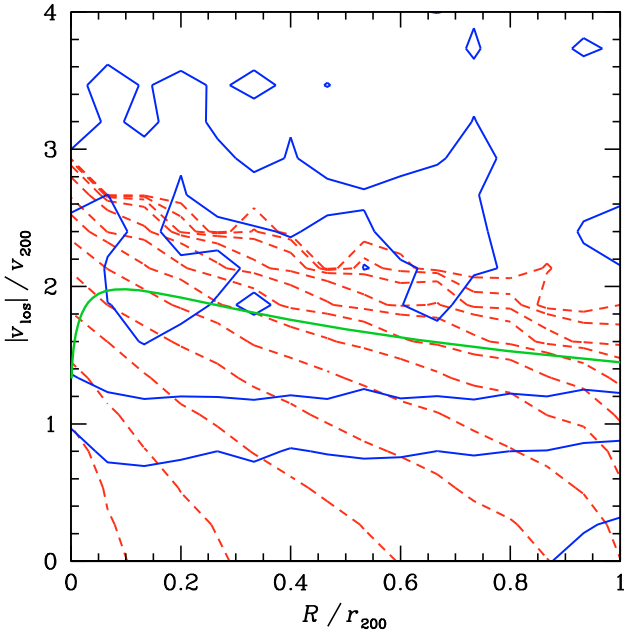


Fig. 4. Contours of projected phase space density of stacked virial cone – in units of $d^2N/(dR/r_{200}\, d|v_{los}|/v_{200})/(2\pi R/r_{200})$ – of halo (*dashed red contours*) and interloper (*solid blue*) particles. The halo contours are equally spaced in log-density, increasing by a factor 2.44 from 0.085 (*upper right*) to 3370 (*lower left*). The interloper contours are taken from the same set as the halo contours, but limited from the 5th highest level halo contour (72 virial units, at the *lower left*) to the 8th highest level halo contour (4.0 virial units, for all contours above the 3rd nearly horizontal one). The *green curve* shows the $2.7\sigma_{los}(R)$ (from Eqs. (21) and (22)) velocity cut for the $c = 4$ NFW model with $r_a = r_{-2}$ ML anisotropy (Eq. (20)).

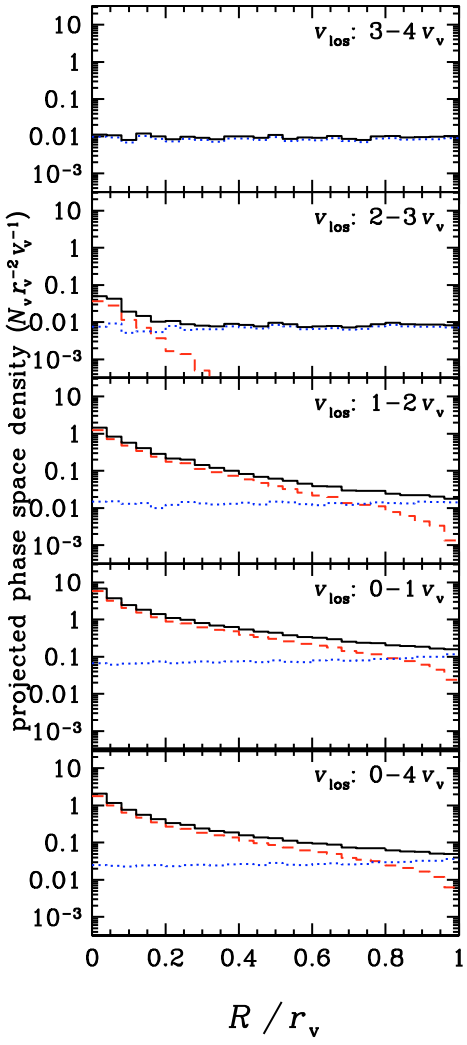


Fig. 5. Phase space density of stacked virial cone as a function of projected radius in bins of absolute line-of-sight velocities (marked on the *upper-right* of each plot). *Dashed* (*red*) and *dotted* (*blue*) *histograms* represent the halo particles ($r < r_v$) and the interlopers ($r > r_v$), respectively, while the *solid histograms* (artificially moved up by 0.06 dex for clarity) represent the full set of particles. There are no halo particles at $v > 3\, v_v$ (*top plot*).

Figure 5 shows how the projected phase space density varies with projected radius in different wide velocity bins. The halo particles display a negative gradient, i.e. a decreasing surface number density profile, as expected. However, one immediately notices that *in all line-of-sight velocity bins, the density of interlopers in projected phase space is roughly independent of projected radius*. In other words, interlopers have a nearly flat surface density profile.

For low velocities, one can notice a small rise of the interloper surface density at high projected radii. This small rise is a geometric effect: the line-of-sight distance between the virial sphere and a sphere of $k > 1$ virial radii is $k - 1$ virial radii at $R = 0$ but $\sqrt{k^2 - 1}$ virial radii at $R = r_v$, which is $\sqrt{(k+1)/(k-1)}$ times greater. We will return to this rise in Sect. 4.2.

Figure 6 displays the distribution of line-of-sight velocities of the halo and interloping particles. The interloper distribution shows a flat component that dominates at large velocities and a Gaussian-like component. Figure 6 confirms, once more (see Figs. 4 and especially 5) that the density of interlopers in projected phase space is fairly independent of radius.

The total surface density of particles in velocity space shows an inflection point at about $2\, v_v$ (bottom plot of Fig. 6). Kinematical modelers attempt to throw out the high-velocity interlopers by identifying this gap by eye (Kent & Gunn 1982; Łokas & Mamon 2003) or automatically (Fadda et al. 1996) or by rejecting 3σ outliers, either using a global criterion (Yahil & Vidal 1977) or a local one (e.g. Łokas et al. 2006; Wojtak & Łokas 2010). Interestingly, the $3\sigma_v$ criterion was first motivated on statistical grounds, but the $v \simeq 2.0\, v_v$ inflection point one sees in the plots of Fig. 6 happens to correspond to $\simeq 3\sigma_v$. In other words, *the $3\sigma_v$ criterion is not only a consequence of statistics, but also motivated by the combination of cluster dynamics and cosmology.*

While visual attempts to separate interlopers from halo particles in projected phase space look for gaps in the line-of-sight velocity distribution, the different panels of Fig. 6 indicate that, on average, one should not expect such gaps in the projected phase space diagram of *stacked* clusters, as the number of interlopers also decreases with velocity to reach a plateau at about $2\, v_v$.

4.2. Universality

The three panels of Fig. 7 illustrate the universality of the line-of-sight velocity distribution of interlopers, in terms of projected

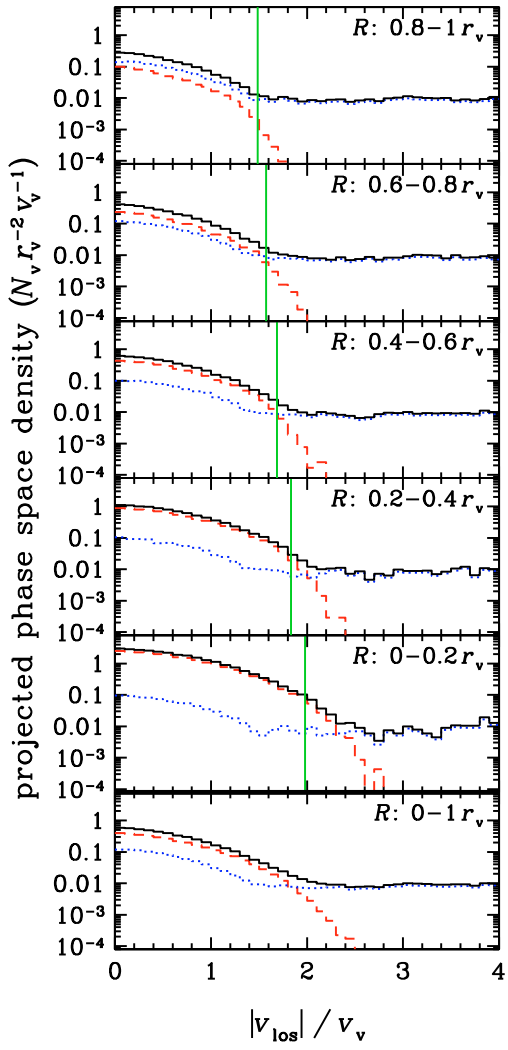


Fig. 6. Phase space density of stacked virial cone as a function of line-of-sight velocity in different radial bins (marked on the lower-left of each plot). *Dashed (red)* and *dotted (blue) histograms* represent the halo particles ($r < r_v$) and the interlopers ($r > r_v$), respectively, while the *solid histograms* (artificially moved up by 0.06 dex for clarity) represent the full set of particles. The *green vertical lines* indicate the $2.7\sigma_{los}(R)$ (from Eqs. (21) and (22)) velocity cut for the $c = 4$ NFW model with $r_a = r_{-2}$ ML anisotropy (Eq. (20)).

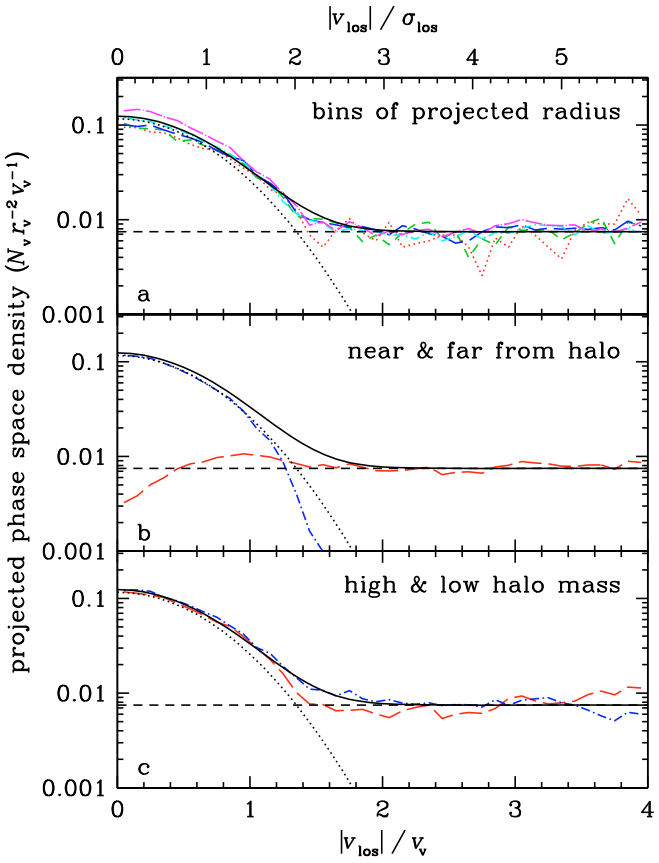


Fig. 7. Phase space density of interlopers as a function of line-of-sight velocity. **a**, *Top*): dependence on projected radius: $R/r_{200} = 0$–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1 (*red dotted*, *green short-dashed*, *blue long-dashed*, *magenta dot-long-dashed* and *cyan dot-long-dashed broken lines*, respectively). **b**, *Middle*): dependence on 3D radial distance: $1 < r/r_{200} < 8$ (*blue dash-dotted curve*) and $r/r_{200} > 8$ (*red dashed curve*). **c**, *Bottom*): dependence on halo mass: high ($h\,M_{200} > 1.87 \times 10^{14}\,M_\odot$, *red long dashed curve*) and low ($h\,M_{200} < 1.87 \times 10^{14}\,M_\odot$, *blue dash-dotted curve*) mass halos. Also shown in all three plots is the MLE (Eqs. (8) and (10)) for the Gaussian (*dotted curve*) and field (*horizontal dashed line*) components, as well as the predicted total interlopers (the sum of these two components, *solid curve*).

radius and halo mass. Figure 7a confirms that the projected phase space density of interlopers depends little on projected radius. This will be quantified later in this sub-section.

As mentioned above, to first order, the density of interlopers in projected phase space has a constant component and a quasi-Gaussian component, which we write

$$g\,(R, |v_{los}|) = A\,\exp\left[-\frac{1}{2}\left(\frac{|v_{los}|}{\sigma_i}\right)^2\right] + B. \quad (8)$$

Maximum likelihood estimation (MLE, see Appendix C) yields

$$\begin{aligned} \sigma_i &= 0.576 \pm 0.003, \\ A &= 0.1164 \pm 0.0006, \\ B &= 0.0075 \pm 0.0001, \end{aligned} \quad (9)$$

where σ_i is in units of v_v, while g, A and B are in units of $N_v\,r_v^{-2}\,v_v^{-1}$, where N_v is the number of particles within the virial sphere. Figures 7a and c show that this Gaussian+constant model is an excellent fit to the distribution of interloper line-of-sight velocities.

The origin of these two components is clarified in Fig. 7b: cutting the halo interlopers in two subsamples at different 3D distances from the halo, one finds that the flat component corresponds roughly to the interlopers beyond $8\,r_{200}$, while the quasi-Gaussian component corresponds to the closer interlopers ($r_{200} < r < 8\,r_{200}$).

Figure 8 shows the radial dependence of the parameters of the interloper phase space distribution. We also show the mean surface densities of interlopers measured on the stacked halo. The normalization of the Gaussian component and the surface density both increase slowly at small projected radii, but sharply near the virial radius. A good fit for the normalization and the standard deviation of the Gaussian component is provided by

$$A(R) \simeq \mathrm{dex}\left[-1.061 + 0.364\,X^2 - 0.580\,X^4 + 0.533\,X^6\right], \quad (10)$$

$$\sigma_i(R) \simeq 0.612 - 0.0653\,X^2, \quad (11)$$

where $X = R/R_{200}$ and $\mathrm{dex}\,x = 10^x$.

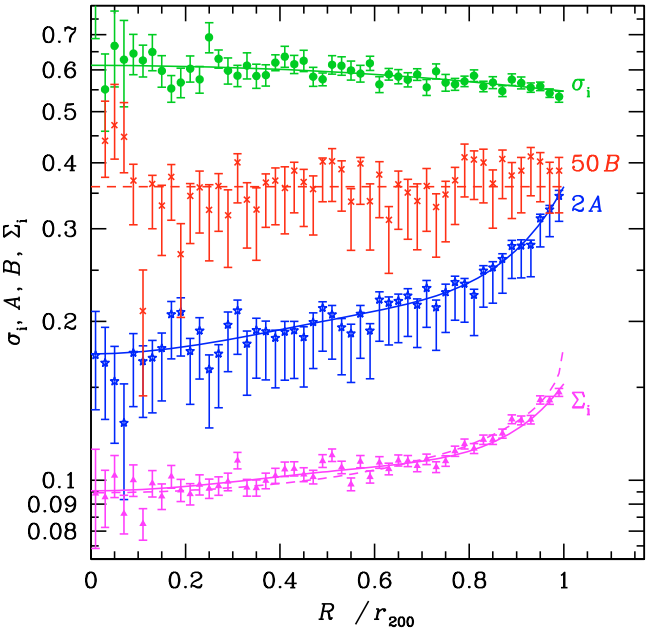


Fig. 8. Variations of MLE parameters of Eq. (8) and measured interloper surface density with projected radius (in units of v_v for σ_i, $N_v\, r_v^{-2}\, v_v^{-1}$ for A and $N_v\, r_v^{-2}$ for Σ_i). The *blue* and *green curves* show the fits of Eqs. (10) and (11), respectively, while the *red horizontal dashed line* shows the mean of 50 B. The *magenta solid curve* shows the prediction for the interloper surface density from Eqs. (12), (10) and (11), respectively, with $B = 0.0075$, while the *magenta dashed curve* shows the interloper mean surface density predicted (Eq. (13)) from a $c = 4$ NFW model. Small radial bins were chosen to capture the rise of Σ_i and A near the virial radius. Errors are from the likelihood ratios in the MLE fit (A, σ_i, and B) or Poisson (Σ_i).

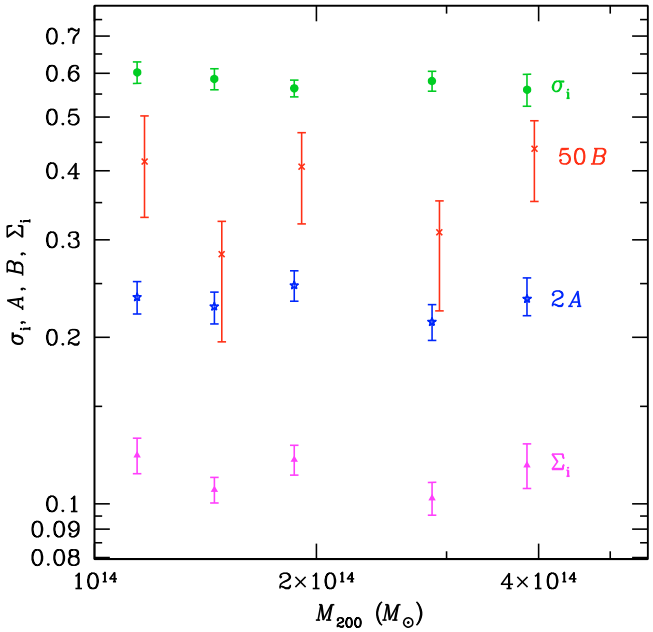


Fig. 9. Variations of MLE parameters of Eq. (8) with halo mass (in units of v_v for σ_i, $N_v\, r_v^{-2}$ for Σ_i and $N_v\, r_v^{-2}\, v_v^{-1}$ for A and B). Errors are from 100 bootstraps on the 93 halos. The symbols for 50 B are displaced by +0.01 dex for clarity.

The solid magenta curve of Fig. 8 indicates that the interloper surface density profile is well recovered from A, σ_i and B by integrating over the model velocity distribution (Eq. (8)) from 0 to $\hat{\kappa}\, v_v$:

$$\Sigma_i(R) = \sqrt{\pi/2}\, A(R)\, \sigma_i(R)\, \mathrm{erf}\left(\frac{\hat{\kappa}}{\sigma_i(R)\,\sqrt{2}}\right) + \hat{\kappa}\, B. \qquad (12)$$

Alternatively, the surface density profile of interlopers is also well recovered (dashed magenta curve in Fig. 8) by the prediction from an NFW model, i.e. as the difference between the standard surface density integrated to infinity and the surface density limited to the virial sphere:

$$\Sigma_i(R) = \Sigma_{\rm NFW}(R) - \Sigma^{\rm sph}_{\rm NFW}(R, R_{200}), \qquad (13)$$

where the expression for $\Sigma^{\rm sph}_{\rm NFW}(R, R_{200})$ is provided in Eq. (B.4). According to Eq. (13), the rise at radii close to the virial radius is almost independent of the concentration (with relative differences of less than 4% at all projected radii). Hence, one can adopt a unique empirical model for A. These two estimates of $\Sigma_i(R)$ are in very good agreement, except that the difference of the cylindrical and spherical NFW surface density profiles (Eq. (13)) rises somewhat faster than our fit at projected radii very close to the virial radius.

The bottom panel of Fig. 7 shows that the density of interlopers in projected phase space as a function of line-of-sight velocity remains the same for low and high mass clusters (where we took the dividing line at the median cluster virial mass of $h\, M_{200} = 1.87 \times 10^{14}\, M_\odot$). The interlopers of high mass clusters have a cluster-outskirts component whose density in projected phase space is roughly 15% lower than that of the low-mass clusters.

Figure 9 provides a closer look at the variation with halo mass of the parameters of Eq. (8). The best fitting logarithmic slopes, obtained by least-squares fits to the points shown in Fig. 9 are -0.042 ± 0.030, -0.037 ± 0.069, 0.155 ± 0.225 and -0.044 ± 0.104 for σ_i, A, B, and Σ_i, respectively. The 90% confidence lower limits on the slopes $\mathrm{d}\ln X/\mathrm{d}\ln M$ are thus -0.08, -0.13, -0.17 and -0.18 for $X = \sigma_i$, A, B, and Σ_i, respectively, while the 90% upper limits are less than 0.1 except for B (where is it 0.5). These shallow limits to the logarithmic slopes illustrate the near-universality of the distribution of interlopers in projected phase space.

However, the universality of the distribution of interlopers in projected phase space may hide an important level of cosmic variance. Performing MLE for parameters σ_i, A, and B, for each of the 93 clusters, each viewed in turn along each of three orthogonal viewing axes, we find standard deviations of

$$\sigma(\log\sigma_i) = 0.11, \quad \sigma(\log A) = 0.23, \quad \sigma(\log B) = 0.40. \qquad (14)$$

So, while the dispersion σ_i of the Gaussian component of the interloper velocity distribution is fairly constant (29% typical variations) from one cluster to the next, there is more scatter in the normalization A of the Gaussian component (factor 1.7 typical variations) and a large scatter in the flat component B (factor 2.5 typical variations).

5. Interloper removal

We remove the interlopers of the stacked cluster proceeding along similar lines as Łokas et al. (2006, see also Wojtak et al. 2007), by clipping the velocities beyond κ times the *local* line-of-sight velocity dispersion. We assume that our stacked cluster

has an NFW profile (Eq. (4)), or, alternatively, an Einasto (1965) profile:

$$\nu(r) = \frac{(2m)^{3m}}{m\,\gamma(3m,2m)}\,\exp\left[-2m\left(\frac{r}{r_{-2}}\right)^{1/m}\right]\left[\frac{M(r_{-2})}{4\pi r_{-2}^3}\right], \quad (15)$$

where $\gamma(a,x)$ is the incomplete gamma function. The density model of Eq. (15) fits the density profiles of ΛCDM halos even better than the NFW model (as first discovered by Navarro et al. 2004), at the expense of an additional parameter, m.

The velocity cut requires an estimate of the line-of-sight velocity dispersion profile of the halo. One could measure this in bins of projected radius, iteratively rejecting the outliers. This gives a profile that shows important radial fluctuations, and we would need to either smooth the profile or fit a smooth analytical function to it.

Instead, we choose to predict the line-of-sight velocity dispersion profile given the typical density and velocity anisotropy profiles of halos. The line-of-sight velocity dispersion profile can be written (Mamon & Łokas 2005b)

$$\sigma_{\rm los}^2(R) = \frac{2}{\Sigma(R)}\int_R^\infty K\left(\frac{r}{R},\frac{r_{\rm a}}{R}\right)\nu(r)\,v_{\rm c}^2(r)\,{\rm d}r, \quad (16)$$

where $v_{\rm c}(r) = \sqrt{GM(r)/r}$ is the circular velocity profile, $r_{\rm a}$ is the anisotropy radius, while the dimensionless kernel K is

$$K(u) = \sqrt{1-\frac{1}{u^2}}, \quad (17)$$

for isotropic orbits (Tremaine et al. 1994; Prugniel & Simien 1997) and

$$K(u,u_{\rm a}) = \begin{cases} \dfrac{1/2}{u_{\rm a}^2-1}\sqrt{1-\dfrac{1}{u^2}} + \left(1+\dfrac{u_{\rm a}}{u}\right)\cosh^{-1}u \\ \quad -{\rm sgn}\,(u_{\rm a}-1)\,u_{\rm a}\dfrac{u_{\rm a}^2-1/2}{\left(u_{\rm a}^2-1\right)^{3/2}}\left(1+\dfrac{u_{\rm a}}{u}\right) \\ \quad \times {\rm C}^{-1}\left(\dfrac{u_{\rm a}u+1}{u+u_{\rm a}}\right) & (u_{\rm a}\neq 1), \\ \left(1+\dfrac{1}{u}\right)\cosh^{-1}u - \dfrac{1}{6}\left(\dfrac{8}{u}+7\right)\sqrt{\dfrac{u-1}{u+1}} & (u_{\rm a}=1), \end{cases} \quad (18)$$

where

$$C^{-1}(x) = \begin{cases} \cos^{-1}x & \text{for } u_{\rm a}<1 \\ \cosh^{-1}x & \text{for } u_{\rm a}>1 \end{cases} \quad (19)$$

(Mamon & Łokas 2005b) for the anisotropy profile

$$\beta(r) = \frac{1}{2}\,\frac{r}{r+r_{\rm a}}, \quad (20)$$

which Mamon & Łokas (2005b, hereafter, ML) found to be a good fit to the anisotropy profiles of ΛCDM halos.

Since the space density model enters Eq. (16) expressing $\sigma_{\rm los}(R)$ (through the tracer density ν and the total mass M, which are related since we are considering single component mass models), we first need to determine the best fitting model to the distribution of particle radii in the stacked halo: we performed MLE of the NFW and Einasto models to the distribution of 3D radii of our stacked cluster. The minimum radius was chosen as $0.03\,r_{200}$ to avoid smaller radii, since our halo centers appear to be uncertain to about 1% of the virial radius. We varied the outer radii of the fit, starting at r_{200}. Since we will later fit the surface density profile out to the virial radius and beyond, we need to remember that the space radii extend beyond the maximum projected radius of the future surface density fits. So we also performed 3D fits beyond the virial radius: at $1.35\,r_{200}$ (which corresponds to the radius where $\Delta = 100$, i.e. the largest radius where the halos should be close to virial equilibrium), and $3\,r_{200}$ for a broader view of halos far beyond r_{200}.

When Prada et al. (2006) fit the density profiles of ΛCDM halos out to $2.7\,r_{200}$, they found them to be well approximated by the sum of an Einasto model and a constant term $\rho_{\rm bg} = \Omega_{\rm m}\rho_{\rm c}$. We therefore also experimented with the addition of a constant background component of density equal to the density of the Universe. In virial units, this background is expected to be equal to $\hat{\nu}_{\rm bg} = 3\,\Omega_{\rm m}/(4\pi\Delta) = 3.6\times10^{-4}$ (with $\Omega_{\rm m} = 0.3$ and $\Delta = 200$).

Table 1 shows the resulting best-fit concentrations[9] obtained by MLE fits of NFW and $m = 5$ Einasto models, plus an optional fixed or free constant background, to the distribution of 3D radii of the stacked halo[10]. The best-fit concentrations increase with the maximum allowed projected radius when the NFW model is used. This indicates the inadequacy of the NFW model at large radii, as it fails to capture the steepening of the slope of the density profile beyond the virial radius (Navarro et al. 2004). In fact, a Kolmogorov-Smirnov test (last column in Table 1) indicates that the NFW model is not an adequate representation of the distribution of radii, whether a constant background is added or not, regardless of the maximum radius used in the fit.

Table 1. MLE fits to the radial distribution of the stacked halo.

Model	$r_{\rm max}$	$100\,\hat{\nu}_{\rm bg}$	c	$P_{\rm KS}$
NFW	1	0	$4.08\pm0.05\pm0.17$	0.0021
NFW	1	(0.036)	$4.10\pm0.05\pm0.17$	0.0023
NFW	1	0.416	$4.26\pm0.07\pm0.23$	0.0069
Einasto	1	0	$\mathbf{4.00\pm0.05\pm0.17}$	0.31
Einasto	1	(0.036)	$4.02\pm0.05\pm0.17$	0.3
Einasto	1	0.002	$4.00\pm0.07\pm0.20$	0.31
NFW	1.35	0	$4.20\pm0.05\pm0.14$	0.0023
NFW	1.35	(0.036)	$4.23\pm0.05\pm0.14$	0.0021
NFW	1.35	0.002	$4.20\pm0.04\pm0.15$	0.002
Einasto	1.35	0	$\mathbf{4.14\pm0.04\pm0.13}$	0.016
Einasto	1.35	(0.036)	$4.17\pm0.04\pm0.13$	0.0058
Einasto	1.35	0.002	$4.14\pm0.04\pm0.14$	0.015
NFW	3	0	$4.80\pm0.04\pm0.24$	0
NFW	3	(0.036)	$5.01\pm0.04\pm0.24$	0
NFW	3	0.001	$4.80\pm0.03\pm0.19$	0
Einasto	3	0	$4.26\pm0.03\pm0.22$	$<10^{-15}$
Einasto	3	(0.036)	$\mathbf{4.50\pm0.03\pm0.23}$	$<10^{-10}$
Einasto	3	0.042	$4.54\pm0.04\pm0.13$	$<10^{-11}$

Notes. The Einasto models are for index $m = 5$. Column (2) is the maximum radius for the fits (the minimum radius is set to $0.03\,r_{200}$). Columns (3) and (4) are respectively the background density (multiplied by 100, fixed if in parentheses) and concentration, while Col. (5) is the Kolmogorov-Smirnov test probability that the model is consistent with the distribution of radii. Virial units (with $\Delta = 200$) are used for $r_{\rm max}$ and $\hat{\nu}_{\rm bg}$ (r_{200} and N_{200}/r_{200}^2, respectively), while $c = r_{200}/r_{-2}$, where r_{-2} is the radius of density slope equal to −2. The outer radius of $1.35\,r_{200}$ corresponds to the virial radius, r_{100}. Errors are 1σ: the first are the statistical errors from likelihood ratios and the second are from cosmic variance using 100 cluster bootstraps. Concentrations in bold face are those providing the highest $P_{\rm KS}$ tests for given $r_{\rm max}$.

[9] In this paper, concentrations refer to r_{200}/r_{-2}.

[10] We also experimented with free index Einasto models: we generally found that the best fit index was in the range $4.6 < m < 5.2$.

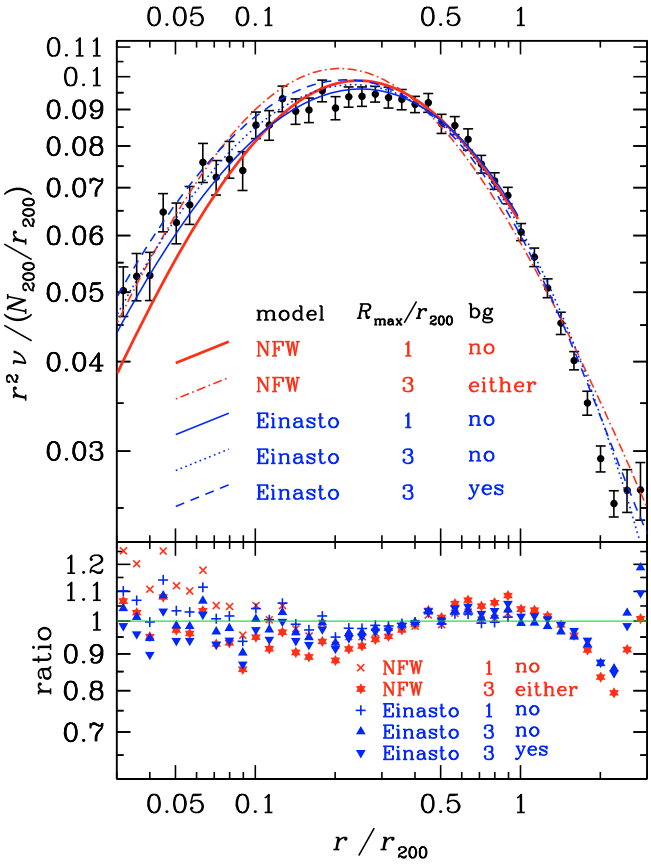


Fig. 10. *Top*: space density profile (multiplied by r^2) of the stack of the 93 halos, with best maximum likelihood fits for r/r_{200} from 0.03 to 1 (*solid*), and 3 (*dotted* without background, *dashed curves* with best-fit background, see Table 1) for the NFW (*red*) and Einasto (*blue*) models. The best fit NFW models to $3\,r_{200}$ with and without background are indistinguishable (see Table 1). The errors are from 100 cluster bootstraps. *Bottom*: ratios of measured to fit densities.

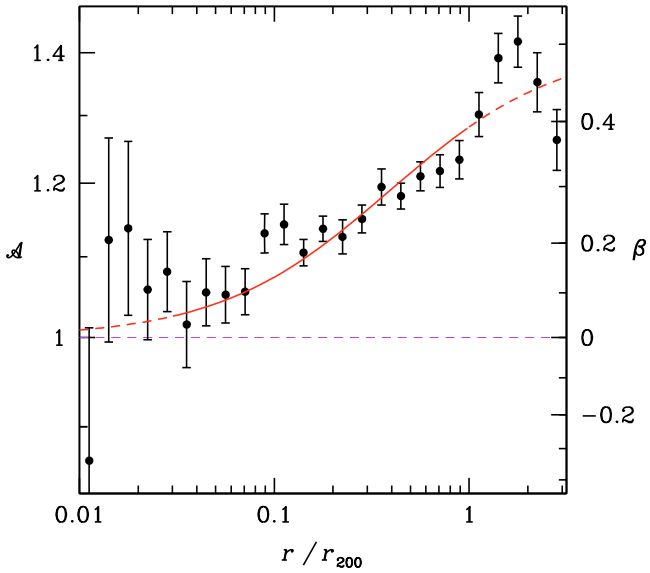


Fig. 11. Velocity anisotropy profile (including streaming motions: Eqs. (6) and (7)) of the stack of the 93 halos. The error bars are from 100 bootstraps on the 93 halos. The *curve* shows the weighted χ^2 fit (in the range $0.03 < r/r_{200} < 1$) of the Mamon-Łokas anisotropy (2005b) model (Eq. (20)) with $r_a = 0.27\,r_{200}$ (the *solid portion* of the curve highlights the region where the fit was performed). The *purple dashed horizontal line* indicates the fully isotropic case.

On the other hand, the concentration of the Einasto model appears to be somewhat less dependent of the outer radius, as also noted by Gao et al. (2008). At $0.03 < r/r_{200} < 1$, the Einasto model is an adequate representation of the distribution of radii. At $0.03 < r < r_{100} = 1.35$, the $m = 5$ Einasto model is inconsistent with the distribution of radii, but not by a large amount. However, the distribution of radii extending far beyond the virial radius, $0.03 < r/r_{200} < 3$ is not consistent with either NFW or $m = 5$ Einasto models[11].

The inclusion of the background in the fits leads to even higher concentrations when the maximum radius considered is $3\,r_{200}$.

The density profile of the 93 stacked clusters is shown in Fig. 10 for maximum fit radii of $r_{max} = 1$ and $3\,r_{200}$. The maximum likelihood NFW model produces clearly worse fits to the density profile than the maximum likelihood Einasto model for $r < 3\,r_{200}$, while the NFW model reproduces better the measured density profile at $r = 3\,r_{200}$. As clearly seen in the bottom panel of Fig. 10, neither model is adequate near $2\,r_{200}$, even when considering the cosmic variance measured by our cluster bootstraps (see the error bars in the top panel of Fig. 10).

In most of what follows, we restrict our analysis to $R < r_{200}$. For these analyses, we adopt the $c = 4$, NFW and $m = 5$ Einasto models, as these models are simple and the latter is consistent with the radial distribution without and with a constant background.

The radial profile of velocity anisotropy (Eq. (6)) of the stacked cluster, shown in Fig. 11, is reasonably well fit by (reduced $\chi^2 = 1.4$) the ML model[12] (Eq. (20)), with anisotropy radius (where β reaches its half-value of 1/4) $r_a = 0.275 \pm 0.020\,r_{200}$, as found with a weighted χ^2 fit for $\log r/r_{200}$ between −1.5 and 0. With a concentration parameter $c = 4.0$, the anisotropy radius is $r_a \simeq 1.1\,r_{-2}$, i.e. *the velocity anisotropy reaches its value intermediate between the center and the outer regions close to the radius of density slope* −2.

We now measure the velocity dispersion of the stacked profile using different schemes for interloper removal to find a scheme that produces a velocity dispersion profile (on the data with Hubble flow and the velocity cut) in agreement with the predictions (with no Hubble flow nor velocity cut) for the $c = 4.0$ NFW and Einasto models whose density profiles we just fit in 3D. Since the surface density profile enters Eq. (16), and no analytical formula is known for the Einasto model, we derived an accurate approximation for the Einasto surface density profile (for a large range of projected radii and of indices m) in Appendix A (Eqs. (A.15) and (A.14)).

The red open triangles in the top panel of Fig. 12 show the line-of-sight velocity dispersion profiles "measured" with our standard velocity cut at 3 times the predicted isotropic

[11] For $r_{max} = 3\,r_{200}$, the KS test showed that the free m Einasto model with free or fixed background (or without any) failed to provide an adequate representation of the distribution of radii.

[12] Other anisotropy models such as constant and Osipkov-Merritt (Osipkov 1979; Merritt 1985) produce much worse best fits (reduced $\chi^2 \simeq 5$ and 20, respectively).

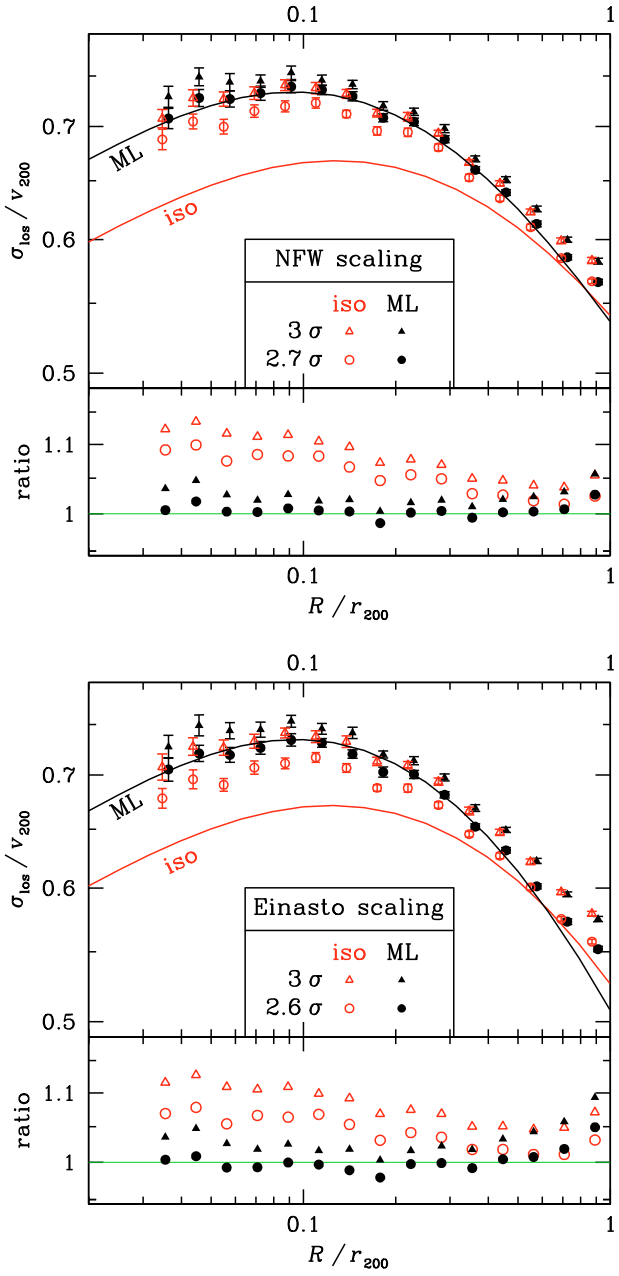


Fig. 12. Line-of-sight velocity dispersion profiles of the stacked virial cone, cutting at 3 (*triangles*) or 2.7 (*circles*) $\sigma_{\rm los}(R)$, assuming the best fitting ($c = 4.0$) NFW (*top*) or $m = 5$ Einasto (*bottom*) model with isotropic velocities (*red open symbols*) or slightly radial ML anisotropy (Eq. (20)) with $r_{\rm a} = r_{-2}$ (*black filled symbols*). The error bars are from 100 bootstraps on the particles within each bin of projected radii. For clarity, the isotropic and ML symbols are shifted by 0.01 dex leftwards and rightwards, respectively. The *curves* show the predicted line-of-sight velocity dispersions (for no Hubble flow, Eq. (16)) assuming isotropy (*red curve*, Eq. (17)) or ML anisotropy with $r_{\rm a} = r_{-2}$ (*black curve*, using Eq. (18)). The ratios of measured to predicted $\sigma_{\rm los}(R)$ are shown in the lower frames of each plot (same colors and symbols as upper frames).

line-of-sight velocity dispersion (hereafter $\sigma_{\rm los}^{\rm iso}$) for an NFW model with concentration $c = 4.0$ (as measured in 3D, see above). In comparison, $\sigma_{\rm los}^{\rm iso}$ (red solid curve in the top panel of Fig. 12) is typically 10% lower than the "measured" velocity dispersion profile for radii $R < 0.1\,r_{200}$. This discrepancy is decreased to 4% if one compares the measured velocity dispersions after clipping at 3 times the line-of-sight velocity dispersion, computed with the ML anisotropy (hereafter $\sigma_{\rm los}^{\rm ML}$, Eqs. (16) and (18), black filled triangles) to $\sigma_{\rm los}^{\rm ML}$ (black solid curve in the top panel of Fig. 12). This suggests that the 3σ clipping generally used is too liberal. A near perfect match (typically better than 1% for $R < 0.8\,r_{200}$) is obtained by cutting at $2.7\sigma_{\rm los}^{\rm ML}$ (black filled circles vs. black solid curve in the top panel of Fig. 12).

When the $m = 5$ Einasto model is used to compute $\sigma_{\rm los}(R)$ before applying the velocity cut, the best match between the measured and predicted line-of-sight velocity dispersion profiles is for a cut at $\kappa = 2.6$ (bottom panel of Fig. 12).

The Hubble flow (HF) causes a shallower slope at projected radii close to the virial radius (one notices in both panels of Fig. 12 an inflection point in the measured profiles (filled circles) of $\log\sigma_{\rm los}$ vs. $\log R$ near half a virial radius). Indeed, we obtained results similar to those of Fig. 12 when we did not incorporate the HF to the peculiar velocities of the simulation: the measured $\sigma_{\rm los}(R)$ fell more sharply, with no inflection point, even somewhat more sharply than predicted by the Einasto model (because the velocity anisotropy without the HF is more radial at $\simeq 4\,r_{200}$ in comparison with the case where the HF is incorporated, where $4\,r_{200}$ roughly corresponds to the turnaround radius where the velocities are mostly tangential). So, although the steeper Einasto density profile ought to catch better the steeper line-of-sight velocity dispersion profile at large projected radii, the NFW model performs slightly better, because its shallower line-of-sight profile mimics better the effects of the Hubble flow.

In summary, Fig. 12 indicates that if one wishes to recover the correct line-of-sight velocity dispersion profile, *one should use 2.6 or 2.7σ clipping instead of 3σ clipping, where the line-of-sight velocity dispersion is either measured or modeled with anisotropic velocities*.

The choice of model and κ is not obvious. We prefer the NFW model, as it is simpler and, with $\kappa = 2.7$, it presents a slightly better match between measured and predicted line-of-sight velocity dispersion profiles than does the $m = 5$ Einasto model (compare the ratios of measured to predicted $\sigma_{\rm los}(R)$ in both plots of Fig. 12, especially at large radii).

Mass modelers of clusters may wish to avoid performing the integral of Eq. (16) with the kernel of Eqs. (18) and (19). The line-of-sight velocity dispersion profile (Eq. (16)) for the NFW and $m = 5$ Einasto models with ML anisotropy with $r_{\rm a} = r_{-2}$ can be approximated as

$$\frac{\sigma_{\rm los}(R)}{\sqrt{GM(r_{-2})/r_{-2}}} \simeq \mathrm{dex}\left\{\sum_{i=0}^{7} a_i \left[\log_{10}\left(\frac{R}{r_{-2}}\right)\right]^i\right\} \qquad (21)$$

where the coefficients are given in Table 2 for both models. These two approximations are accurate to better than 0.5% (rms) for $0.0032 < R/r_{-2} < 32$.

Then, one can write $\sigma_{\rm los}(R)$ in terms of v_v using Eq. (21) and

$$\frac{v_v^2}{GM(r_{-2})/r_{-2}} = \begin{cases} \dfrac{\ln(c+1) - c/(c+1)}{(\ln 2 - 1/2)\,c} & \text{(NFW)}, \\ \dfrac{\gamma\left(3m, 2m\,c^{1/m}\right)}{\gamma(3m, 2m)\,c} & \text{(Einasto)}, \end{cases} \qquad (22)$$

(e.g., Navarro et al. 1996; for NFW and trivially derived from the Einasto mass profile first derived by Mamon & Łokas 2005a).

In the absence of information on the mass profile (e.g. from X-ray observations), neither the concentration parameter, c, nor

Table 2. Coefficients for $\sigma_{los}(R)$ approximation (Eq. (21)).

	NFW	$m = 5$ Einasto
a_0	–0.1478	–0.1520
a_1	–0.1109	–0.1242
a_2	–0.1357	–0.1637
a_3	0.001948	–0.01688
a_4	0.02317	0.01892
a_5	0.0006310	0.001844
a_6	–0.003234	–0.002044
a_7	–0.0006370	–0.0004103

the scale radius r_{-2} are known, so one has to work iteratively, first guessing a plausible value of c, applying the velocity filter, then re-estimating c from the data and re-applying the velocity filter. This process should converge in a one or two iterations.

6. Interloper statistics after the velocity cut

We now show the statistics of interlopers after our adopted velocity cut. The motivation is to allow observers to compare with their own data. With the velocity cut ($\kappa = 2.7$), the line-of-sight distances are now effectively limited to $\pm 17\, r_{200}$ from the center of the stacked halo (Eq. (5)). The qualitative features of the projected phase space distribution are robust to variations of the method to cut the velocities.

The green curves in Fig. 3 show the velocity cut at $\pm 2.7\sigma_{los}(R)$ – with our adopted NFW model with ML anisotropy with anisotropy radius $r_a = r_{-2}$ – on top of the projected phase space (using Eqs. (21) and (22)). Only 0.4% of the halo particles are rejected by the 2.7σ velocity cut, which is a low enough fraction that the shot noise in the structural and kinematical modeling is not significantly increased. The velocity cut in Fig. 3 seems very reasonable as it is close to optimizing the completeness of the selection of particles within the virial sphere. However, less than 17% of the interlopers are identified as such by the velocity cut. Therefore, *the great majority of interlopers cannot be removed by a velocity cut*.

Figure 4 shows the velocity cut on top of the phase space density distribution. The velocity cut appears to occur in a region where the interloper phase space density is roughly constant.

This can be seen in a clearer fashion in Fig. 6, where the velocity cut is shown as green vertical lines. While the highest velocity interlopers are removed by the velocity cut, there remains signs of the field component, which we had identified with particles beyond 8 virial radii, at low ($R < 0.4\, r_{200}$) projected radii.

The fraction of particles outside the virial sphere is displayed in Fig. 13. Interestingly, at $R > 0.8\, r_{200}$ (magenta histogram) the interlopers account for over 60% of all particles, regardless of the particle velocity up to the velocity cut (filled circles). But even at smaller radii, $0.4 < R/r_{200} < 0.6$, interlopers account for over 20% of all particles again for all velocities up to the cut. So, unless one limits one's kinematical analysis to very small cluster apertures, one cannot avoid being significantly contaminated by interlopers.

While there is no gap in the velocity distribution of particles (Fig. 6), Fig. 13 shows local minima of the interloper fraction for all bins of projected radii, except the outermost one. Regardless of the application of a velocity cut, these local minima occur at lower velocity (1.3 to 1.5 v_v) than the inflection points of the interloper density in projected phase space (1.6 to 2.6 v_v as seen in Fig. 6). The local minima occur at velocities that decrease with projected radius, suggesting that our local $\kappa\sigma_{los}$ cut is preferable

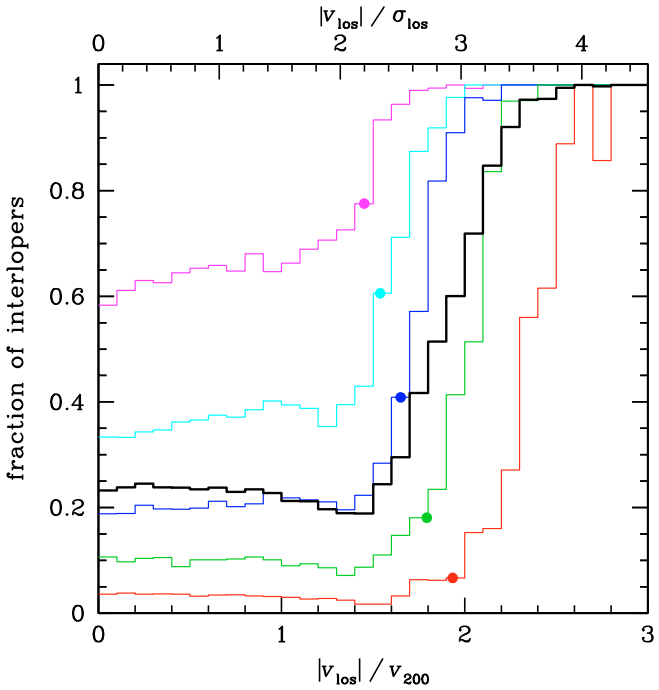


Fig. 13. Fraction of interlopers as a function of line-of-sight velocity for all projected radii $R < r_{200}$ (*thick black histogram*) and in bins of projected radius: $R/r_{200} = 0$–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1 (*thin histograms*), increasing upwards. The *filled circles* show the $2.7\sigma_{los}(R)$ (from Eqs. (21) and (22)) velocity cut for the $c = 4$ NFW model with $r_a = r_{-2}$ ML anisotropy (Eq. (20)).

to a global one, since $\sigma_{los}(R)$ decreases with R for $R > 0.1\, r_{200}$ (see Fig. 12). These local minima arise because the interloper system has a lower velocity dispersion than the halo system: $\sigma_i = 0.58$ (Eq. (10)) while after the velocity cut the aperture velocity dispersion of the global stacked virial cone (thus including both halo particles and interlopers) is $\eta = 0.65$, i.e. 5% higher than predicted by Mauduit & Mamon (2007) for an isotropic NFW model, which is not surprising given the radial anisotropy of the halos (Fig. 11).

The surface density profile of the stacked halo is shown in Fig. 14. The surface density profile of the interlopers is flat with small fluctuations around the mean values $\Sigma_i = 0.114$ and $0.096\, N_v\, r_v^{-2}$, measured in the stacked virial cone, respectively before and after the velocity cut[13]. In comparison, our model of the surface density of interlopers (Eq. (12)) combined with our MLE values for A, B and σ_i yields mean interloper densities of 0.114 and $0.094\, N_v\, r_v^{-2}$, respectively before and after the $\kappa = 2.7$ ($\hat{\kappa} = 1.76$ with $\eta = 0.65$) velocity cut. The general agreement is excellent.

Note that, at projected radii beyond the virial radius, all particles are interlopers, so the surface density of interlopers is not constant but decreases, to first order, as the NFW or Einasto models. While the total surface density profiles of the popular NFW and Einasto models for ΛCDM halos are convex in log surface density vs. log projected radius, an important additional background term in the surface density would lead to an inflection point and subsequent concavity at some radius. Such a feature would lead to poor fits of single NFW or Einasto profiles.

[13] Figure 14 shows interloper surface densities that are lower, at $R < 0.7\, r_{200}$, than 0.114 and $0.096\, N_v\, r_v^{-2}$, respectively before and after the velocity cut, but most of the particles lie within the highest bins of log projected radius.

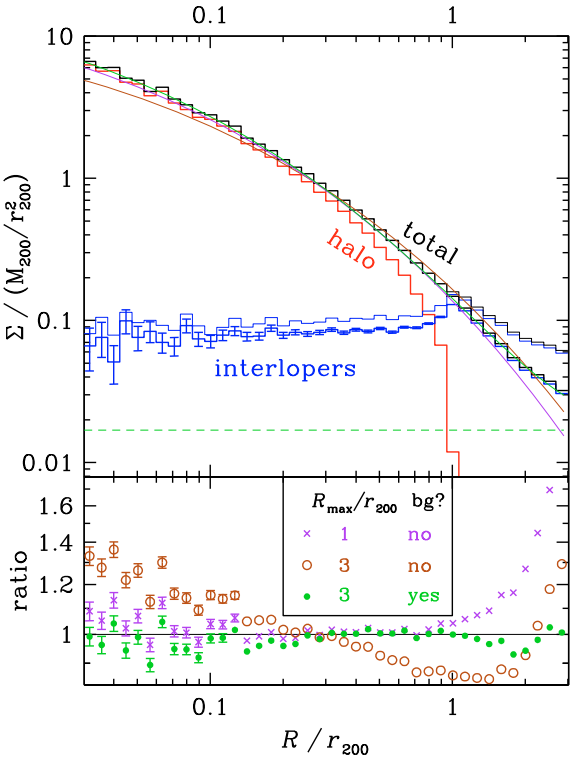


Fig. 14. *Top panel*: surface density profile of global stacked cone (*black histogram*, raised up by 0.02 dex for clarity), as well as halo members ($r \le r_{200}$, *red histogram*) and interlopers ($r > r_{200}$) before (*thin*) and after (*thick blue histograms*) the velocity cut. Poisson errors are only shown for the interlopers after the velocity cut for $R < r_{200}$. Also shown (*curves*) are maximum likelihood $m = 5$ Einasto model fits (after the velocity cut), in the range 0.03 r_{200} to 1 (*purple*) or 3 (*brown and green*) r_{200}, without (*purple and brown*) or with (*green*) an additional free constant background component (*dashed green line*). *Bottom panel*: ratios of measured to fit surface densities (after the velocity cut). For clarity, Poisson errors are only shown if larger than the symbol size.

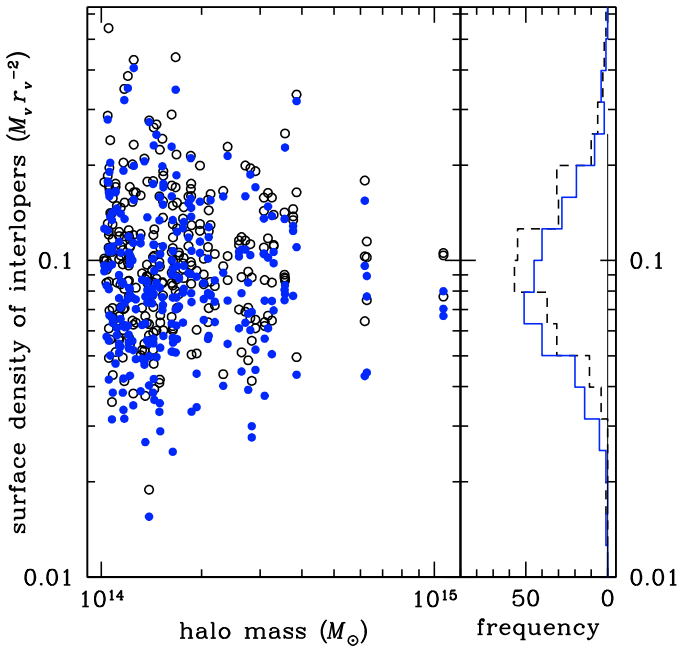


Fig. 15. Mean interloper surface density (in virial units) versus halo mass (for $\Delta = 200$). The *open black* and *filled blue circles* show the 93 halos each measured along 3 orthogonal viewing directions, before and after the velocity cut, respectively. The *right panel* provides the frequency of the mean interloper surface density (with logarithmic bins), before (*dashed black*) and after (*solid blue histograms*) the velocity cut.

Table 3. Statistics of interloper surface densities Σ_i (in $N_v\, r_v^{-2}$).

Velocity cut	No	Yes
Arithmetic mean	0.116	0.098
Geometric mean	0.102	0.085
Standard deviation of $\log \hat{\Sigma}_i$	0.216	0.229
Spearman rank correlation	0.025	−0.008
Probability	0.34	0.44

Notes. Spearman rank correlation is between mean surface density of interlopers and halo mass. Probability is of having a stronger correlation by chance.

Now, within the virial radius, no such inflection point and outer concavity are seen in Fig. 14 for the *total* surface density profile. However, extending the surface density profile out to three virial radii, as illustrated in Fig. 14, one does see the inflection point of the total surface density profile (near 1.6 r_{200} after the velocity cut and right at r_{200} before). This points to an additional background of surface density. This requirement for the additional background component is confirmed by the fairly flat ratios of data over model (bottom panel of Fig. 14) for the case where the background is fit, in comparison with larger residuals for $R > r_{200}$ for the fits without a background.

Such a background is expected, since the density profiles of ΛCDM halos is the sum of an Einasto model and a constant background corresponding to the mean density of the Universe (Prada et al. 2006). Integrating along the line-of-sight (Eq. (1)) within the sphere of radius $r_{\rm max}$ (Eq. (5)), one deduces that the surface density profile of (foreground/)background structures is

$$\Sigma_{\rm bg}(R) = 2\,\Omega_{\rm m}\,\rho_{\rm c}\,\sqrt{r_{\rm max}^2 - R^2}. \quad (23)$$

Since $r_{\rm max} \gg r_v$ (Eq. (5)), $\Sigma_{\rm bg}$ is roughly constant for $R \lesssim r_v$:

$$\Sigma_{\rm bg} \simeq 2\,\Omega_{\rm m}\,\rho_{\rm c}\,r_{\rm max}, \quad (24)$$

which in dimensionless virial units ($N_v\, r_{\rm v}^{-2}$) becomes

$$\hat{\Sigma}_{\rm bg} = \frac{\Sigma_{\rm bg}}{N_v/r_v^2} \simeq \frac{3}{\pi}\,\frac{\eta\,\kappa\,\Omega_{\rm m}}{\sqrt{8\,\Delta}}, \quad (25)$$

using Eq. (3). With the velocity cut, $\kappa = 2.7$ and Eq. (25) yields $\hat{\Sigma}_{\rm bg} \simeq 0.0126$ for $\Omega_{\rm m} = 0.3$, $\Delta = 200$, and $\eta = 0.65$ (see above). Without the velocity cut $\hat{\kappa} = \eta\,\kappa = 4$ and one obtains $\hat{\Sigma}_{\rm bg} \simeq 0.0286$. According to Eq. (23), the relative drops of $\Sigma_{\rm bg}$ from $R = 0$ to $R_{\rm max} = 1$ or $3\,r_v$ are 0.2% and 1.6%, respectively, so the approximation of a constant surface density background is adequate.

This background corresponds to the velocity-independent component of the interloper surface phase space density, which in our model (Eq. (8)) is the B term, which produces a mean surface density of $\hat{\Sigma}_{\rm bg} = \eta \times 2.7\,B = 0.0132$ (with $B = 0.0075$ from Eq. (10) and again $\eta = 0.65$). This agrees with the previous value to within 4%. This means that the constant field term in the velocity distribution corresponds precisely to the additional halos outside the test halo. In any event, *the total surface density of a cosmological structure is the sum of the surface density of that structure and a constant background*.

Although the mean surface density of interlopers is roughly independent of halo mass (Fig. 9), it may vary from cluster to cluster. Figure 15 shows the surface density of each of the 93 halos. As seen in Table 3, the arithmetic mean value of Σ_i matches

well the value of the stacked virial cone, regardless of the velocity cut, which has only a minor effect on the statistics of interlopers (while the geometric means are ≃15% lower). But the dispersion in $\log\hat{\Sigma}_i$ is as high as 0.22, close to $\sigma(\log A)$ (Eq. (14)), so that the relative dispersion of Σ_i is as high as a factor $10^{0.22} \simeq 5/3$.

The fraction of interlopers in the stacked virial cone is

$$f_i = \frac{N_i}{N_h + N_i} = \frac{\pi\hat{\Sigma}_i}{1 + \pi\hat{\Sigma}_i} \quad (26)$$

(where indices "h" and "i" correspond to halo and interloper particles, respectively), where the second equality of Eq. (26) made use of $\hat{\Sigma}_h = \Sigma_h/[N_v/r_v^2] = 1/\pi$, by definition. This yields $f_i = 27\%$ before and $23.1 \pm 0.1\%$ after the velocity cut, where the error is both from binomial statistics and from a bootstrap on all the particles; a bootstrap on the halos leads to an error of 0.6%; propagating (with Eq. (26)) the error on the mean of $\hat{\Sigma}_i$ (from the standard deviation of $\hat{\Sigma}_i$ given in Table 3) also leads to an error on f_i of 0.6%; finally, the standard deviation of the interloper fraction for the three stacked cartesian virial cones is 1.7%. We adopt the error estimate of 0.6% on f_i for the later discussion.

The small decrease in interloper fraction from before to after the velocity cut confirms our finding that the large majority of interlopers have too low velocities to be filtered by velocity. Cosmic variance causes huge fluctuations in the fraction or surface density of interlopers (2/3 of the mean value, independent of the presence of a velocity cut), with roughly a log-normal distribution (see the right panel of Fig. 15 and Table 3). The last two lines of Table 3 indicate that there is no statistically significant correlation of surface density of interlopers with halo mass.

7. Biases in concentration and anisotropy?

What are the effects of the Hubble flow on estimates that observers make on halos, e.g. the concentration and the velocity anisotropy of the distribution of their tracer constituents (i.e. galaxies in clusters)?

7.1. Effects of the Hubble flow

We begin by a naïve comparison of the observable distributions with and without the Hubble flow, before comparing the concentration and anisotropy measured by an observer with the corresponding quantities we directly infer in 3D from the cosmological simulations. Admittedly, the distribution of velocities without the Hubble flow is not fully realistic, since the cosmological simulation solved equations for comoving coordinates in an expanding universe[14].

Figure 16 shows the changes in the radial profiles of surface density and line-of-sight velocity dispersion and kurtosis[15], once the Hubble flow is added to the peculiar velocities. One striking feature of Fig. 16 is that the Hubble flow leads to a lower surface density profile at large radii. This cannot be a consequence of the restriction of the line-of-sight of the *halo* component to $\pm 19\, r_{200}$ (see Sect. 1), because the NFW surface density with

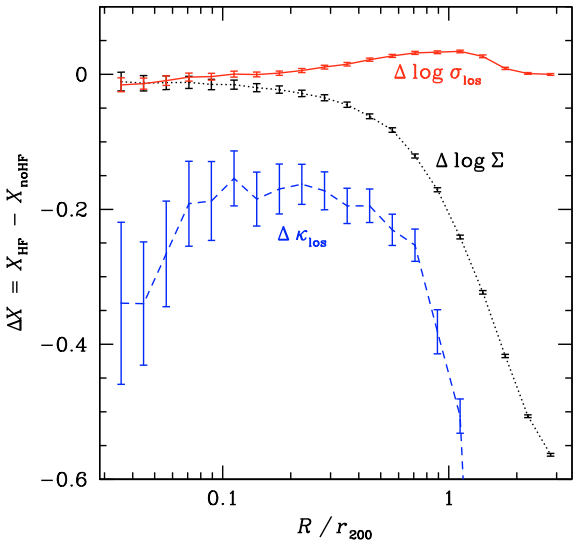


Fig. 16. Difference of velocity moments with Hubble flow and velocity cut (HF) and without Hubble flow or velocity cut (noHF): log surface density (*dotted black*), log line-of-sight velocity dispersion (*solid red*), and line-of-sight velocity kurtosis (*dashed blue*). The error bars are based upon Poisson errors for the surface density and bootstraps within the radial bin for the log dispersion and the kurtosis, where the error on the difference is the square root of the sum of the square errors.

line-of-sight limited to the sphere of that radius (Appendix B) matches the NFW surface density projected to infinity to better than 1.4% relative accuracy for $R < 3\, r_{200}$ (for $c = 4$). Instead, it is the integral along the line of sight of the *constant density* (foreground/background) *component* that diverges when no Hubble flow is present, and is limited to half a box size here: $L/2 = 96\, h^{-1}$ Mpc, which corresponds to roughly 100 virial radii. In any event, the lower (40% lower at r_{200}) surface density profile found when the Hubble flow is added to the peculiar velocities might explain the lack of concavity in the (log−log) surface density profile within the virial radius (Fig. 14).

We noticed in Sect. 5 that the line-of-sight velocity dispersion profile showed an excess at radii near the virial radius. Figure 16 confirms that the line-of-sight velocity dispersion profile is gradually overestimated at large radii, while the surface density profile is much more biased beyond half a virial radius, being underestimated at large radii. The effects of the Hubble flow on the surface density and line-of-sight velocity dispersion profile are both small at very low projected radii.

Finally, our sharp cut (Sect. 5) in the distribution of line-of-sight velocities when the Hubble flow is incorporated implies that *the line-of-sight velocity kurtosis is underestimated (especially at large projected radii)*.

7.2. Concentration

Does the excess of interlopers at large projected radii lead to lower values of the concentration parameter in the fits of the projected NFW and Einasto profiles to the surface density profiles of clusters?

Table 4 shows the MLE fits (see Appendix C) of the NFW and $m = 5$ Einasto surface density profiles to the distribution of projected radii of the three cartesian stacked cones. Different fits were performed with variations in the maximum allowed projected radius, R_{max}, the presence of a constant (fixed or free) background term, and the possible removal of high-velocity outliers. The NFW surface density and projected number (or

[14] In fact, static universes are never simulated in a cosmological context, because of their lack of realism, given the expansion of the Universe as seen in the Hubble law, and also because of the lack of knowledge of suitable initial conditions in such a static universe.

[15] The reader should not confuse the line-of-sight velocity kurtosis κ_{los} with the velocity cutoff in units of line-of-sight velocity dispersion (κ) or of virial velocity ($\hat{\kappa}$).

Table 4. MLE fits to the distribution of projected radii of the three cartesian stacked cones.

Model	R_{max}	Vel. cut	No bg		Fixed bg		Free bg			
			c_{2D}	P_{KS}	c_{2D}	P_{KS}	$100\,\hat{\Sigma}_{bg}$	$(\sigma(\log\hat{\Sigma}_{bg}))$	c_{2D}	P_{KS}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		(9)	(10)
NFW	1	N	$3.46 \pm 0.04 \pm 0.21$	$<10^{-5}$	$4.06 \pm 0.05 \pm 0.26$	0.18	2.4	(0.20)	$3.95 \pm 0.05 \pm 0.03$	0.08
NFW	1	Y	$3.84 \pm 0.05 \pm 0.19$	0.059	$4.10 \pm 0.05 \pm 0.21$	0.19	1.0	(0.07)	$4.06 \pm 0.11 \pm 0.17$	0.2
Einasto	1	N	$3.30 \pm 0.04 \pm 0.21$	0.00053	$3.90 \pm 0.05 \pm 0.25$	0.46	2.6	(0.18)	$3.83 \pm 0.05 \pm 0.08$	0.27
Einasto	1	Y	$3.70 \pm 0.04 \pm 0.18$	0.32	$3.96 \pm 0.05 \pm 0.20$	0.42	0.8	(0.01)	$3.87 \pm 0.12 \pm 0.19$	0.65
NFW	1.35	N	$3.21 \pm 0.03 \pm 0.24$	$<10^{-14}$	$4.07 \pm 0.05 \pm 0.31$	0.11	2.9	(0.08)	$4.06 \pm 0.07 \pm 0.15$	0.25
NFW	1.35	Y	$3.78 \pm 0.04 \pm 0.23$	0.052	$4.19 \pm 0.05 \pm 0.26$	0.06	0.5	(0.29)	$3.96 \pm 0.07 \pm 0.16$	0.22
Einasto	1.35	N	$3.00 \pm 0.03 \pm 0.22$	$<10^{-13}$	$3.83 \pm 0.04 \pm 0.28$	0.32	3.3	(0.08)	$3.95 \pm 0.06 \pm 0.16$	0.52
Einasto	1.35	Y	$3.59 \pm 0.04 \pm 0.20$	0.04	$3.98 \pm 0.04 \pm 0.22$	0.46	0.9	(0.10)	$3.89 \pm 0.08 \pm 0.17$	0.73
NFW	3	N	$1.67 \pm 0.01 \pm 0.16$	0	$3.82 \pm 0.04 \pm 0.37$	0	3.8	(0.03)	$4.42 \pm 0.04 \pm 0.45$	$<10^{-6}$
NFW	3	Y	$3.06 \pm 0.02 \pm 0.12$	0	$4.31 \pm 0.04 \pm 0.17$	$<10^{-9}$	1.2	(0.12)	$4.34 \pm 0.05 \pm 0.38$	0.00003
Einasto	3	N	$1.42 \pm 0.01 \pm 0.12$	0	$3.19 \pm 0.03 \pm 0.34$	0	4.3	(0.05)	$4.15 \pm 0.03 \pm 0.67$	0.00003
Einasto	3	Y	$2.65 \pm 0.02 \pm 0.10$	0	$3.75 \pm 0.03 \pm 0.12$	$<10^{-13}$	1.5	(0.03)	$4.02 \pm 0.01 \pm 0.12$	0.00055

Notes. Column 1: model (NFW or $m = 5$ Einasto); Col. 2 (R_{max}): projected radius of the cone in which the stacked cluster is built, in units of r_{200}; Col. 3 (v-cut): presence (Y) or absence of the velocity cut with $\kappa = 2.7$ (NFW) or 2.6 (Einasto); Cols. 4–5, 6–7, 9–10: mean best-fit concentration ($c = r_{200}/r_{-2}$) from projected radii and probability that distribution of projected radii is consistent with model using a Kolmogorov-Smirnov test (P_{KS}), for fits without a background (Cols. 4 and 5), with a fixed background ($\hat{\Sigma}_{bg} = 0.0286$ (no velocity cut) or 0.0126 (with velocity cut), Cols. 6 and 7) or a free background (Cols. 9 and 10), with best-fit value (100 times the geometric mean and error on its logarithm in parentheses) given in Col. 8. The minimum projected radius is set to $0.03\,r_{200}$. For the Einasto model, we adopt the approximation to the surface density and projected number (mass) profiles given in Appendix A. The errors on c are statistical (first) and a measure of the cosmic variance term estimated by the gapper (Wainer & Thissen 1976, see Beers et al. 1990) standard deviation of the MLE values for the three projection axes.

equivalently projected mass) profiles, required for the normalization of the probability used in the MLE, are given by Bartelmann (1996) and, in another form by Łokas & Mamon (2001). The surface density and projected number (mass) profiles of the Einasto model are not known in analytical form, so we have derived accurate approximations in Appendix A. For these MLE, we adopt Eqs. (A.17) with (A.6) for the surface density profile and Eqs. (A.8) with (A.6) for the projected mass profile. When a constant background term is included in the fits, it is either free (Col. 8) or fixed at $\hat{\Sigma}_{bg} = \hat{\kappa}\,B = 0.0286$ and 0.0126 without ($\hat{\kappa} = 4$) and with ($\hat{\kappa} = \eta\,\kappa = 1.75$) the velocity cut (see Sect. 6).

The concentrations measured on the projected radii with single component fits are always smaller than the characteristic value found in 3D ($c = 4.0$, 4.1, and 4.5, see bold values in Table 1). The best-fit concentrations are very low when the fits are performed out to $3\,r_{200}$ (unless a velocity cut is performed or background component added to the model). Note that when the background is fitted together with the concentration and no velocity cut is performed, the best-fit value for the background can be over a factor two off from the value expected from Eq. (24), even with maximum projected radii of $3\,r_{200}$. The KS tests indicate that for most combinations of maximum projected radius, presence or absence of the velocity cut and how the background is handled, the $m = 5$ Einasto model usually provides a better representation of the distribution of projected radii than does the NFW model. Finally, the errors in Table 4 indicate that the cosmic variance of stacks of 93 clusters, measured using the standard deviation of the three cartesian stacked cones with the gapper[16] estimate of dispersion, which is most robust to small sample sizes (Beers et al. 1990), are much greater than the intrinsic fitting errors.

Table 5 shows the bias in concentrations, where the bias is the ratio between the concentration measured with the 2D fit (Table 4) and the best-fitting of the concentrations found with the 3D fits (highlighted in bold in Table 1), for given r_{max}/r_{200} and different models and backgrounds. The concentration parameters found in the fits of the surface density profile are underestimated by typically $16\% \pm 7\%$ when we make no velocity cut, limit the projected radii to r_{200}, and do not incorporate a constant background in the fit. This underestimate of the concentration is statistically significant (marginally so for NFW). The concentration bias gets worse as we increase the maximum projected radius: the concentration is underestimated by 1/4 at $R < r_{100}$ and by as much as 2/3 with $R < 3\,r_{200}$.

But when we make the velocity cut at 2.7 (NFW) or 2.6 (Einasto) $\sigma_{los}(R)$, where $\sigma_{los}(R)$ is obtained from Eqs. (21) and (22), the biases in the concentration parameters are typically reduced by a factor two. When we limit the analysis at $R < r_{200}$, *the concentration parameters found in the fits of the surface density profiles are (only) underestimated by typically* $6 \pm 6\%$. This low bias is no longer statistically significant given our fit and cosmic variance errors[17]. However, extending the analysis to $R = r_{100}$, the concentration is still biased low by as much as 11% (which is marginally significant), despite the velocity cut. And if we go all the way to $3\,r_{200}$, the bias is still very strong, as the concentration is underestimated by over 1/3.

One may wonder whether one can recover the concentration parameter more accurately with two-component fits to the set of projected radii than with the single-component fit, since we found an additional background needs to be added (Sect. 6). For example, Lin et al. (2004) fit a projected NFW model plus a constant surface density background (hereafter NFW+bg), with no velocity cut, for projected radii $0.02 < R_{max}/r_{200} < 2.5$. We tested the single and double-component models in the optimistic case of our stacked cluster with nearly 3×10^5 particles. As seen in Table 5, *the 3D concentration parameter is recovered better by the two-component models, regardless of the velocity cut*,

[16] The gapper dispersion of a vector $\boldsymbol{x}$ of length n is $s = \sqrt{\pi}/[n(n-1)]\ \sum_{i=1}^{n-1} i(n-i)\,(x_{i+1} - x_i)$ (Wainer & Thissen 1976).

[17] We also experimented with concentration fits on the projected radii of all particles within $3\sigma_{los}(R)$ instead of 2.7, but the changes were very small (less than 1%, with slightly worse underestimates of the concentration for the single component fits).

Table 5. Concentration bias of 2D fits.

Model	$R_{\rm max}$	$c_{\rm 2D}/c_{\rm 3D}^{\rm best}$					
		no v-cut			v-cut		
		no bg	fixed bg	free bg	no bg	fixed bg	free bg
NFW	1	0.86 ± 0.07	$\mathbf{1.01 \pm 0.08}$	$\mathbf{0.99 \pm 0.05}$	$\mathbf{0.96 \pm 0.06}$	$\mathbf{1.02 \pm 0.07}$	$\mathbf{1.01 \pm 0.07}$
Einasto $m = 5$	1	0.82 ± 0.06	$\mathbf{0.97 \pm 0.08}$	$\mathbf{0.96 \pm 0.05}$	$\mathbf{0.93 \pm 0.06}$	$\mathbf{0.99 \pm 0.07}$	$\mathbf{0.97 \pm 0.07}$
NFW	1.35	0.78 ± 0.06	$\mathbf{0.98 \pm 0.08}$	$\mathbf{0.98 \pm 0.05}$	$\mathbf{0.91 \pm 0.06}$	$\mathbf{1.01 \pm 0.07}$	$\mathbf{0.96 \pm 0.05}$
Einasto $m = 5$	1.35	0.72 ± 0.06	$\mathbf{0.93 \pm 0.07}$	$\mathbf{0.95 \pm 0.05}$	0.87 ± 0.06	$\mathbf{0.96 \pm 0.06}$	$\mathbf{0.94 \pm 0.05}$
NFW	3	0.37 ± 0.04	0.85 ± 0.09	0.98 ± 0.11	0.68 ± 0.04	0.96 ± 0.06	0.96 ± 0.10
Einasto $m = 5$	3	0.32 ± 0.03	0.71 ± 0.08	0.92 ± 0.16	0.59 ± 0.04	0.83 ± 0.05	0.89 ± 0.05

Notes. The biases highlighted in bold (respectively blue italics) show the cases where the best-fit surface density profile was consistent with the data to better than 5% (between 0.01% and 5%) confidence (see last column of Table 4). The errors are the statistical (MLE fit) and cosmic variance (from the 3 cartesian stacked halos) errors of Table 4 added in quadrature together and with the analogous errors on the best-fit (bold in Table 1) 3D concentration, with $\sigma^2(c_{\rm 2D}/c_{\rm 3D}^{\rm best}) = \sigma^2(c_{\rm 2D})/\left\langle c_{\rm 3D}^{\rm best}\right\rangle^2 + \sigma^2(c_{\rm 3D}^{\rm best})\,\langle c_{\rm 2D}\rangle^2 / \left\langle c_{\rm 3D}^{\rm best}\right\rangle^4$.

although the improvement is not always statistically significant. Note that the background is well recovered with $R_{\rm max} = 3\,r_{200}$ and a velocity cut for both the NFW+bg and Einasto+bg models (Table 4).

In summary, *the concentrations measured in 2D recover best the 3D values once the velocities are filtered and especially once a background is included in the fit (even when limited to the virial radius)*.

Note also that one should not attempt to model the surface density profile as the sum of the halo term (with *line-of-sight limited to the sphere*, with the formulae of Appendix B for the NFW model) and a constant background, because the *total* surface density profile decreases smoothly beyond the virial radius in ways that are not simple to model, for example with a spherical halo and a constant background. Moreover, for fits where the projected radii are limited to the virial radius, these spherical plus background fits are not recommended because the background is not constant but rises with radius (Figs. 5 and 8) and is less easy to model than the surface density with line-of-sight integrated to infinity.

7.3. Velocity anisotropy

The signature of the Hubble flow on the shape of the line-of-sight velocity dispersion profile (Fig. 12) suggests that the velocity anisotropy that is recovered may be affected. Figure 17 shows the result of the non-parametric anisotropy inversion (first developed by Binney & Mamon 1982; but we use here the simpler algorithm by Solanes & Salvador-Solé 1990), which computes the anisotropy profile assuming a smooth representation of the line-of-sight velocity dispersion profile and a mass model. Here, we adopt a $c = 4$ NFW model and fit polynomials to the binned $\log\sigma_{\rm los}$ vs. $\log R$[18].

In the region where the order of the polynomial fit does not matter ($0.1 < r/r_{200} < 1$), the recovered anisotropy profile reproduces very well the one measured in three dimensions (points in Fig. 17), although, beyond $0.2\,r_{200}$, the recovered anisotropy profiles are slightly more radial than that measured in 3D. This bias towards more radial motions appears statistically significant, since in all six radial bins where there is a $\simeq 1\sigma$ offset, this offset is in the same direction (probability of $2^{-5} = 3\%$). Therefore, *the Hubble flow produces only a slight radial velocity anisotropy bias in the envelopes of halos.*

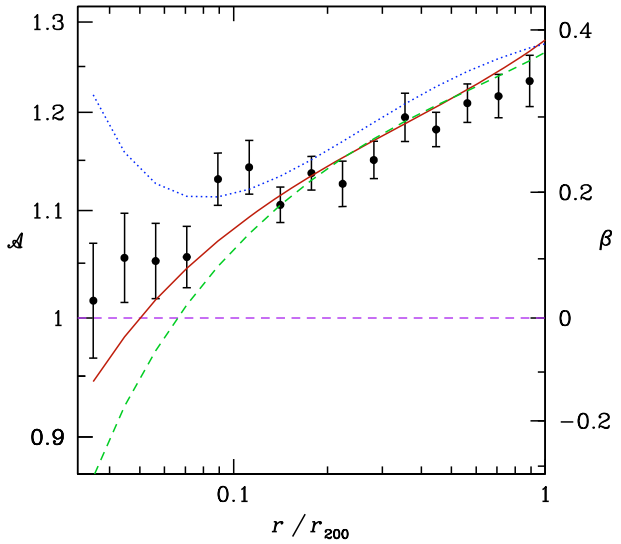


Fig. 17. Velocity anisotropy profiles (including streaming motions: Eqs. (6) and (7)) of the stack of the 93 halos. The *points* are the measured velocity anisotropy (same as in Fig. 11, again with uncertainties from 100 bootstraps on the 93 halos) and the *curves* are recovered from anisotropy inversion assuming the $c = 4$ NFW model (*solid curves*), for three polynomial fits (orders 2, 3, and 4 in log−log space) to the measured line-of-sight velocity dispersion profile (after the $\kappa = 2.7$ velocity cut using the NFW model with ML anisotropy: *solid dark red*, *dashed green*, and *dotted blue* for orders 2 to 4, respectively). The *purple dashed horizontal line* indicates the fully isotropic case. Note that the 4th order polynomial (*blue*) extrapolates poorly the line-of-sight velocity dispersion profile at very low and very high projected radii.

8. Summary and discussion

This work analyzes the distribution of particles in projected phase space around dark matter halos in cosmological simulations. The particles are split among halo particles within the virial sphere and interlopers within the virial cone but outside the virial sphere (Fig. 2). The reader should be careful that the analyses presented here cannot be directly applied to observations of clusters of galaxies, as they work with halo particles instead of galaxies within clusters, and assume the halo centers to be determined quite precisely (from real space measurements).

[18] We cannot employ the analytical approximation of Eq. (21) to the line-of-sight velocity dispersion profile for an NFW model with $r_{\rm a} = r_{-2}$ ML anisotropy, because we place ourselves in the context of an observer who wishes to measure the velocity anisotropy with no prior on it: (s)he is thus forced to use a smooth representation of the observed line-of-sight velocity dispersion profile.

We find a universal distribution of interlopers in projected phase space, i.e. with little dependence on halo mass (Figs. 7c and 9). In particular, we note that velocity cuts cannot distinguish the quarter of particles that are interlopers from those in the virial sphere (Fig. 13), as was previously noted by Cen (1997). We find that the distribution of interlopers in projected phase space displays a roughly constant surface density (Figs. 5 and 8) and a distribution of line-of-sight velocities that is the sum of a quasi-Gaussian component, caused by the halo outskirts (out to typically 8 virial radii, Fig. 7b) and a uniform component caused by particles at further distances from the halo (Figs. 6 and 7).

The cosmological simulations allow us to optimize the ratio of maximum velocity to line-of-sight velocity dispersion that recovers the latter quantity. Although this may seem to be a circular argument (since $\sigma_{\rm los}(R)$ depends on the velocity cut), it has been widely used in the past, usually in iterative form, with a 3σ cutoff. We find that this cutoff is not restrictive enough and causes an overestimate of the line-of-sight velocity dispersion profile (based upon mass and velocity anisotropy models derived from the cosmological simulations): up to 10% for the isotropic NFW velocity cut, which is reduced to 5% for the ML anisotropy velocity cut (Fig. 12). We recommend instead a velocity cut at $2.7\sigma_{\rm los}(R)$ on the best iterative fit to the line-of-sight velocity dispersion for the NFW model with $r_{\rm a} = r_{-2}$ ML anisotropy. Alternatively, one can use a velocity cut at $\kappa = 2.6$ for the $m = 5$ Einasto model, modeled (again) with $r_{\rm a} = r_{-2}$ ML anisotropy, but this underestimates the line-of-sight velocity dispersion near the virial radius (Fig. 12).

We illustrate (Figs. 3, 4, 6, 13, and 14) how the distribution of particles in projected phase space is altered once the high velocity interlopers are rejected with this new velocity filter (besides limiting the line-of-sight to typically $\pm 17\, r_{200}$, the main effect is to remove the flat velocity component). The fraction of interlopers within the virial cone drops from 27% (with an observer at distance $D = 90\, h^{-1}$ Mpc) to 23.1 ± 0.6% (independent of D for $D \gtrsim 17\, \langle r_{200} \rangle$) when the velocity cut is applied (where the uncertainty is taken from the end of Sect. 6).

This fraction of interlopers can be directly inferred from the NFW or Einasto model

$$f_{\rm i} = \frac{\hat{M}_{\rm p}(r_{200}) - 1 + \pi\, \hat{\Sigma}_{\rm bg}}{\hat{M}_{\rm p}(r_{200}) + \pi\, \hat{\Sigma}_{\rm bg}} \tag{27}$$

where $\hat{M}_{\rm p} = M_{\rm p}/M_{200}$ is the projected virial mass in virial units (i.e. in units of the mass within the virial sphere), while $\hat{\Sigma}_{\rm bg}$ is given in Eq. (25). For the NFW model, one then obtains $f_{\rm i} =$ 26.8% and 24.0%, respectively before ($\hat{\Sigma}_{\rm bg} = 0.0286$) and after ($\hat{\Sigma}_{\rm bg} = 0.0126$) the velocity cut, while with the $m = 5$ Einasto model with $c = 4$, the corresponding percentages of interlopers are 25.7% and 22.8%. These theoretical predictions are in excellent agreement with the fractions obtained from the simulations. Note that the omission of the background ($\hat{\Sigma}_{\rm bg}$) term in Eq. (27) reduces $f_{\rm i}$ by typically 10% in relative terms, relative to the fractions after the velocity cut. Applying Eq. (27) to models of different concentrations leads to roughly a power-law variation of $f_{\rm i}$ with slope −0.32 (NFW) or −0.49 ($m = 5$ Einasto). Therefore, *the fraction of interlopers should be (slightly) more important in the more massive halos*, since they have (slightly) lower concentrations (Navarro et al. 1997; Macciò et al. 2008).

In comparison, using 62 clusters from the same simulation as the one we have analyzed (we have 53 clusters in common), Biviano et al. (2006) found that among particles selected in cones of projected radius $1.5\, h^{-1}$ Mpc around cluster-mass halos (after their velocity cut), 18 ± 1.4% of them lie outside the sphere of the same radius[19]. Their halos have a median virial radius of $r_{200} = 0.93\, h^{-1}$ Mpc (1.08 times our median) and hence a virial mass of $M_{200} = 1.9 \times 10^{14}\, h^{-1}\, M_\odot$. Their Figure 7 indicates that their velocity cut is roughly 1180, 1105, and 780 km s^{-1}, at projected radii 0.6, 1.0 and $1.5\, h^{-1}$ Mpc, respectively. Since NFW concentration scales as $M^{-0.1}$ (Navarro et al. 1997; Macciò et al. 2008), their median concentration should be 3.9, hence their scale radius should be $930/3.9 = 238\, h^{-1}$ kpc. Assuming an NFW model, we deduce that their median circular velocity at the scale radius is 914 km s^{-1}, and find that their velocity cuts correspond to $\kappa = 2.1, 2.2$, and 1.8, at the three projected radii chosen above. These fractions are consistent with the values of $\hat{\kappa}$ one can read off of Fig. 3 of Wojtak et al. (2007) that illustrates the same velocity cut model (den Hartog & Katgert 1996). We then considered a cone of projected size $1.5/0.93 = 1.6\, r_{200}$. Adopting their typical $\kappa = 2$, we then found that after a 2σ velocity cut, the fraction of particles with $r > 1.6\, r_{200}$ is now 21.3%. This fraction is still marginally significantly larger than Biviano et al.'s fraction of 18% (assuming the same errors as above). We attribute this discrepancy to their variable κ velocity cut, which differs from our fixed κ one. Wojtak et al. (2007) tried several interloper removal schemes and definitions (using a different ΛCDM cosmological simulation). Their local 3σ cut leads to 20.4 ± 1.7% of interlopers remaining within the virial cone. Given the quoted errors, the lower fraction of interlopers found by Wojtak et al. is marginally consistent with ours.

This fraction of 23% of interlopers after the velocity cut is surprisingly close to the fraction of blue galaxies (i.e. galaxies off the Red Sequence) observed within SDSS clusters, as Yang et al. (2008) find roughly 22% of blue galaxies within SDSS clusters of masses $>10^{14}\, h^{-1}\, M_\odot$. Admittedly, it is dangerous to match the dark matter distribution with the galaxy distribution, since galaxies are biased tracers of the matter distribution. In fact, galaxies are biased relative to dark matter halos (e.g. Conroy et al. 2006), which in turn are biased relative to the dark matter particle distribution (e.g. Mo & White 1996; Catelan et al. 1998). If, in the end, the SDSS galaxies analyzed by Yang et al. are unbiased tracers of the dark matter distribution, then this close agreement would be expected if all blue galaxies are caused by projection effects. But if projections also pick up red galaxies in groups, then some blue galaxies would need to survive within the virial sphere for the match to hold. However, the Yang et al. group finder is fairly efficient in separating groups along the line-of-sight, so we conclude that the fraction of blue galaxies within the virial sphere should be small. In other words, star formation appears to be strongly quenched when galaxies penetrate the virial spheres of clusters.

When no velocity cut is performed, a maximum likelihood fit of the concentration of the projected NFW model to the projected radii of a stacked cluster of nearly 300 000 particles out to r_{200} (r_{100}) leads to a 14 ± 7% (22 ± 6%) underestimate of the true concentration parameter (Table 5, where most of the uncertainty comes from cosmic variance). Similar biases occur with the $m = 5$ Einasto model. But after the velocity cut, these biases decrease by a factor two, and are no longer statistically significant (Table 5). Moreover, the inclusion in such fits of a constant background as an extra parameter also strongly decreases the bias, even when the maximum projected radius is as low as r_{200} (Table 5). In fact, inspection of Table 5 indicates that, for $R_{\rm max} = r_{100}$ or $3\, r_{200}$, the background (fixed or free) has a greater influence than the velocity filter in removing the bias on

[19] The error is taken as their dispersion over the square root of their number of halos.

measured concentration. Surprisingly, for $R_{\rm max} = 3\,r_{200}$, a physically motivated fixed background added to the NFW model is slightly less effective in reducing the concentration bias than is a free background.

When the maximum radius is $3\,r_{200}$ and no velocity cut is performed, the NFW model with a free (respectively fixed) background underestimates the concentration (Table 5) by $2 \pm 11\%$ ($15 \pm 9\%$). This insignificant (marginally significant) bias is caused by the strong decrease of the surface density profile once the Hubble flow is added to the peculiar velocities (Fig. 16). These small biases suggest that the fairly low concentration ($c_{200} = 2.9 \pm 0.2$) for the galaxy distribution in clusters found by Lin et al. (2004), who fit an NFW model with a free constant background to the distribution of projected radii in the range $0.02 < R/r_{200} < 2.5$, but who did not make a velocity cut for lack of velocity data, is incompatible with true cluster concentrations of $c = 4.0$ at the 2σ level. The lower concentration bias with the two-component model is expected, because the single component NFW or Einasto models cannot capture the flat surface density at large radii (Fig. 14), because other halos are projected along the line-of-sight.

While a two-component model of halo (to infinity) + constant background is better able to recover the halo concentration than a single-component model (Table 5), it is not wise to estimate the halo concentration from a two-component model with a halo term whose line-of-sight is limited to the sphere (Appendix B) plus a near constant background term arising from our universal interloper surface density model (Eqs. (12) or simply (13)): the single-component NFW captures better the *total* surface density profile than this halo+background model, especially if the maximum projected radius is beyond the virial radius, as the interloper surface density has a discontinuous slope at the virial radius (Fig. 14). On the other hand, the universal distribution of interlopers in projected phase space might be useful to model the internal kinematics (hence total mass profile) of clusters of galaxies, where the full distribution of galaxies in projected phase space is the sum of these interlopers and an NFW-like model projected onto the virial sphere. We are preparing tests of the mass/anisotropy modeling of clusters, groups, and galaxies (through their satellites) using this interloper model.

We also performed 2D fits to individual halos of typically 700 particles (not shown here). The dispersion of the concentrations were much larger (typically 0.16 dex) than the biases obtained from the stacked virial cone (typically 0.05 dex, i.e. 10% errors, see Table 5), which means that shot noise and cosmic variance dominate the bias caused by the Hubble flow.

The line-of-sight velocity dispersion profile shows a concavity (in log-log) near the virial radius (Fig. 12), which is caused by the Hubble flow (Fig. 16). The velocity anisotropy profile recovered from this velocity dispersion profile, assuming the correct mass distribution, is close to the true anisotropy profile, with a slight, marginally significant, radial bias in the envelopes of clusters in comparison with the anisotropy profile recovered in 3D (Fig. 17), as was previously noted by Biviano (2007).

In summary, the density profile of ΛCDM halos falls fast enough that the effects of the Hubble flow perturbing the standard projection equations produce only small biases in comparison with the shot noise of clusters with less than 1000 galaxies, as well as the large cosmic variance of the halos.

These results have been obtained with the dark matter particles of a cosmological N-body simulation (with additional gas and galaxy components). They will need to be confirmed with future more realistic simulations of the *galaxy* distribution.

Acknowledgements. We thank Marisa Girardi for providing the positions and masses of the mock clusters, Mike Hudson and Raphael Gavazzi for helpful comments, and Richard Trilling for a critical reading of an early version of the manuscript. We also warmly thank an anonymous referee for his thorough reading of the manuscript and several important comments, especially his insistence on our use of cluster bootstraps to estimate the errors from cosmic variance. A.B. acknowledges the hospitality of the Institut d'Astrophysique de Paris. This research has been partly financially supported by INAF through the PRIN-INAF scheme. The simulation has been carried out at the Centro interuniversitario el Nord-Est per il Calcolo Elettronico (CINECA, Bologna) with CPU time assigned thanks to an INAF-CINECA grant.

References

Bartelmann, M. 1996, A&A, 313, 697
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Binney, J., & Mamon, G. A. 1982, MNRAS, 200, 361
Biviano, A. 2007, http://adlibitum.oats.inaf.it/biviano/sesto07.pdf
Biviano, A., & Girardi, M. 2003, ApJ, 585, 205
Biviano, A., Murante, G., Borgani, S., et al. 2006, A&A, 456, 23
Borgani, S., Murante, G., Springel, V., et al. 2004, MNRAS, 348, 1078
Catelan, P., Lucchin, F., Matarrese, S., & Porciani, C. 1998, MNRAS, 297, 692
Cen, R. 1997, ApJ, 485, 39
Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201
den Hartog, R., & Katgert, P. 1996, MNRAS, 279, 349
Einasto, J. 1965, Trudy Inst. Astroz. Alma-Ata, 51, 87
Fadda, D., Girardi, M., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1996, ApJ, 473, 670
Faltenbacher, A., & Diemand, J. 2006, MNRAS, 369, 1698
Gao, L., Navarro, J. F., Cole, S., et al. 2008, MNRAS, 387, 536
Haardt, F., & Madau, P. 1996, ApJ, 461, 20
Katgert, P., Biviano, A., & Mazure, A. 2004, ApJ, 600, 657
Kent, S. M., & Gunn, J. E. 1982, AJ, 87, 945
Lin, Y.-T., Mohr, J. J., & Stanford, S. A. 2004, ApJ, 610, 745
Łokas, E. L., & Mamon, G. A. 2001, MNRAS, 321, 155
Łokas, E. L., & Mamon, G. A. 2003, MNRAS, 343, 401
Łokas, E. L., Wojtak, R., Gottlöber, S., Mamon, G. A., & Prada, F. 2006, MNRAS, 367, 1463
Macciò, A. V., Dutton, A. A., & van den Bosch, F. C. 2008, MNRAS, 391, 1940
Mamon, G. A., & Boué, G. 2010, MNRAS, 401, 2433
Mamon, G. A., & Łokas, E. L. 2005a, MNRAS, 362, 95
Mamon, G. A., & Łokas, E. L. 2005b, MNRAS, 363, 705
Mauduit, J.-C., & Mamon, G. A. 2007, A&A, 475, 169
Merritt, D. 1985, MNRAS, 214, 25P
Miralda-Escude, J. 1991, ApJ, 370, 1
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Navarro, J. F., Hayashi, E., Power, C., et al. 2004, MNRAS, 349, 1039
Osipkov, L. P. 1979, SvA Lett., 5, 42
Prada, F., Klypin, A. A., Simonneau, E., et al. 2006, ApJ, 645, 1001
Prugniel, P., & Simien, F. 1997, A&A, 321, 111
Saro, A., Borgani, S., Tornatore, L., et al. 2006, MNRAS, 373, 397
Solanes, J. M., & Salvador-Solé, E. 1990, A&A, 234, 93
Springel, V. 2005, MNRAS, 364, 1105
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289
Tremaine, S., Richstone, D. O., Byun, Y.-I., et al. 1994, AJ, 107, 634
van der Marel, R. P., Magorrian, J., Carlberg, R. G., Yee, H. K. C., & Ellingson, E. 2000, AJ, 119, 2038
Wainer, H., & Thissen, D. 1976, Psychometrica, 41, 9
Wojtak, R., & Łokas, E. L. 2010, MNRAS, in press [arXiv:1004.3771]
Wojtak, R., Łokas, E. L., Mamon, G. A., et al. 2007, A&A, 466, 437
Yahil, A., & Vidal, N. V. 1977, ApJ, 214, 347
Yang, X., Mo, H. J., & van den Bosch, F. C. 2008, ApJ, 676, 248

Pages 18 to 22 are available in the electronic edition of the journal at http://www.aanda.org

Appendix A: Projected mass, surface density and tangential shear of the Einasto model

In this appendix, we derive an approximation to the surface density and projected mass (or, equivalently, projected number) profiles for the Einasto model.

A.1. Projected mass profile

For any density model, the projected mass is

$$\begin{aligned} M_{\rm p}(R;m) &= \int_0^R 2\,\pi\,S\,\Sigma(S\,;m)\,{\rm d}S \\ &= 4\,\pi\left[\int_0^R r\,\rho(r)\,{\rm d}r\int_0^r \frac{S\,{\rm d}S}{\sqrt{r^2-S^2}} + \int_R^\infty r\,\rho(r)\,{\rm d}r\int_0^R \frac{S\,{\rm d}S}{\sqrt{r^2-S^2}}\right] \\ &= 4\,\pi\left[\int_0^R r^2\,\rho(r)\,{\rm d}r + \int_R^\infty r\left(r-\sqrt{r^2-R^2}\right)\rho(r)\,{\rm d}r\right] \qquad \text{(A.1)} \\ &= M_\infty - 4\,\pi\int_R^\infty r\,\sqrt{r^2-R^2}\,\rho(r)\,{\rm d}r, \qquad \text{(A.2)} \end{aligned}$$

where the second equality is obtained after reversing the order of integration. Equation (A.1) is general, while Eq. (A.2) is only valid for models with finite total mass M_∞. For the Einasto model of total mass M_∞, the 3D mass profile is

$$M(r;m) = P\left[3m, 2m\left(\frac{r}{r_{-2}}\right)^{1/m}\right] M_\infty, \qquad \text{(A.3)}$$

where $P(a,x) = \gamma(a,x)/\Gamma(a)$ is the regularized incomplete gamma function. The ratio

$$\mu(R,m) = \frac{M_{\rm p}(R;m)}{M(R;m)}, \qquad \text{(A.4)}$$

determined from Eqs. (A.3) and (A.2), with Eq. (15), varies little, as seen in the right panel of Fig. A.1. We fit again a two-dimensional fourth-order polynomial in m and $u = \log_{10}(R/r_{-2})$ and find

$$\begin{aligned} \mu(R,m) &\simeq \mu_{\rm apx}(u,m) \qquad \text{(A.5)} \\ \mu_{\rm apx}(u,m) &= {\rm dex}\big(0.0001219\,m^4 + 0.0007400\,m^3u - 0.003209\,m^3 + 0.002976\,m^2u^2 - 0.01560\,m^2u \\ &\quad +0.02966\,m^2 + 0.0003307\,m\,u^3 - 0.04434\,m\,u^2 + 0.1273\,m\,u - 0.1149\,m \\ &\quad +0.001036\,u^4 - 0.003133\,u^3 + 0.1905\,u^2 - 0.5241\,u + 0.3525\big). \qquad \text{(A.6)} \end{aligned}$$

In the interval $3.5 \le m \le 6.5$ and $-2 \le u \le 2$, Eqs. (A.5) and (A.6) are accurate to better than 1.5% everywhere (0.23% rms).

The projected mass of the Einasto model can thus be written

$$\begin{aligned} M_{\rm p}(R;m) &\simeq \mu_{\rm apx}(u,m)\,M(R;m) \qquad \text{(A.7)} \\ &= \mu_{\rm apx}\left[\log_{10}\left(\frac{R}{r_{-2}}\right), m\right] P\left[3m, 2\,m\left(\frac{R}{r_{-2}}\right)^{1/m}\right] M_\infty, \qquad \text{(A.8)} \end{aligned}$$

where again $u = \log_{10}(R/R_{-2})$.

A.2. Surface density profile

Inserting Eqs. (15) into (1), the surface density of the Einasto model of total mass M and index m is

$$\begin{aligned} \Sigma(R;m) &= \frac{M\,(r_{-2})}{\pi r_{-2}^2}\,\widetilde{\Sigma}\left(\frac{R}{r_{-2}};m\right), \qquad \text{(A.9)} \\ \widetilde{\Sigma}(X;m) &= \frac{(2m)^{3m-1}}{\gamma(3m,2m)}\int_X^\infty \exp\left(-2\,m\,x^{1/m}\right)\frac{x\,{\rm d}x}{\sqrt{x^2-X^2}}. \qquad \text{(A.10)} \end{aligned}$$

Writing the dimensionless mass density as

$$\widetilde{\rho}(x;m) = \frac{\rho(xr;m)}{M\,(r_{-2};m)\,/\left(4\pi r_{-2}^3\right)} = \frac{(2m)^{3m}}{m\,\gamma(3m,2m)}\,\exp\left(-2m\,x^{1/m}\right), \qquad \text{(A.11)}$$

where the second equality derives from Eq. (15), we can express the ratio of dimensionless surface to space densities as

$$\mathcal{R}(X,m) = \frac{\widetilde{\Sigma}(X;m)}{\widetilde{\rho}(X;m)} = \frac{1}{2}\,\exp\left(2\,m\,X^{1/m}\right)\int_X^\infty \exp\left(-2m\,x^{1/m}\right)\frac{x\,{\rm d}x}{\sqrt{x^2-X^2}}, \qquad \text{(A.12)}$$

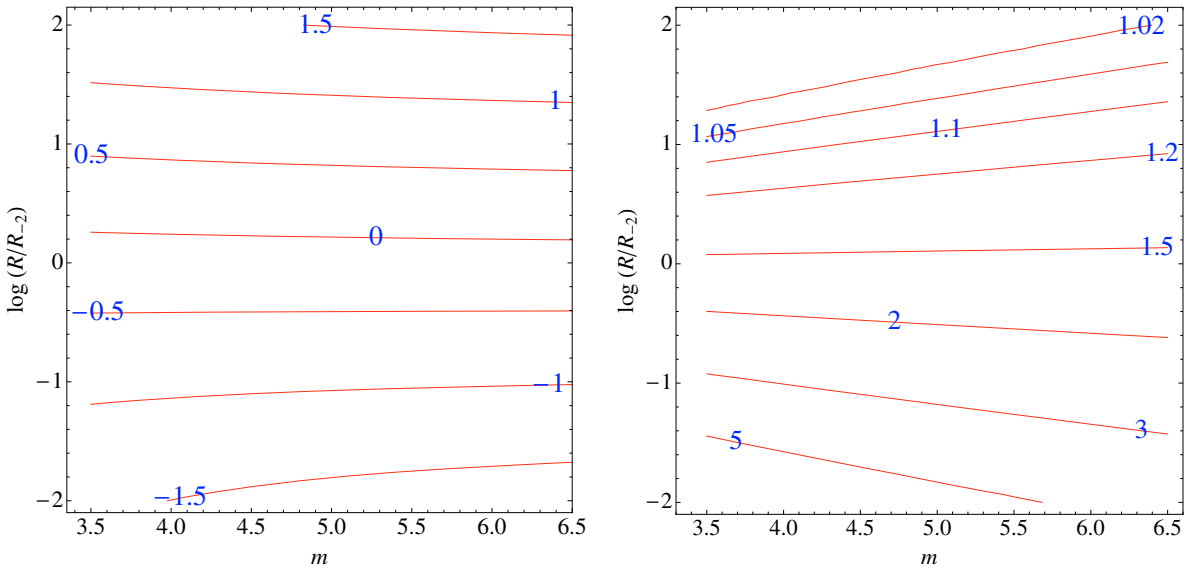


Fig. A.1. Contours of $\log_{10}\mathcal{R}$ (*left*, Eq. (A.12)) and μ (*right*, Eq. (A.4), with Eqs. (15), (A.2) and (A.3)) for the Einasto model.

where $X = R/r_{-2}$. In the range $3.5 \le m \le 6.5$ (spanned by ΛCDM halos in the redshift range $0 \le z \le 3$ according to Gao et al. 2008) and $-2 \le \log_{10} X \le 2$, $\mathcal{R}$ varies little and regularly, as seen in the left panel of Fig. A.1. We fit a two-dimensional 4th-order polynomial in m and $u = \log_{10}(R/r_{-2})$ to $\log_{10}\mathcal{R}$. We find

$$\mathcal{R}(X,m) \simeq \mathcal{R}_{\rm apx}(u,m) \tag{A.13}$$

$$\begin{aligned}\mathcal{R}_{\rm apx}(u,m) = \mathrm{dex}\Big(&6.286\times10^{-6}\,m^4 + 0.001178\,m^3u - 0.0002251\,m^3 + 0.001524\,m^2u^2 - 0.02427\,m^2u\\ &+0.0008538\,m^2 + 0.001861\,m\,u^3 - 0.02323\,m\,u^2 + 0.1849\,m\,u + 0.01577\,m\\ &+0.0006014\,u^4 - 0.01506\,u^3 + 0.1056\,u^2 + 0.3406\,u - 0.2515\Big).\end{aligned} \tag{A.14}$$

In the interval $3.5 \le m \le 6.5$ and $-2 \le u \le 2$, Eqs. (A.13) and (A.14) are accurate to better than 0.8% everywhere (0.12% rms).

The dimensionless surface density can then be written as

$$\widetilde{\Sigma}(X;m) = \frac{(2\,m)^{3m}}{m\,\gamma(3\,m,2\,m)}\,\exp\left(-2\,m\,X^{1/m}\right)\mathcal{R}_{\rm apx}(\log_{10}X,m), \tag{A.15}$$

or equivalently, with $M\,(r_{-2}) = P(3m,2m)\,M_\infty$, where $P(a,x) = \gamma(a,x)/\Gamma(a)$ is the regularized incomplete gamma function and M_∞ the total mass of the Einasto model:

$$\frac{\Sigma(R)}{M_\infty/\left(\pi r_{-2}^2\right)} = \frac{(2\,m)^{3m}}{m\,\Gamma(3\,m)}\,\exp\left(-2\,m\,X^{1/m}\right)\mathcal{R}_{\rm apx}(\log_{10}X,m). \tag{A.16}$$

Alternatively, the surface density profile can be, self-consistently, estimated from Eq. (A.7) by differentiation over the projected mass profile, yielding after some algebra

$$\begin{aligned}\Sigma(R;m) &= \frac{1}{2\pi R}\,\frac{\mathrm{d}M_{\rm p}(R;m)}{\mathrm{d}R}\\ &\simeq \frac{\mu_{\rm apx}(u,m)}{2\pi R^2}\left[4\pi R^3\,\rho(R;m) + \frac{\mathrm{d}\log_{10}\mu_{\rm apx}}{\mathrm{d}u}\,M(R;m)\right]\\ &= \left[\frac{M\,(r_{-2})}{\pi r_{-2}^2}\right]\left[X^3\,\widetilde{\rho}(X;m) + \frac{\mathrm{d}\log\mu_{\rm apx}}{\mathrm{d}\log X}\,\frac{P\left(3m,2m\,X^{1/m}\right)}{P(3m,2m)}\right]\frac{\mu_{\rm apx}(u,m)}{2\,X^2},\end{aligned} \tag{A.17}$$

where $\widetilde{\rho}$ is given in Eq. (A.11).

Equation (A.17) has the advantage of providing an approximation for the surface density profile that is consistent with that of the projected mass profile. This is crucial for maximum likelihood estimation of concentration (and possibly Einasto index and background level). On the other hand, the accuracy of Eq. (A.17) is about 5 times worse than that of Eq. (A.15).

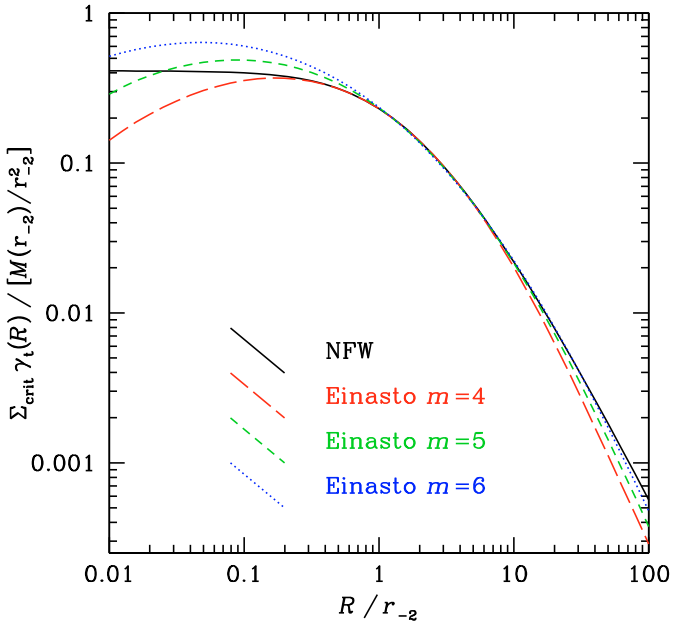


Fig. A.2. Dimensionless tangential shear profile for the NFW model (*black*) and the $m = 4$ (*red, long-dashed*), 5 (*green, short-dashed*) and 6 (*blue, dotted*) Einasto models, using Eq. (A.18) with Eqs. (A.9), (A.14), (A.15), (A.6), and (A.8).

A.3. Tangential shear profile

For any density model, the tangential shear measured by weak lensing can be written (e.g. Miralda-Escude 1991)

$$\gamma_{\rm t}(R;m) = \frac{\overline{\Sigma}(R;m) - \Sigma(R;m)}{\Sigma_{\rm crit}}, \tag{A.18}$$

where $\overline{\Sigma}(R;m) = M_{\rm p}(R;m)/(\pi R^2)$ is the mean surface density, while $\Sigma_{\rm crit} = c^2/(4\pi G)\,D_{\rm S}/(D_{\rm L}D_{\rm LS})$ is the critical surface density, with c the velocity of light, and where $D_{\rm S}$, $D_{\rm L}$, and $D_{\rm LS}$ are the angular diameter distances between the observer and the source, the observer and the lens, and the lens and the source, respectively. Equation (A.18) indicates that adding a constant term to the surface density (Eq. (24)) has no effect on $\gamma_{\rm t}$ (this is the mass-sheet degeneracy). For the Einasto model, the tangential shear (Eq. (A.18)) is readily computed using Eqs. (A.15) with (A.14) and (A.8) with (A.6). Figure A.2 shows the subtle differences in the shear profile between the NFW and Einasto models of index $m = 4$, 5, and 6. While the tangential shear of the four models is indistinguishable in the wide range $0.8 < R/r_{-2} < 10$, there are potentially measurable differences at $R > 10\,r_{-2}$ (at $100\,r_{-2}$, the NFW shear is 1.5 times greater than that for the $m = 5$ Einasto model) and possibly at $R < 0.8\,r_{-2}$ (as long as the weak linear approximation assumed for the measured shear to match the expression of $\gamma_{\rm t}$ of Eq. (A.18) remains valid).

Appendix B: Surface density and projected mass of the NFW model with lines of sight limited to a sphere

In this appendix, we derive the surface density and projected mass (or, equivalently, projected number) profiles of the NFW model, with the lines of sight restricted to a sphere (which we conveniently choose as the virial sphere) instead of extending to infinity.

B.1. Surface density profile

In an analogous manner as the case with line-of-sight extending to infinity (Eq. (1)), the surface density at projected radius R within the sphere of radius $r_{\rm max}$ is

$$\Sigma^{\rm sph}(R;r_{\rm max}) = 2\int_R^{r_{\rm max}} \rho(r)\frac{r\,{\rm d}r}{\sqrt{r^2-R^2}}\cdot \tag{B.1}$$

We now consider the case of the virial sphere: $r_{\rm max} = r_v$. The surface density can then be written

$$\Sigma^{\rm sph}(R;r_v) = \frac{M(r_{-2})}{\pi r_{-2}^2}\,\widetilde{\Sigma}^{\rm sph}\left(\frac{R}{r_{-2}},\frac{r_v}{r_{-2}}\right), \tag{B.2}$$

where

$$\widetilde{\Sigma}^{\rm sph}(X,c) \;=\; \frac{1}{2\,\ln 2-1}\int_X^c \frac{{\rm d}x}{(1+x)^2\,\sqrt{x^2-X^2}} \tag{B.3}$$

$$= \frac{1}{2\,\ln 2-1}\begin{cases} \dfrac{1}{(1-X^2)^{3/2}}\,\cosh^{-1}\left[\dfrac{c+X^2}{(c+1)\,X}\right] - \dfrac{1}{(c+1)}\,\dfrac{\sqrt{c^2-X^2}}{1-X^2} & 0<X<1, \\ \dfrac{\sqrt{c^2-1}(c+2)}{3(c+1)^2} + \dfrac{\left(-2c^3-4c^2-c+2\right)(X-1)}{5(c+1)^2\,\sqrt{c^2-1}} & X=1<c, \\ \dfrac{1}{(c+1)}\,\dfrac{\sqrt{c^2-X^2}}{X^2-1} - \dfrac{1}{(X^2-1)^{3/2}}\,\cos^{-1}\left[\dfrac{c+X^2}{(c+1)\,X}\right] & 1<X<c\,, \\ 0 & X=0 \text{ or } X>c, \end{cases} \tag{B.4}$$

where Eq. (B.3) is found by inserting the NFW density profile (Eqs. (4)) into (B.1).

B.2. Projected mass profile

For the NFW model, the projected mass within the virial sphere is

$$M_{\rm p}^{\rm sph}(R;r_v) = \int_0^R 2\,\pi\,S\,\Sigma^{\rm sph}(S\,;r_v)\,{\rm d}S \;=\; M(r_{-2})\,\widetilde{M}_{\rm p}^{\rm sph}\left(\frac{R}{r_{-2}},\frac{r_v}{r_{-2}}\right), \tag{B.5}$$

where

$$\widetilde{M}_{\rm p}^{\rm sph}(X,c) \;=\; 2\int_0^X Y\,\widetilde{\Sigma}^{\rm sph}(Y,c)\,{\rm d}Y \tag{B.6}$$

$$= \frac{1}{\ln 2-1/2}\begin{cases} 0 & X=0, \\ \dfrac{\sqrt{c^2-X^2}-c}{c+1} + \ln\left[\dfrac{(c+1)\left(c-\sqrt{c^2-X^2}\right)}{X}\right] + \dfrac{1}{\sqrt{1-X^2}}\cosh^{-1}\left[\dfrac{c+X^2}{(c+1)X}\right] & 0<X<1 \text{ and } X<c, \\ \dfrac{\sqrt{c^2-X^2}-c}{c+1} + \ln\left[\dfrac{(c+1)\left(c-\sqrt{c^2-X^2}\right)}{X}\right] + \dfrac{1}{\sqrt{X^2-1}}\cos^{-1}\left[\dfrac{c+X^2}{(c+1)X}\right] & 1<X<c, \\ \ln\left[(c+1)\left(c-\sqrt{c^2-1}\right)\right] - \dfrac{c}{c+1} + 2\sqrt{\dfrac{c-1}{c+1}} & 1=X<c, \\ \ln(c+1) - \dfrac{c}{c+1} & X\ge c, \end{cases} \tag{B.7}$$

where Eq. (B.7) was found by inserting Eq. (B.4) into Eq. (B.6). For $X \ge c$, one recovers the mass within the virial sphere.

Appendix C: Maximum likelihood estimates

In this appendix, we illustrate the maximum likelihood calculations that we have performed.

Given parameters $\boldsymbol{\theta}$, and data points $\mathbf{x}$ the MLE is found by minimizing

$$-\ln\mathcal{L} = -\sum_j \ln p(x_j|\boldsymbol{\theta}), \tag{C.1}$$

where $\mathcal{L} = \prod_j p(x_j|\boldsymbol{\theta})$ is the likelihood.

C.1. Density profile

The probability of measuring an object (galaxy or dark matter particle) at radius r in a spherical model of concentration c is

$$p(r_j|c) = \frac{4\pi r^2\,\left[\nu(r_j;c)+b\right]}{N(r_{\rm max};c)-N(r_{\rm min};c)+4\,\pi\,b\,\left(r_{\rm max}^3-r_{\rm min}^3\right)/3}, \tag{C.2}$$

where $\nu(R)$ and $N(R)$ are respectively the density and number (proportional to mass) profiles, $r_{\rm min}$ and $r_{\rm max}$ are respectively the minimum and maximum radii, c is the concentration, while b is the constant density background.

C.2. Surface density profile

The probability of measuring a galaxy at projected radius R in a spherical model of concentration c and background b is

$$p(R_j|c,b) = \frac{2\pi R_j \left[\Sigma(R_j;c) + \Sigma_{\rm bg}\right]}{N_{\rm p}(R_{\rm max};c) - N_{\rm p}(R_{\rm min};c) + \pi\Sigma_{\rm bg}\left(R^2_{\rm max} - R^2_{\rm min}\right)}, \tag{C.3}$$

where $\Sigma(R)$ and $N_{\rm p}(R)$ are respectively the surface density and projected number (proportional to projected mass) profiles, $R_{\rm min}$ and $R_{\rm max}$ are respectively the minimum and maximum projected radii, c is the concentration, while $\Sigma_{\rm bg}$ is the constant surface density background.

For the surface density profile $\Sigma(R)$ and the projected number (mass) profile $N_{\rm p}(R)$, we use the formulae of Łokas & Mamon (2001) and of Appendix A for the NFW and Einasto models, respectively.

C.3. Distribution of interloper velocities

According to Eq. (8), the distribution of interloper line-of-sight *absolute* velocities, $v_j \equiv |v_{{\rm los},j}|$, is to first order the sum of a Gaussian and a constant term:

$$p(v_j|\sigma_{\rm i},A,B) = \frac{A\,\exp\left[-v_j^2/\left(2\sigma_{\rm i}^2\right)\right] + B}{\sqrt{\pi/2}\,A\,\sigma_{\rm i}\,{\rm erf}\left[\hat{\kappa}/(\sigma_{\rm i}\sqrt{2})\right] + \hat{\kappa}\,B}, \tag{C.4}$$

where the denominator is found by ensuring $\int_0^{\hat{\kappa}} p(v_j)\,{\rm d}v_j = 1$ (Eq. (12)), and where $\hat{\kappa}$ is the maximum considered value of $|v_{\rm los}|/v_{\rm v}$ (so $\hat{\kappa} = 4$ in Figs. 6). If A and B are expressed in virial units, then the denominator of Eq. (C.4) is the surface density of particles under consideration in virial units, which we directly measure from the simulation as $\Sigma = (N/N_v)/S$, where N_v is the number of particles within the virial sphere, while N is the number of particles in the radial bin (or within the full virial cone), and S is the surface of the radial bin (i.e. π for the full virial cone). Hence, substituting for $A = \Sigma\,(1 - \hat{\kappa}\,B')/[\,\sqrt{\pi/2}\sigma_{\rm i}\,{\rm erf}[\hat{\kappa}/(\sigma_{\rm i}\sqrt{2})]$, we can write the probability of measuring an interloper absolute velocity as

$$p(v_j|\sigma_{\rm i},B) = \frac{(1 - \hat{\kappa}\,B')\,\exp\left[-v_j^2/\left(2\sigma_{\rm i}^2\right)\right]}{\sqrt{\pi/2}\,\sigma_{\rm i}\,{\rm erf}\left[\hat{\kappa}/(\sigma_{\rm i}\sqrt{2})\right]} + B', \tag{C.5}$$

where $B' = B/\Sigma$. Then given the respective uncertainties $\epsilon(\sigma_{\rm i})$ and $\epsilon(B')$ in $\sigma_{\rm i}$ and B', we deduce the uncertainties in B and A as

$$\epsilon(B) = \Sigma\,\epsilon(B'), \tag{C.6}$$

$$\epsilon(A) = \sqrt{\left(\frac{\partial A}{\partial\sigma_{\rm i}}\right)^2 \epsilon^2(\sigma_{\rm i}) + \left(\frac{\partial A}{\partial B'}\right)^2 \epsilon^2(B')} = \Sigma\,\frac{\sqrt{2\hat{\kappa}^2\,\pi\,\sigma_{\rm i}^4\mathcal{E}^2\,\epsilon^2(B') + (1 - \hat{\kappa}\,B')^2\,\left\{\sqrt{2\pi}\,\sigma_{\rm i}\,\mathcal{E} - 2\hat{\kappa}\,\exp\left[-\hat{\kappa}^2/\left(2\sigma_{\rm i}^2\right)\right]\right\}^2\,\epsilon^2\,(\sigma_{\rm i})}}{\pi\,\sigma_{\rm i}^3\,\mathcal{E}^2}, \tag{C.7}$$

where $\mathcal{E} = {\rm erf}\left[\hat{\kappa}/(\sigma_{\rm i}\sqrt{2})\right]$.

C.4. Practical considerations

For one-parameter fits, we first search on a wide linear grid of equally-spaced 11 points for θ_j, then we consider the three points with the lowest values of $\ln\mathcal{L}$ and create a subgrid of 11 equally-spaced points (thus typically zooming in by a factor of 5), and iterate with finer subgrids until the two values of the parameter θ_j with the highest likelihoods differ by less than 0.0001 or when the lowest $-\ln\mathcal{L}$ decreases by less than 10^{-12}. We then obtain the 1σ confidence interval fitting a cubic spline to the points below and above the best-fit parameter to solve for $-\ln\mathcal{L} = -\ln\mathcal{L}_{\rm ML} + 0.5$.

For two-parameter (three-parameter) fits, we first search on a wide rectangular (cuboidal) grid of equally-spaced 11 points. Then we consider the rectangle (cuboid) obtained by searching for the lowest values of $-\ln\mathcal{L}$, such that there are at least 3 different values for both (all three) parameters. We create a sub-grid in this rectangle (cuboid) with again 11×11 ($11\times11\times11$) points, and iterate with finer subgrids until the pair of each of the two (three) parameters with the highest two likelihoods differ by less than 0.0001 or when the lowest $-\ln\mathcal{L}$ decreases by less than 10^{-12}. We then obtain the 1σ contour by considering those points in parameter space for which $-\ln\mathcal{L} = -\ln\mathcal{L}_{\rm ML} + 1.15$ (1.77), and then define as the minimum and maximum values for each parameter the extreme values in this contour.