Mon. Not. R. Astron. Soc. **000**, 1–10 () Printed 26 June 2000 (MN LATEX style file v1.4)

Mapping the substructure in the Galactic halo with the next generation of astrometric satellites

Amina Helmi and P. Tim de Zeeuw
Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands

Accepted ... Received ...; in original form ...

ABSTRACT
We run numerical simulations of the disruption of satellite galaxies in a Galactic potential to build up the entire stellar halo, in order to investigate what the next generation of astrometric satellites will reveal by observing the halo of the Milky Way. We generate artificial DIVA, FAME and GAIA halo catalogues, in which we look for the signatures left by the accreted satellites. We develop a method based on the standard Friends-of-Friends algorithm applied to the space of integrals of motion. We find this simple method can recover about 50% of the different accretion events, when the observational uncertainties expected for GAIA are taken into account, even when the exact form of the Galactic potential is unknown. The recovery rate for DIVA and FAME is much smaller, but these missions, like GAIA, should be able to test the hierarchical formation paradigm on our Galaxy by measuring the amount of halo substructure in the form of nearby kinematically cold streams with for example, a two–point correlation function in velocity space.

Key words: The Galaxy: formation, kinematics and dynamics, halo

1 INTRODUCTION

Hierarchical theories of structure formation in the Universe propose that galaxies are the result of mergers and accretion of smaller building blocks (White & Rees 1978). Detailed studies of the properties of a galaxy built in this way have shown that such events leave fossil signatures in the present day components, which for a galaxy like our own would be clearly detectable with future astrometric missions (Helmi & White 1999). In particular the stellar halo would be the natural place to look for such substructures, since a spheroidal component is formed by the trails of stars left by disrupted satellite galaxies. Moreover, recent observations have shown that indeed considerable structure is still present in Milky Way's halo, indicating that accretion events have had some role in its formation history (e.g. Ibata, Gilmore & Irwin 1994; Majewski, Munn & Hawley 1996; Helmi et al. 1999; Ivezic et al. 2000).

In the next ten years, several satellite missions will be devoted to measure with very high accuracy the motions of thousands to many millions of stars in our Galaxy. NASA's Space Interferometry Mission (SIM) is a targeted mission which will obtain parallaxes and proper motions for about 10000 stars. With somewhat different goals, and more similar to the HIPPARCOS satellite, the Full-sky Astrometric Mapping Explorer (FAME, Horner et al. 1999) promises to measure positions and parallaxes for stars brighter than $V \sim 9$ to better than 50 μas and proper motions to 50 μas yr^{-1}. At $V \sim 15$ these accuracies will be degraded by an order of magnitude. The resulting astrometric database will have 4×10^7 stars, and may be combined with the radial velocities from the Sloan Digital Sky Survey or from other ground based catalogues to obtain full phase-space information. Less ambitious but still an improvement over HIPPARCOS is the German DIVA mission (Röser 1998). If launched it will observe of the order of 3.5×10^7 stars, at four times the precision of HIPPARCOS ($\sigma_\pi = 0.25$ mas and $\sigma_\mu = 0.4$ mas yr^{-1} at $V = 10$), thereby completing the knowledge of nearby stars. Like FAME, DIVA will not measure radial velocities. On the other hand, the proposed ESA astrometric satellite GAIA (Gilmore et al. 1998) will provide very precise astrometry (<10 μas in parallax and <10 μas yr^{-1} in proper motion at $V \sim 15$, increasing to 0.2 mas yr^{-1} at $V \sim 20$) and multicolour photometry, for all 1.3 billion objects to $V \sim 20$, *and* radial velocities with accuracies of a few km s^{-1} for most stars brighter than $V \sim 17$, so that full and homogeneous six-dimensional phase-space information will be available. These satellite missions will thereby provide a very large and statistically reliable sample of stars, from which the fundamental questions concerning the origin and evolution of the Galaxy may finally be answered.

In this paper we shall focus on what GAIA will tell us about the history and formation of the stellar halo of the Milky Way. We will also discuss the impact of DIVA and FAME, and leave aside SIM as this mission will not provide a survey but a hand-picked catalogue of stars. Even though we focus on the stellar halo, the method that we shall propose for finding substructures in phase-space may also be extended to find, for example, disk moving groups (e.g. de Zeeuw et al. 1999; Chereul, Crézé & Bienaymé 1999).

There are several methods for detecting moving groups. The Great Circle Cell Counts method (G3C) proposed by Johnston, Hernquist & Bolte (1996) uses the position on the sky, and employs the fact that satellite galaxies in orbits that probe only the outer (spherical) halo conserve the orientation of their plane of motion, thereby leaving their debris along great circles on the sky, if observed from the Galactic centre. The methods used in the Solar neighbourhood for detection of disk moving groups and open clusters use also proper motions (and sometimes parallax), and assume that all the stars belonging to the same system have the same velocity vector (e.g. Hoogerwerf & Aguilar 1998; de Bruijne 1998). Lynden-Bell & Lynden-Bell's method (1995) needs the position on the sky and the radial velocity, and has been used, for example, to link globular clusters which lie in the same plane to some of the (disrupted) dwarf companions of our Galaxy (see also Lynden-Bell, 1999). The applicability of the above mentioned methods is questionable in the inner parts of the halo. In this regime, the Galactic potential is significantly flattened so that the debris does not remain on a fixed plane, the situation where G3C works. As noted by Helmi & White (1999) no spatial correlations should be expected for satellites disrupted several Gyr ago. On the other hand, even though the velocity dispersions in a stellar stream do decrease with time, and therefore, very strong correlations are to be expected, in the inner halo strong phase-mixing takes place. For example in the Solar neighbourhood several hundred (mainly) cold streams originating in disrupted satellites may be present, but it may in practice be difficult to resolve each one of such moving groups completely. Clearly, before exploring the full capabilities of the next generation of astrometric satellite missions, we first need to identify where the clustering that is characteristic of a satellite manifests itself in the debris that we observe after many galactic orbits. As shown in Helmi et al. (1999) a method based on the lumpiness in integrals of motion space seems to be a promising tool for unveiling the merger history of our Galaxy.

2 BUILDING UP A STELLAR HALO

Our main goal is to test whether with the next generation of astrometric satellites, we would be able to find the signatures left by merger events in the Galactic stellar halo. We will assume that the whole stellar halo is the result of the superposition of several disrupted satellite galaxies which fell onto the Milky Way about 10 Gyr ago. We shall here discuss the initial conditions and the numerical methods used to generate this version of the stellar halo.

2.1 Initial conditions for the satellites

2.1.1 Orbital properties

The stellar halo has a density profile (Kinman, Suntzeff & Kraft 1994)

$$\rho_\star(r) = \rho_0 \left(\frac{r}{r_0}\right)^{-3.5}, \tag{1}$$

a total luminosity of about $10^9\,\mathrm{L}_\odot$, and a half light radius which probably lies around 3 kpc from the Galactic centre. For $r = r_0 = 8$ kpc (the distance to the Galactic Centre from the Sun), ρ_0 corresponds to the local stellar halo density, for which we take $\rho_0 = 1.5 \times 10^4\,\mathrm{M}_\odot\,\mathrm{kpc}^{-3}$ (Fuchs & Jahreiß 1998).

The initial orbital conditions of our satellites should be drawn from the Galactic halo distribution function (DF), which we assume to be a function of energy E and angular momentum L: $f(E, L)$. For simplicity here we shall assume that the stellar halo is a power-law tracer population embedded in a singular isothermal sphere, representing the dark-matter halo of the Milky Way. Following van den Bosch et al. (1998), we assume that

$$f(E, L) = g(E) h_\alpha(\eta), \qquad \text{where} \qquad \eta = L/L_c(E),$$

and $L_c(E)$ is the angular momentum of a circular orbit with energy E: $L_c(E) = r_c(E) V_c$, where $r_c(E) = \mathrm{e}^{-1/2} \exp\left[E/V_c^2\right]$. The function $h_\alpha(\eta)$ is known as the circularity, and determines the degree of anisotropy of the DF. We choose a simple parametrization of $h_\alpha(\eta)$ (Gerhard 1991):

$$h_\alpha(\eta) = \begin{cases} \tanh\left(\frac{\eta}{\alpha}\right) / \tanh\left(\frac{1}{\alpha}\right) & \alpha > 0 \\ 1 & \alpha = 0 \\ \tanh\left(\frac{1-\eta}{\alpha}\right) / \tanh\left(\frac{1}{\alpha}\right) & \alpha < 0 \end{cases} \tag{2}$$

so that for $\alpha = 0$, the DF is isotropic, for $\alpha < 0$ it is radially anisotropic and for $\alpha > 0$ is is tangentially anisotropic. We shall take $\alpha = -0.5$, since the halo appears to be radially anisotropic.

For a singular isothermal sphere

$$\rho(r) = \frac{V_c^2}{4\pi G r^2}, \qquad \phi(r) = V_c^2 \ln \frac{r}{r_s} \tag{3}$$

The corresponding DF is

$$g(E) = \frac{\mathrm{e}}{16\pi^2 G V_c \kappa} \exp\left[-\frac{2E}{V_c^2}\right], \tag{4}$$

(Gerhard 1991) where

$$\kappa = \int_0^\infty \mathrm{d}u\, \mathrm{e}^{-u} \int_0^{\eta_{\max}} h_\alpha(\eta) \frac{\eta \mathrm{d}\eta}{\sqrt{\eta_{\max}^2 - \eta^2}}. \tag{5}$$

Here $\eta_{\max} = \sqrt{2\mathrm{e}}\sqrt{u}\mathrm{e}^{-u}$, with $u = (E - \phi)/V_c^2$. Since the density profile may be derived from the initial distribution function as

$$\rho(r) = \frac{4\pi}{r} \int_{\phi(r)}^\infty \mathrm{d}E g(E) L_c(E) \int_0^{\eta_{\max}} h_a(\eta) \frac{\eta \mathrm{d}\eta}{\sqrt{\eta_{\max}^2 - \eta^2}}, \tag{6}$$

the joint probability distribution of E and η at a given radius r_p is

$$P(E, \eta) = \frac{4\pi}{r_p \rho(r_p)} g(E) L_c(E) \frac{\eta h(\eta)}{\sqrt{\eta_{\max}^2 - \eta^2}},$$

(van der Marel, Sigurdsson & Hernquist 1997). Using Eq.(4), we find that the normalised cumulative probability distribution of E is

$$\hat{P}(< E) = 1 - \exp\left[-\frac{E - \phi}{V_c^2}\right]. \tag{7}$$

We may derive the initial positions of the satellites by assuming the profile given in Eq.(1), and using, from Eq.(7), the energy as $E = \phi(r_p) - V_c^2 \ln(1 - \mathcal{R})$, with $\mathcal{R}$ a uniform random variable. With the energy (or $\mathcal{R}$) we can compute $\eta_{\max} = \sqrt{2\mathrm{e}}\sqrt{\gamma}(1-\mathcal{R})^{\gamma}\sqrt{-\ln(1-\mathcal{R})}$, for $\gamma = 1/2.5$ to mimic the stellar halo power law. Using $\eta_{\max}$ and the probability distribution for η we may derive the non-circularity of the orbits, and in this way fully determine the phase-space initial position of a satellite.

2.1.2 *Internal properties of the satellites*

For the satellites we assume they initially have King profiles, as do most of the satellites in the Local Group. Their present day total luminosity is fixed to be that of the stellar halo. The luminosity of each satellite is drawn from a Gaussian distribution with mean 2.5×10^7 L$_\odot$ and dispersion 10^7 L$_\odot$, thus reproducing the characteristic luminosities of Local Group dwarf spheroidals. The initial number of satellites is thus set to be 33. We assume that the satellites follow the scaling relations (Burstein et al. 1997)

$$\begin{aligned}\log L &= 5.35 + 1.80 \log \sigma_v, \\ \log R &= -0.82 + 0.51 \log \sigma_v.\end{aligned}$$

These relations allow us to derive from the luminosity L, the core radius R, and the central velocity dispersion σ_v. Our King models have a concentration parameter $c = \log r_t/r_K \sim 0.72$, where r_t and r_K are the tidal and King radii respectively. The initial mass of the satellite is now also fully determined.

2.2 Galactic potentials

We will consider two different Galactic potentials. In both cases, our Galaxy has three components: a dark halo, a disk and a bulge, but we take different functional forms for the potential. In Model I, we take a dark logarithmic halo

$$\Phi_{\mathrm{halo}} = v_{\mathrm{halo}}^2 \ln(r^2 + d^2), \tag{8}$$

a Miyamoto–Nagai disk

$$\Phi_{\mathrm{disk}} = -\frac{GM_{\mathrm{disk}}}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}, \tag{9}$$

and a spherical Hernquist bulge

$$\Phi_{\mathrm{bulge}} = -\frac{GM_{\mathrm{bulge}}}{r + c}, \tag{10}$$

where d=12 kpc and $v_{\mathrm{halo}} = 131.5$ km s^{-1}; $M_{\mathrm{disk}} = 10^{11}$ M$_\odot$, $a = 6.5$ kpc and $b = 0.26$ kpc; $M_{\mathrm{bulge}} = 3.4 \times 10^{10}$ M$_\odot$ and $c = 0.7$ kpc. This choice of parameters gives a flat rotation curve with an asymptotic circular velocity of 186 km s^{-1}.

In Model II, we represent the disk density profile with a double exponential (Quinn & Goodman 1986)

$$\rho_D(R, z) = \frac{M_D}{4\pi R_D^2 z_o} \mathrm{e}^{-R/R_D} \mathrm{e}^{-\beta|z|},$$

where $R_D = 3.5$ kpc is the disk scale length, z_o is its scale height, $\beta = 1/z_o$ and $M_D = 5.5 \times 10^{10}$ M$_\odot$ the total disk mass. The associated potential is

$$\Phi_D(R, z) = -\frac{GM_D}{R_D^3} \times$$

$$\int_0^\infty \frac{\mathrm{d}k k J_0(kR)}{k^2 + 1/R_D^2} \frac{\beta^2}{\beta^2 - k^2} \left\{ \frac{\mathrm{e}^{-k|z|}}{k} - \frac{\mathrm{e}^{-\beta|z|}}{\beta} \right\}. \tag{11}$$

For the halo we choose (Hernquist 1993)

$$\rho_h(r) = \frac{M_h}{2\pi^{3/2}} \frac{\alpha_q}{r_c} \frac{\mathrm{e}^{-r^2/r_c^2}}{r^2 + \gamma_q^2},$$

where $M_h = 1.5 \times 10^{12}$ M$_\odot$, $\alpha_q = \left(1 - \sqrt{\pi}\, q \mathrm{e}^{q^2}(1 - \mathrm{Erf}[q])\right)^{-1}$ with $q = \gamma_q/r_c$ and $r_c = 200$ kpc is the cutoff radius. The corresponding potential is

$$\Phi_h(r) = -\frac{GM_h(r)}{r} + \frac{GM_h}{\sqrt{\pi} r_c} \mathrm{Ei}\left[-\left(\frac{r}{r_c}\right)^2 - q^2\right], \tag{12}$$

where

$$M_h(r) = \frac{2M_h\alpha_q}{\sqrt{\pi}} \int_0^{r/r_c} \frac{x^2 \mathrm{e}^{-x^2}}{x^2 + q^2} \mathrm{d}x \tag{13}$$

and Ei(x) is the exponential integral (e.g. Gradshteyn & Ryzhik 1965). For the bulge we use Eq.(10) but we take $M_b = 1.1 \times 10^{10}$ and $c = 0.525$ kpc (following Velázquez & White 1995). Figure 1 shows the circular velocity curves produced by each of the two potentials.

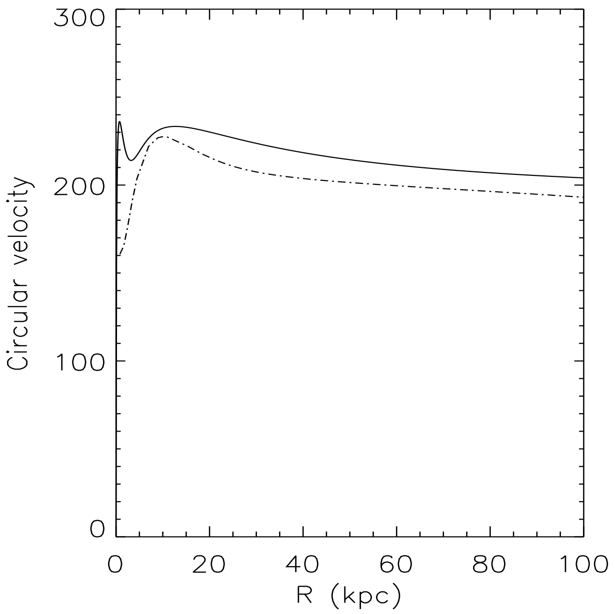

Figure 1. The circular velocity profile as a function of distance from the Galactic centre. The solid curve represents the potential used in the simulations. The dashed curve corresponds to the alternative potential.

2.3 Numerical methods

In our numerical simulations, we use the potential of Model I for our Galaxy. We represent the satellite galaxy by a collection of 10^5 particles and model their self-gravity by a multipole expansion of the internal potential to fourth order (White 1983; Zaritsky & White 1988). This type of code has the advantage that a large number of particles can be followed in a relatively small amount of computer time. In this quadrupole expansion, higher than monopole terms are

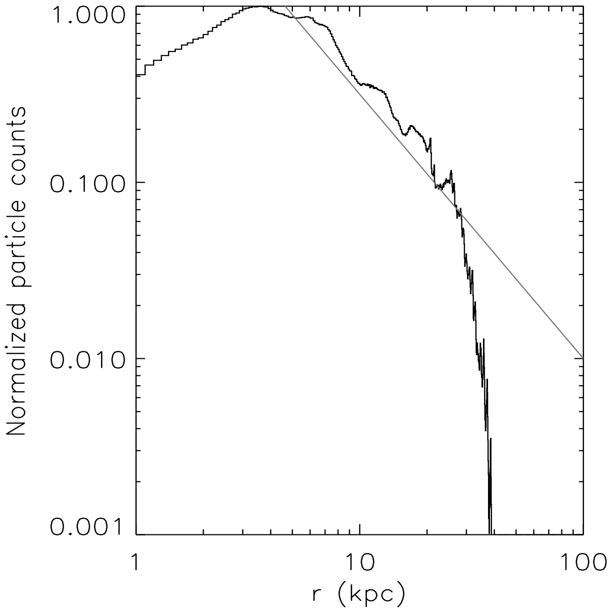

Figure 2. Number counts profile $N(r) = r^2\rho(r)$ for the simulated stellar halo resulting from the superposition of 33 disrupted satellite galaxies, after 12 Gyr of evolution. The straight line represents the expected $r^{-1.5}$ law, arbitrarily shifted.

softened more strongly. We choose $\epsilon_1 \sim 0.2 - 0.25R$ for the monopole term (R is the core radius of the system) and $\epsilon_2 = 2\epsilon_1$ for dipole and higher terms and for the centre of expansion. The centre of expansion is a particle which, in practice, follows the density maximum of the satellite closely at all times.

After letting our satellite relax in isolation, we integrate each simulation for $\sim$ 12 Gyr. In Figure 2 we show the final particle counts in radial bins $N(r) = r^2\rho(r)$ as a function of distance from the Galactic centre resulting from the superposition of all our experiments. For guidance, we also plot the expected $r^{-1.5}$, arbitrarily shifted. We see that within the range of 3 to 30 kpc, our simulations follow relatively well the profile. Outside this range we see a sharp drop, due to the fact that we are (intentionally) not populating the outer halo. Since the properties of the inner stellar halo are not so well constrained, we do not worry about the fact that we find a shallower slope in the inner few kiloparsecs (this is also the result of our initial conditions). More important is the fact that our simulations can reproduce very well the regime where the astrometric missions promise to give accurate six dimensional phase-space information.

2.4 Generating catalogues of halo stars

To generate an artificial catalogue for the Galaxy we assume that each particle in our simulations represents a giant star of absolute magnitude $M_V = 1$. The total number of particles in our simulations corresponds well to the expected number of giant stars in the Galactic halo (based on the luminosity function, derived for the age and metallicity characteristic of halo stars). We prefer to take only giant stars at this stage because they are bright enough to be easily observable from the Sun. We need to determine a limiting

Table 1. Estimated precision in parallax (σ_π, in μas) and proper motion (σ_μ, in μas yr^{-1}) as a function of V magnitude. For FAME and DIVA we assume $\sigma_\pi = \sigma_\mu$ (based on Horner (1999) and Röser (1998), respectively). In the case of GAIA, the estimated precisions correspond to a K3 III star with no reddening, and increase to 0.2 mas at $V \sim 20$ (Gilmore et al. 1998).

		9	10	11	12	13	14	15
GAIA	σ_π	3.65	3.65	3.65	3.65	4.83	7.05	10.8
	σ_μ	2.74	2.74	2.74	2.74	3.62	5.28	8.10
FAME	σ_π	24	36	56	90	146		
DIVA	σ_π	200	250	300				

magnitude $V_{\rm lim}$ for our "artificial" catalogue, which we define so that all giant stars brighter than $V_{\rm lim}$ have accurate full 6-dimensional phase-space information. In the case of GAIA, we take $V_{\rm lim} = 15$, a limit set by the accuracy in the radial velocity. For FAME, $V_{\rm lim} = 12.5$ as all stars brighter than this magnitude will have relative parallax errors σ_π/π smaller than (or of the order of) 25%. For DIVA, we take $V_{\rm lim} = 11$ for which $\sigma_\pi/\pi \sim 0.3$. Our GAIA, FAME and DIVA catalogues have 386144, 12497 and 1742 "stars" with $M_V = 1$ respectively.

The positions and velocities of each particle are first transformed into the observables (α, δ, π) and (μ_α, μ_δ, v_r); the expected observational "errors" are then added to the parallax, the radial velocity and the proper motion, according to Table 1. For GAIA the precision in the radial velocity is taken to be 5 km s^{-1} for $V < 14$, and to vary like $\sigma_v = 10(V - 14) + 10\,{\rm km\,s^{-1}}$ up to $V = 15$. Since FAME and DIVA will not measure radial velocities on board, for these we estimate the error $\sigma_v = 15\,{\rm km\,s^{-1}}$, as achievable from the ground for such large samples. These "observed" quantities are then transformed back to "observed" positions and velocities. We repeat this procedure 5 times to obtain 5 different realizations of the data.

3 FINDING DISRUPTED GALAXIES

3.1 Integrals of motion space

Our satellites disrupt relatively quickly, in only a few pericentric passages. Therefore we may consider each of the 33 satellites as an ensemble of particles with very similar integrals of motion (energy, angular momentum). As we show in Figure 3, initially satellites are both clumps in configuration and velocity space, as they are in (E, L, L_z) space. If these are conserved quantities, or evolve only slightly, this initial clumping should be present even after the system has phase-mixed completely. Thus the space of integrals or adiabatic invariants seems to be the natural space to look for the substructure produced by an accreted satellite.

There are a few issues we should address here before fully discussing a method based on clumping in the integrals of motion space. To compute the energy of the particles (or stars that will be observed by GAIA for example) we need to assume a Galactic potential. To determine the success of such a method we need to understand how our lack of

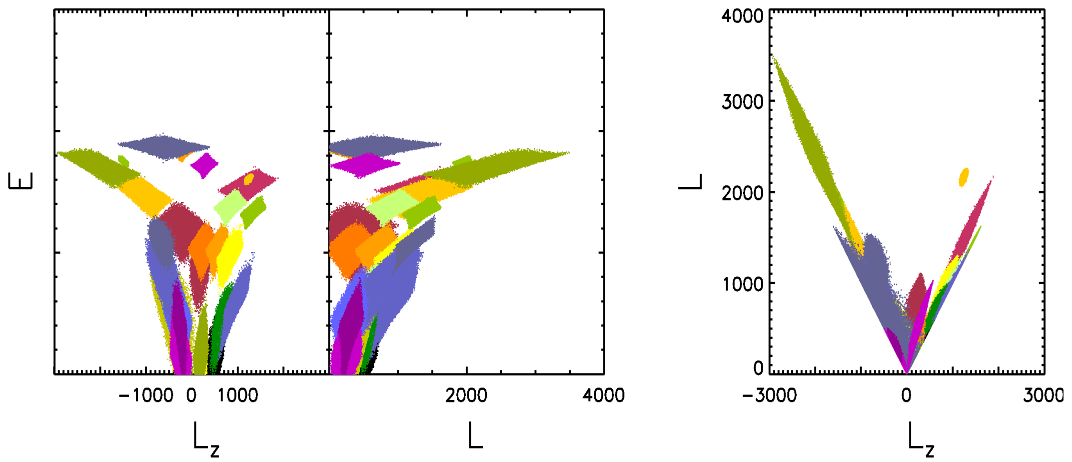

Figure 3. Initial distribution of particles in the integrals of motion space. The different colours represent different satellites.

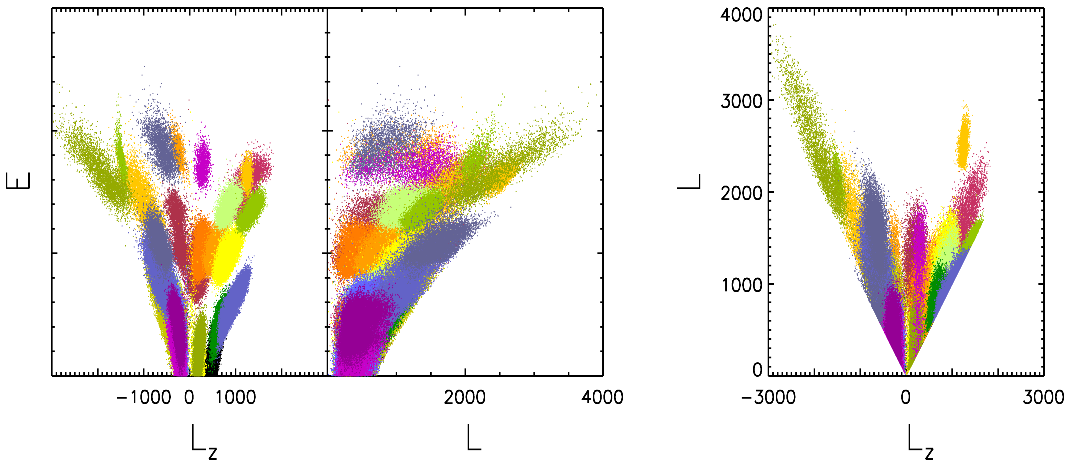

Figure 4. Final distribution of particles in the integrals of motion space after 12 Gyr, after convolution with the errors expected for GAIA for the original potential. Here we include all particles brighter than $V = 15$ (i.e. within roughly 6 kpc from the Sun).

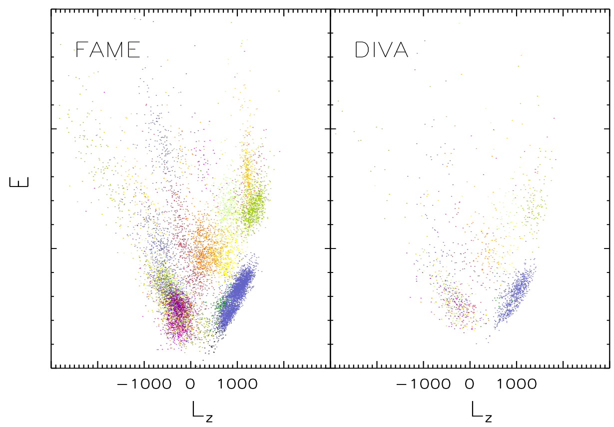

Figure 5. Final distribution of particles in the $L_z - E$ space after error convolution for FAME (left panel) and for DIVA (right panel), with energies computed using the original potential. A comparison to the left panel of Figs. 3 and 4 shows that the expected errors for these missions tend to erase much of the substructure left in the integrals of motion space.

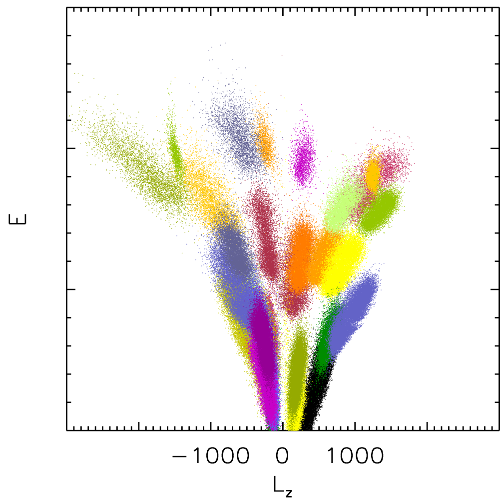

Figure 6. Final distribution of particles in the $L_z - E$ space after GAIA error convolution for the alternative potential. Compare to left panel of Fig.4.

knowledge on the precise form of the Galactic potential influences our results. We shall therefore proceed in two steps. In the first step, we take the same potential as that used in the simulations, Model I. In the final step we use the alternative potential introduced in Sec. 2.2, our Model II. This last step, in which we do not know the exact form of the Galactic potential but we make a reasonable guess, is most likely to represent the real situation.

Secondly, even though the total angular momentum is not fully conserved for an axisymmetric potential (only L_z is), it evolves preserving a certain degree of coherence. The advantage of using the integrals of motion space is that the number of clumps detected in this way will represent well the total number of accretion/merging events, since unlike other methods which are only local, it singles out all the stars from a given accreted object, independently of how different their phases and velocities might be. We choose to make use of all three integrals to reduce the chances of overlap amongst different lumps, since this probability clearly depends on the dimensionality of the space.

The analogue of Figure 3 for particles "brighter than 15th magnitude" (roughly within 6 kpc from the Sun) in the simulations, after 12 Gyr of evolution and for the original potential, shows that, even though there is some degree of evolution, clumping remains in the integrals of motion space. In Figure 4 we plot the integrals of motion space for one realization of the GAIA catalogue, i.e. after error convolution. A number of substructures are clearly visible, many of which can be directly related to the initial distribution, even with the GAIA observational uncertainties taken into account. This shows that the expected observational errors for GAIA will not affect the chances of detecting such substructures. In the case of FAME the situation is not as good, as illustrated in the left panel of Figure 5, where the different lumps are less populated (because of the magnitude limit) and considerably more smeared out (because of the larger observational errors). For DIVA the clumping has disappeared almost completely, as shown in the right panel of the same figure.

Figure 6 corresponds to the same realization of the GAIA catalogue as used before, but with the energies calculated using the case of the potential of Model II. Clearly, even though the two considered potentials are different, the substructure remains. The uncertainty in the precise form of the Galactic potential therefore does not affect the likelihood of finding disrupted satellites.

3.2 Method: FOF in integrals of motion space

We use a Friends-of-Friends (FOF) algorithm to find clumps in the integrals of motion space. This method has been used frequently to find bound halos in cosmological N-body simulations. The basic idea is that all particle pairs separated by less than a fraction ℓ of the mean interparticle distance are linked. Disjoint sets of connected particles are then identified as halos (Efstathiou et al. 1988). These halos correspond approximately to the regions interior to isodensity contours at an overdensity of $2/\ell^3$. This FOF procedure allows a rapid identification of halos, and moreover, all members of a given halo found for a particular value of ℓ are members of the same halo in any list generated for a larger value of ℓ. In the case of cosmological simulations, the linking distance is defined so that the mean density of a halo is about 200 times the density of the Universe at the time of identification.

In our case, it is less clear how we should define the interparticle distance, or the linking length. Because the energy and angular momentum have, by definition, different scales, it seems natural to try to reduce everything to the same scale, or equivalently, to use instead of spheres an ellipsoidal configuration. Even though the angular momentum and its z-component have the same scale, lumps are generally elongated in the L-direction with a 2:1 ratio, as can be seen from Figure 3. We therefore search for lumps whose characteristic size would be defined as:

$$\Delta L \sim 2\Delta L_z, \qquad \frac{\Delta E}{(\mathrm{km\,s^{-1}})^2} \sim 20\frac{\Delta L_z}{\mathrm{kpc\,km\,s^{-1}}},$$

where now ΔL_z would be related to the linking length. This implies that we re-scale the variables according to

$$E \rightarrow E/20, \qquad L \rightarrow L/2, \qquad L_z \rightarrow L_z.$$

The factor 20 in the energy scaling may be derived (heuristically) from the fact that the typical energy range in the Solar neighbourhood is $1.6 \times 10^5 (\mathrm{km\,s^{-1}})^2$, whereas the range of L_z is 8000 kpc $\mathrm{km\,s^{-1}}$ (from -4000 to 4000 kpc $\mathrm{km\,s^{-1}}$).

We will apply the FOF algorithm for two different linking lengths, to allow for different characteristic sizes of the halos and resolutions in the algorithm. Note that there are particular regions in this space which are occupied by more than one satellite, even in this 3-dimensional space (this is even worse if only the $L_z - L$ plane is used), so that not each of the lumps found may correspond to only one satellite, but may have contributions of a few.

3.3 Results

We apply the FOF algorithm to our GAIA catalogue, including error convolution for all particles brighter than 15th magnitude, and using the original potential. We take two different values for the FOF linking length: $\ell = 16$ and $\ell = 30$, where $\Delta L_z = 5\ell$. We consider groups with at least 500 "stars". We combine the two group catalogues to obtain a new group catalogue which contains all lumps detected. If some particles are found to belong to two different clumps (one from each catalogue) we keep the lump which has the smallest size. We now iterate one more time on the catalogue defined by the particles that do not belong to any of the lumps found by our FOF, again for the two values of ℓ and with a minimum of 250 "stars". Some of the newly found lumps can be related to those previously detected, and some others are found to resolve some of the largest lumps in our initial group catalogue. In the left panel of Figure 7 we show the distribution of energy E and L_z for our final group catalogue.

We find 17 different groups with this method. Not all the groups may be associated exclusively with one of our original satellites. As can be seen from Fig.3, there is quite a bit of superposition in this three–dimensional space, and so not all the original satellites can be recovered, or equivalently, not all lumps can be resolved with just two iterations. If we analyse how the particles in the different lumps can be related to particles in the initial satellites we find that, out of the 17 groups discovered, 14 can be associated

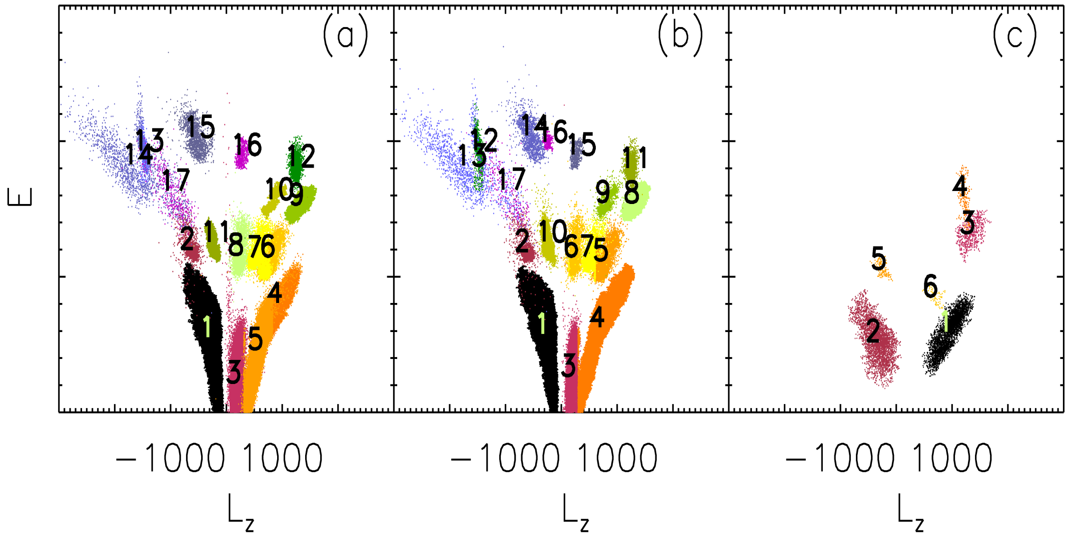

Figure 7. Lumps detected with our FOF algorithm. In panel (a) we show the final group catalogue for the original potential used in the simulations after convolving with the observational errors expected for GAIA. Panel (b) corresponds to our alternative potential and also to the GAIA catalogue. In both cases the recovery rate is about 50%. Panel (c) shows the lumps recovered by our FOF applied to the FAME catalogue generated as described in the text and for energies computed with the original potential. Compare to Fig.4 in the case of GAIA and to the left panel in Fig.5 for FAME.

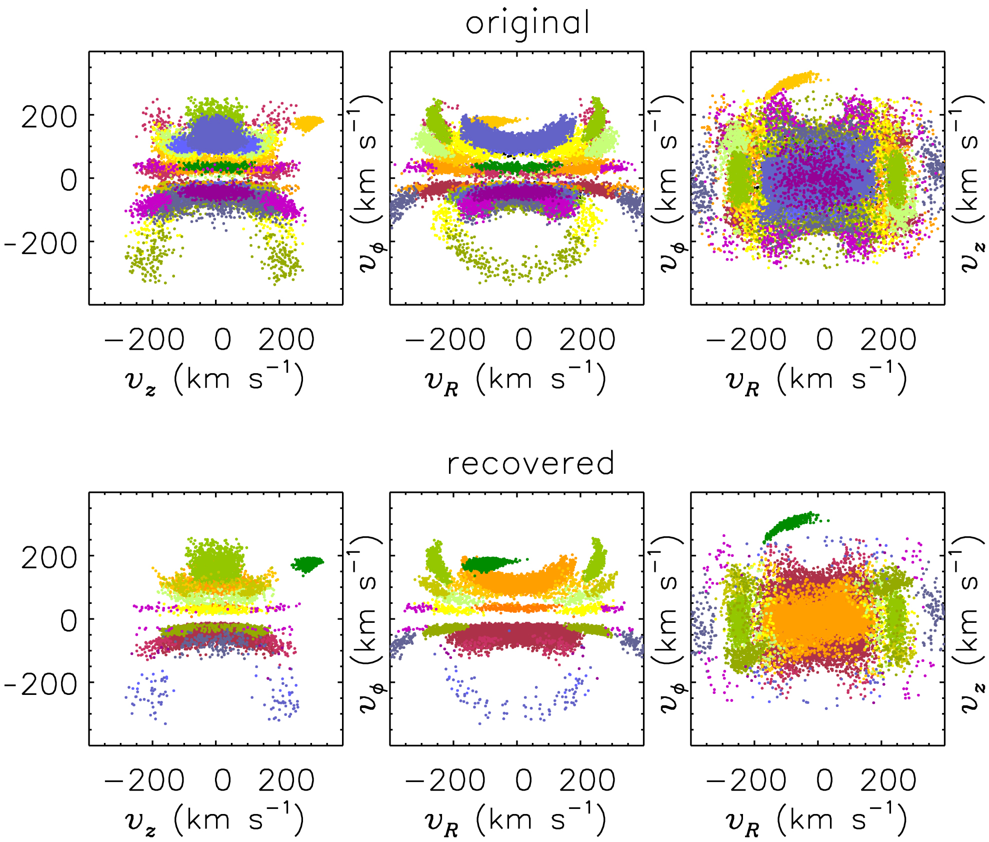

Figure 8. The velocity space distribution for particles in a cubic volume of 2 kpc on a side centered on the Sun, and for one realization of the GAIA catalogue. In the upper panels different colours indicate particles associated with different satellites (using the same colour coding as in Figure 3). In the lower panels, the colours are used to show particles associated to the lumps recovered by our FOF algorithm applied to the GAIA catalogue in the case of the alternative Galactic potential. (Here the colour coding corresponds to that used in panel (b) of Figure 7.)

almost uniquely to one satellite*. This means that our simple method is capable of finding more than 40% of all the satellites that were accreted by our "Galaxy".

Similarly, we apply the FOF algorithm to the same GAIA catalogue but now compute the energy E of the particles with the alternative potential. In this case we again find 17 different lumps (after two iterations, and combining the results of the two different values of the linking length). Of these 17 groups, 14 can be uniquely associated to one satellite. This is shown in the central panel of Figure 7. Our method is thus quite successful in identifying disrupted satellite galaxies in integrals of motion space, even when we only have a guess for the Galactic potential and when the observational uncertainties are taken into account. Further uncertainties such as the distance from the Sun to the Galactic centre and the velocity of the Local Standard of Rest are also negligible.

When we apply the same method for the original potential on the FAME catalogue we are able to find 6 different groups. For 5 of these a unique correspondence with an accreted satellite exists, as shown in the rightmost panel of Figure 7. In the case of DIVA we find only 1 group (with at least 20 particles), which can be easily identified visually from the right panel in Figure 5. In the cases of DIVA and FAME we used slightly larger linking lengths to take into account the smearing out of the lumps caused by the larger observational uncertainties. Because the samples are also smaller (because of the limiting magnitude), we consider groups with at least 50 particles in the case of FAME, and 20 particles for DIVA.

4 CLUMPINESS IN THE KINEMATICS OF HALO STARS

In Figure 8 we show the velocities of all the particles contained in a volume of 2 kpc on a side for one realization of the GAIA catalogue. There is considerable substructure, which is visible thanks to the great precision that GAIA will achieve. From the upper panels it is clear that, as discussed in the introduction, distinguishing the satellites that gave rise to each one of the different moving groups is a non-trivial task in this space. In the lower panels we have coloured the different contributions from the 14 groups detected by our FOF algorithm in the case of the alternative potential. A comparison between upper and lower panels also shows how successful our method is.

The kinematically cold streams visible in Figure 8 remain as coherent structures for longer than a Hubble time. This is true even when mergers, rather than simple satellite accretion, are dominant (Helmi, White & Springel 2000). The clumpiness in the kinematics of halo stars should thus be a distinct feature of the hierarchical formation of our Galaxy. It is therefore also interesting to determine the degree of the clumpiness and whether it will be measurable with future astrometric missions. We will determine this clumpiness using the two–point correlation function ξ in velocity space for a sphere of 1 kpc radius around the Sun. We estimate ξ from

$$\xi = \frac{\langle DD\rangle\langle RR\rangle}{\langle DR\rangle^2} - 1 \tag{14}$$

(e.g. Hamilton 1993) where $\langle DD\rangle$ is the normalised number of pairs of particles with velocities in a given velocity range (or bin), i.e.

$$\langle DD\rangle = \frac{\sum \text{pairs of particles } i,j \text{ with } v < |\mathbf{v}_i - \mathbf{v}_j| < v + \Delta}{N_D(N_D-1)} \tag{15}$$

and where N_D is the number of particles in the sphere. $\langle RR\rangle$ is defined analogously but for N_R random points. The random variates are drawn from a trivariate Gaussian distribution determined from the "data" in the principal axes velocity frame. Here we take $N_R = 10N_D$. Finally $\langle DR\rangle$ are the normalised counts for "data"–random pairs. We estimate the uncertainty from

$$\Delta_\xi = (1+\xi)\sqrt{\frac{2}{N_D(N_D-1)\langle DD\rangle}}. \tag{16}$$

If the sample contains kinematically cold streams, we should find an excess of pairs in the bins corresponding to small velocity differences, i.e. the correlation function should be significantly different from zero (which corresponds to the absence of correlations). We proceed by measuring ξ for our GAIA, FAME and DIVA catalogues including error convolution as described in Sec. 2.4. We also vary the position of the 1 kpc sphere around the "Sun", keeping the same distance from the "Galactic centre". That is, we place the Sun at $(x,y) = \{(8,0),(0,-8),(-8,0),(0,8)\text{kpc}\}$ and $z = 0$. This allows us to account for the natural variations one may have from volume to volume. We then make five realizations for each "Sun" position for each catalogue. In Figure 9 we show the correlation function obtained by averaging over all the realizations, for each 1 kpc sphere and for each catalogue. The average ξ for each volume is the weighted mean, where the weights are given by $1/\Delta_\xi^2$, and the error bars indicate the (weighted) dispersion around the (weighted) mean. We find an excess of pairs of stars with similar motions, the signature indicating the presence of cold streams as expected for a stellar halo built by disrupted satellites. Note that it will even be possible to determine that the halo is not a smooth distribution in the Solar neighbourhood even with velocity errors of the order of 20 $\mathrm{km\,s^{-1}}$ such as those expected for DIVA for a star with $M_V = 1$ at 500 pc from the Sun.

5 DISCUSSION

We simulated the entire stellar halo of the Galaxy starting from disrupted satellite galaxies, and "observed" it with the next generation of astrometric satellites (DIVA, FAME and GAIA). We analysed the observations with the aim of recovering the different accretion events our "Galaxy" experienced over its lifetime. We used a FOF algorithm to find clumps in the integrals of motion space, which we expected would correspond to the disrupted satellites. Our integrals of motion space is defined by energy E, total angular momentum L and its z-component L_z, even though strictly

* We say that a group is almost uniquely associated to one satellite if more than 70% of the particles in the group belong to only that satellite.

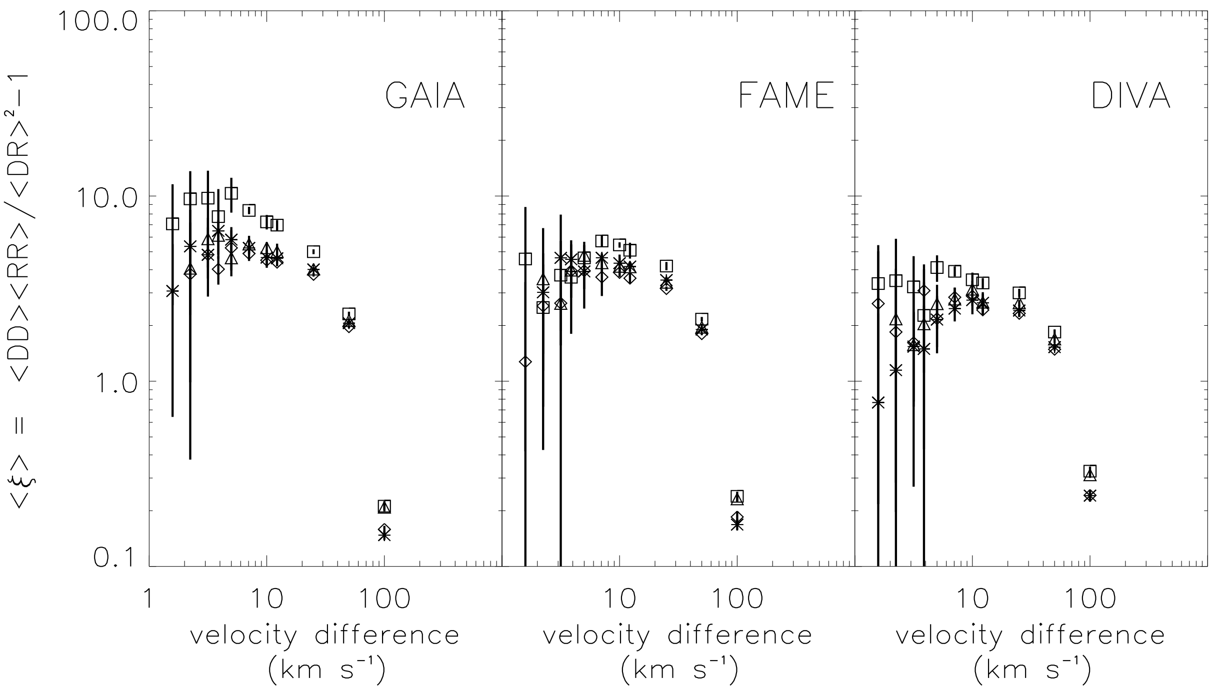

Figure 9. The two-point correlation function for "giant stars" inside spheres of 1 kpc radius around the Sun (defined as 8 kpc from the Galactic centre on the Galactic disk) computed as the weighted average over five realizations of the DIVA, FAME and GAIA catalogues. The different symbols correspond to ξ measured inside spheres at different locations of the "Sun" on the Solar circle.

speaking these are not fully conserved quantities (because of interaction of the stars while still bound to the satellite, and because of the axisymmetry of our Galaxy). We have shown that the initial clumping in this space is maintained to a great extent even after 12 Gyr of evolution.

After using our FOF algorithm we find that we can only recover a couple of accreted satellites (in our analysis just one) for the DIVA catalogue, whereas for FAME we recover about 15% of all satellites. In both cases we assume that the astrometry is complemented by ground based radial velocity measurements. The situation is significantly different in the case of the GAIA catalogue, for which we recover almost half of all disrupted satellites with this simple algorithm. The improvement generally lies in the larger volume for which full 6-D information is available, in particular when comparing FAME and GAIA. The use of 6-D information appears to be essential to recover all the events, as there is a large fraction of phase-space where these are superposed. This is particularly clear from Figure 4 (rightmost panel), where angular momentum alone cannot be used to distinguish the different satellites. Whereas by eye inspection in the (E, L, L_z) space we may recover five or six events, for the space (L, L_z) this is reduced to one or two events. Our results are unlikely to be strongly dependent on the particular choice of the luminosity distribution of the disrupted satellites. This is because a large number of small satellites occupy basically the same phase-space volume as a small number of large ones.

The evolution of the Galactic potential may be the most crucial simplification in our analysis. In hierarchical cosmologies the number of objects that form a galaxy like our own is in the range of $5-20$, with comparable masses. The process of formation is likely to be very violent and the potential is surely not static, quite probably not axisymmetric, and therefore the initial clumping of the system may not be reflected in clumping in our defined integrals of motion space. However, if this happened during the first few Gyrs, any object infalling later, ought to have perceived a fairly static (or adiabatically changing) Galaxy, and then our method would still be useful. Indeed, some preliminary analysis of the formation of a halo in a ΛCDM cosmology indicates that particles from different satellites may be recovered as lumps in this space (Helmi et al. 2000), though the structure is less evident than in the plots shown here, where even by simple eye inspection one may recover about 1/5 of all satellites.

What will anyway remain as signatures of the merger history of our Galaxy will be the kinematically cold streams originating in disrupted halos. An interesting observational test is the comparison of the kinematics of a smooth, possibly Gaussian, distribution (which may be expected in the case of a monolithic collapse) to the kinematics observed in the stellar halo built by disrupted satellites. Our analysis of the correlation function in velocity space indicates the presence of a larger number of streams with very small velocity dispersions in a sphere of 1 kpc radius around the Sun. This test will be feasible even for DIVA. The key to the success of this test lies in the complete and large sample of stars with 3-D velocities which will be available.

In this paper we have focused on determining the merger history of the Milky Way, rather than the precise form of the Galactic potential or to what extent it may have varied. However, these are key questions that will be solved very likely by SIM and GAIA (e.g. Johnston et al. 1999). We may add here that after finding the different satellites we will be able to determine the conditions and characteristics of objects that fell onto the Milky Way more than 10 Gyr ago.

ACKNOWLEDGMENTS

We wish to thank Volker Springel for his FOF algorithm, Simon White for many useful discussions, and Anthony Brown for comments on an earlier version of this manuscript.

REFERENCES

Burstein D., Bender R., Faber S.M., Nolthenius R., 1997 , AJ, 114, 1365
Chereul E., Crézé M., Bienaymé O., 1999, A&AS, 135, 5
de Bruijne J., 1999, 306, 381
de Zeeuw P.T., Hoogerwerf R., de Bruijne J. H. J., Brown A. G. A., Blaauw A., 1999, AJ, 117, 354
Efstathiou G., Frenk C.S., White S.D.M., Davis M., 1988, MNRAS, 235, 715
Fuchs B., Jahreiß H., 1998, A&A, 329, 81
Gerhard O., 1991, MNRAS, 250, 812
Gilmore G. F., Perryman M. A., Lindegren L., Favata F., Hoeg E., Lattanzi M., Luri X., Mignard F., Röser S., de Zeeuw P. T., 1998, SPIE, 3350, 541
Gradshteyn I.S., Ryzhik I.M., 1965, Table of integrals, series and products, Academic Press
Hamilton A.J., 1993, ApJ, 417, 19
Helmi A., White S.D.M., 1999, MNRAS, 307, 495
Helmi A., White S.D.M., de Zeeuw P.T., Zhao H.S., 1999, Nature, 402, 53
Helmi A., Springel V., White S.D.M., 2000, in preparation
Hernquist L., 1993, ApJS, 86, 389
Hoogerwerf R., Aguilar L., 1999, MNRAS, 306, 394
Horner S.D., Germain M.E. et al., 1999, in "Working on the Fringe: An International Conference on Optical and IR Interferometry from Ground and Space", ASP Conf. Series (in press)
Ibata R., Gilmore G., Irwin G., 1994, Nature, 370, 194
vezic Z., Goldston J., Finlator K., et al. (for the SDSS), 2000, submitted to AJ (astro-ph/0004130)
Johnston K.V., Hernquist L., Bolte M., 1996, ApJ, 465, 278
Johnston K.V., Zhao H.S., Spergel D.N., Hernquist L., 1999, ApJ, 512, L109
Kinman T.D., Suntzeff N.B., Kraft R.P., 1994, AJ, 108, 1722
Lynden-Bell D., Lynden-Bell R.M., 1995, MNRAS, 275, 429
Lynden-Bell D., in "The Galactic Halo", Proc. of the 3rd. Stromlo Symposium, ASP Conf. Series 165, p.17
Majewski S.R., Munn J.A., Hawley S.L., 1996, ApJ, 459, L73
Mateo M., 1998, ARA&A, 36, 435
Quinn P.J., Goodman J., 1986, ApJ, 309, 475
Röser S., 1998, in Jahrestagung der Astronomischen Gesellschaft, Heidelberg: "DIVA - Beyond HIPPARCOS and Towards GAIA" (http://www.aip.de/groups/DIVA/)
van den Bosch F.C., Lewis G.F., Lake G., Stadel J., 1998,
van der Marel R.P., Sigurdsson S., Hernquist L., 1997, ApJ, 487, 153
Velázquez H., White S.D.M., 1995, MNRAS, 275, 23L
White S.D.M., 1983, ApJ, 274, 53
hite S.D.M., Rees M.J., 1978, MNRAS, 183, 341
Zaritsky D., White S.D.M., 1988, MNRAS, 235, 289