Mon. Not. R. Astron. Soc. **000**, 1–12 (2010) Printed 10 November 2010 (MN LATEX style file v2.2)

# A Two-Point Correlation Function For Galactic Halo Stars

A.P. Cooper[1,2⋆], S. Cole[1], C.S. Frenk[1], A. Helmi[3]
[1] *Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham, DH1 3LE, UK*
[2] *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748, Garching, Germany*
[3] *Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, Netherlands*



Accepted xxxx. Received xxxx; in original form xxxx

**ABSTRACT**

We describe a correlation function statistic that quantifies the amount of spatial and kinematic substructure in the stellar halo. We test this statistic using model stellar halo realizations constructed from the Aquarius suite of six high-resolution N-body simulations in combination with the GALFORM semi-analytic galaxy formation model. These simulations show considerable scatter in the properties of stellar haloes. We find that our statistic can distinguish between these plausible alternatives for the global structure of the Milky Way stellar halo. We compare with observational data and show that pencil beam surveys of $\sim 100$ tracer stars (such as the Spaghetti Survey) are not sufficient to constrain the degree of structure in the Milky Way halo with this statistic. Larger area surveys with $\geqslant 1000$ tracer stars (such as BHB stars in the Sloan Digital Sky Survey) provide much tighter constraints on comparisons between models and data. In our simulations, we find examples of haloes with spatial and kinematic substructure consistent with the available Milky Way data.

**Key words:** galaxies – Milky Way: formation.

## 1 INTRODUCTION

In the Cold Dark Matter (CDM) cosmogony, galactic stellar haloes are built up in large part from the debris of tidally disrupted satellites (e.g. Searle & Zinn 1978; White & Springel 2000; Bullock & Johnston 2005; Cooper et al. 2010). Discovering and quantifying halo structures around the Milky Way may provide a useful diagnostic of the Galaxy's merger history (e.g. Helmi & de Zeeuw 2000; Johnston et al. 2008; Gómez & Helmi 2010). Upcoming Milky Way surveys (for example with PanSTARRS1, LAMOST, HERMES and the LSST) will provide large datasets in which to search for structure, and the *Gaia* mission will determine six-dimensional phase-space coordinates for all stars brighter than $V \sim 17$, from which it should be possible to untangle even well-mixed streams in the nearby halo (Gómez et al. 2010).

Testing the CDM model by comparing these observations with numerical simulations of stellar halo formation requires that the 'abundance of substructure' be quantified in a straightforward way, with a method equally applicable to simulations and observations. Algorithms already exist for identifying substructure in huge multi-dimensional datasets (e.g. Sharma & Johnston 2009), such as those expected from *Gaia*, supplemented by chemical abundance measurements (Freeman & Bland-Hawthorn 2002). These algorithms can also be applied to simulations, although this is not straightforward. One problem is that current (cosmological) hydrodynamic simulations still fall short of the star-by-star 'resolution' of the *Gaia* data, particularly in the Solar neighbourhood (e.g. Brown, Velázquez & Aguilar 2005).

In the outer halo, longer mixing times allow ancient structures to remain coherent in configuration space for many gigayears. However, 6D *Gaia* data will be restricted to relatively bright stars. In the near future studies of the outer halo (beyond $\sim 20$ kpc) will continue to rely on a more modest number of 'tracers' (giant and horizontal branch stars). For these stars, typically only angular positions and (more uncertain) estimates of distance and radial velocity are available. Current simulations contain as many particles as there are (rare) tracer stars in observational samples. This enables the comparison that we present here between models and data that are already available. We focus on quantifying the degree of structure in rare tracer stars in a generic way, which we apply to these data and to simulations of stellar haloes.

Most studies of spatial and kinematic structure in the Milky Way halo have given priority to the discovery of individual overdensities (exceptions include Bell et al. 2008, Xue, Rix & Zhao 2009 and Helmi et al. in preparation). Relatively few have investigated global statistical quantities for the entire stellar halo, although several authors have suggested an approach based on clustering statistics. Re Fiorentin et al. (2005) analysed the velocity-space clustering of a small number of halo stars in the Solar neighbourhood, using a correlation function statistic. Following early work by Doinidis & Beers (1989), Brown et al. (2004) examined the angular two-point correlation function of photometrically selected blue

⋆ E-mail: acooper@mpa-garching.mpg.de

horizontal branch (BHB) stars in the 2MASS[1] catalogue, probing from $\sim 2-9$ kpc. They detected no significant correlations at latitudes $|b| \gtrsim 50^\circ$, but did detect correlations on small scales ($1^\circ$, $\sim 100$ pc) at lower latitudes, which they attributed to structure in the thick disc. Lemon et al. (2004) performed a similar analysis for nearby F stars in the Millennium Galaxy Catalogue and found no significant clustering.

Starkenburg et al. (2009) used a correlation function in *four* dimensions (which we discuss in detail below) to identify substructures in the Spaghetti pencil-beam survey of the distant halo, from which they obtained a significant detection of clustering and set a lower limit on the number of halo stars in all substructures. Similarly, Schlaufman et al. (2009) constrained the mass fraction of the halo in detectable substructure by estimating the completeness of their overdensity detection algorithm. Starkenburg et al. and Schlaufman et al. concluded respectively that $> 10\%$ (by number of stars) and $\sim 30\%$ (by volume) of the Milky Way halo belongs to groups meeting their own definitions of phase space substructure. These methods were tested on 'mock catalogues' of tracer stars constructed from simplified models of the stellar halo. Also of particular relevance to this work is the study of Xue et al. (2009), who considered the pairwise radial velocity separation of a large sample of halo BHB stars as a function of their separation in space, but found no evidence of clustering.

The statistic we develop below is more general than the otherwise similar approach of Starkenburg et al. (2009) and more sensitive than that of Xue et al. (2009). Following Starkenburg et al. (2009), we define a two-point correlation function using a metric that combines pairwise separations in the four most readily obtained phase space observables for halo stars (angular position, radial distance and radial velocity). We apply this statistic to the data analysed by Xue et al. (2009) and demonstrate that a significant signal can be extracted.

A metric of the kind we propose can be tuned to probe a specific scale of clustering by adjusting the weight given to each of its components (e.g. Starkenburg et al. 2009). However, it is not clear what signal is to be expected from a 'typical' $\Lambda$CDM stellar halo which is a superposition of many sub-components with a complex assortment of morphologies in phase space. We can identify no clearly 'optimal' metric. Instead, we make a fiducial choice of scaling which we test using the self-consistent accreted halo models of Cooper et al. (2010). These incorporate a full $\Lambda$CDM galaxy formation model and are based on high-resolution cosmological simulations from the Aquarius project (Springel et al. 2008). Having defined our metric, we are able to make direct comparisons between these simulations and the data of Xue et al. (2008). We show that even though both the metric and our choice of scaling are simple, this approach has the power to discriminate quantitatively between qualitatively different stellar haloes.

We describe the basis of our method in Section 2 and the observational data to which we compare in Section 3. In Section 4 we describe our simulations and our procedure for constructing mock catalogues. In Section 4.3 we discuss how our technique relates to the similar approach of Starkenburg et al. (2009) in the context of the Spaghetti Survey (Morrison et al. 2000). We show that the number of stars in this survey is too small to give useful constraints with our approach. In Section 6 we apply our method to a much larger sample of BHB stars (Xue et al. 2008) from the 6th data release (DR6) of the Sloan Digital Sky Survey (SDSS SEGUE), and compare these data to our simulations. Our conclusions are given in Section 7.

[1] Two Micron All Sky Survey (Skrutskie et al. 2006)

## 2 METRICS FOR PHASE-SPACE DISTANCE

The most readily obtained phase-space observables for halo stars are their Galactic angular coordinates, $l$ and $b$, heliocentric radial distance, $d$, and heliocentric line-of-sight velocity, $v_{\rm hel}$. From their angular position and distance estimate, each star can be assigned a three-dimensional position vector in galactocentric Cartesian coordinates, $\boldsymbol{r}$ $(X, Y, Z)$, and a radial velocity corrected for the Solar and local standard of rest motions, $v_{\rm GSR}$ (hereafter $v$). We begin by defining a scaling relation (metric), $\Delta$, which combines these observables into a simple 'phase-space separation' between two stars:

$$\Delta_{ij}^2 = |\boldsymbol{r_i} - \boldsymbol{r_j}|^2 + w_v^2 (v_i - v_j)^2. \quad (1)$$

Here, $|\boldsymbol{r_i} - \boldsymbol{r_j}|$ is the separation of a pair of stars in coordinate space (in kiloparsecs), and $v_i - v_j$ is the difference in their radial velocities (in kilometres per second). The scaling factor $w_v$ has units of $\mathrm{kpc\,km^{-1}\,s}$, such that $\Delta$ has units of kpc. The choice of $w_v$ is arbitrary unless a particular 'phase space scale' of interest can be identified. This is not straightforward; we discuss some possible choices below.

The aim of this paper is to explore $\xi(\Delta)$, the cumulative two-point correlation function of halo stars in the metric defined by Equation 1. Throughout, we use the estimator

$$1 + \xi(\Delta) = \frac{DD(<\Delta)}{\langle RR(<\Delta)\rangle}. \quad (2)$$

Here $DD(<\Delta)$ counts the number of pairs in the sample separated by less than $\Delta$, and $\langle RR(<\Delta)\rangle$ is the equivalent count of pairs of points randomly placed in the volume of the survey, averaged over a number of realisations.

The definition of $\Delta$ given by Eqn. 1 is similar but not identical to the *4-distance* metric of Starkenburg et al. (2009). These authors define the phase-space separation of two stars, $\delta_{\rm 4d}$, to be

$$\delta_{4d,ij}^2 = \tilde{w}_\phi \phi_{ij}^2 + \tilde{w}_d (d_i - d_j)^2 + \tilde{w}_v (v_i - v_j)^2. \quad (3)$$

The observables in Eqn. 3 are the same as those in Eqn. 1. In this definition, the distance in configuration space between two stars is split into an angular component, $\phi$, and a radial component, $d$. The constant scaling factors, $\tilde{w}$, normalize each component to its maximum observable value in the Spaghetti survey (described below). Starkenburg et al. choose $\tilde{w}_\phi = 1/\pi^2$, $\tilde{w}_d = \eta_{d,ij}/(130\,\mathrm{kpc})^2$, and $\tilde{w}_v = \eta_{v,ij}/(500\,\mathrm{km\,s^{-1}})^2$. The $\eta_{ij}$ terms are intended to incorporate into the metric itself the observational errors, $\sigma_{d,v}$, on $d$ and $v$. They are defined relative to the typical error of a star in the survey:

$$\eta_{d,ij} = \frac{(\sigma_{d,i}/d_i)^2 + (\sigma_{d,j}/d_j)^2}{2\langle \sigma_d/d\rangle^2} \quad (4)$$

$$\eta_{v,ij} = \frac{\sigma_{v,i}^2 + \sigma_{v,j}^2}{2\langle \sigma_v\rangle^2}\,. \quad (5)$$

If these $\eta$ terms are neglected (i.e. $\eta_{d,v} = 1$), the metric of Starkenburg et al. can be related to Eqn. 1 by separating radial distance into transverse and parallel components, i.e. $r^2 = r_\parallel^2 + r_\perp^2$. The $r_\parallel$ term is exactly equivalent to $d$ in Eqn. 3, but transforming the angular separation of the stars to a transverse distance, $r_\perp$, is

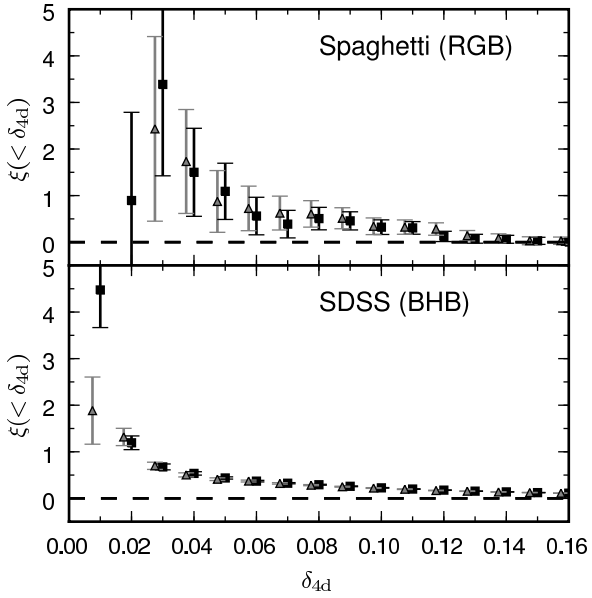


**Figure 1.** Upper panel: the *4-distance* cumulative correlation function $\xi(<\delta_{4\rm d})$ defined by Starkenburg et al. (black squares with Poisson error bars) for 101 RGB stars from the Spaghetti survey (Starkenburg et al. 2009). Grey triangles (slightly offset in $\delta_{4\rm d}$ for clarity) show the same function setting $\eta = 1$ in the metric of Starkenberg et al. (see text). Lower panel: correlation functions in the same metric for 2558 BHB stars ($6 < r < 60\,{\rm kpc}$) from the SDSS DR6 sample of Xue et al. (2008).

less straightforward[2]. At small angles, where $r_\perp^2 = r^2 - r_\parallel^2 \approx w_\phi^2 \phi_{ij}^2$, the scaling between our metric and that of Starkenburg et al. corresponds to $\Delta = (130\,{\rm kpc}) \times \delta_{4\rm d}$, with $w_\phi = 130/\pi \sim 41.4\,{\rm kpc\,rad^{-1}}$ and $w_v = 0.26\ {\rm kpc\,km^{-1}\,s}$.

Starkenburg et al. suggest that pairs of stars separated by a suitably small $\delta_{4\rm d}$ can be regarded as a 'group'. To determine the optimum value of $\delta_{4\rm d}$ to define groups in a given survey[3], they examine the cumulative two-point correlation function with a $DD/\langle RR\rangle$ estimator equivalent to Eqn. 2. Starkenburg et al. compute $\langle RR\rangle$ by repeatedly 'reshuffling' the distances and velocities of stars in the sample amongst their (fixed) angular coordinates $(l, b)$. As described below, we adopt the same procedure when computing our correlation function, $\xi(\Delta)$.

Fig. 1 shows two correlation functions in the Starkenburg et al. *4-distance* metric. The first of these is for giant branch stars from the Spaghetti survey (as in Fig. 1 of Starkenburg et al. 2009) and the second is for a much larger sample of BHB stars in SDSS DR6 extending[4] to $\sim 60$ kpc (Xue et al. 2008). These datasets are described in more detail in the following sections.

We also show in Fig. 1 the equivalent correlation functions setting $\eta_d = 1$ and $\eta_v = 1$ in Eqn. 3 as described above. These terms make interpretation of the metric distance more complicated. When they are included, pairs with larger errors have larger $\delta_{4\rm d}$, and so are assigned to a higher-separation bin in the cumulative correlation function. Including these terms has a practical advantage if $\delta_{4\rm d}$ is used only as a 'structure finder', as they 'clear out' dubious pairs from the smallest separation bins. However, $\delta_{4\rm d}$ then depends not only on the physical phase-space coordinates of two stars, but on how well those coordinates are measured. For example, increasing the radial separation of a pair by 10 kpc and improving the measurement of distance for both stars by a factor of 10 (relative to the average of the sample) would result in the same $\delta_{4\rm d}$. Thus $DD/RR$ is not a straightforward measurement of physical clustering when these weighting factors are used. Fig. 1 shows that setting $\eta = 1$ affects the significance of the correlation function signal for these two surveys at the smallest separations but makes little difference to the overall shape. This is especially true in the case of the SDSS BHB stars.

Both samples show a significant signal in $\xi(<\delta_{4\rm d})$ at small separations. Starkenburg et al. adopted $\delta_{4\rm d} = 0.05$ as a suitable 'linking length' to identify meaningful groups in the Spaghetti survey. Here we are not concerned with the identification of individual groups. Instead, our aim is to use our $\xi(<\Delta)$ correlation function to quantify in a statistical sense the *overall* nature of structure in the halo.

## 3 OBSERVATIONAL SAMPLES FROM SPAGHETTI AND SDSS

Spaghetti is a survey of the stellar halo in 134 pencil-beam fields covering $\sim 52\,{\rm sq.deg.}$ at high Galactic latitudes (Morrison et al. 2000; Dohm-Palmer et al. 2000). Of these 134 fields, 101 were targeted randomly within the region defined in Galactic coordinates by $b > 30^\circ$ and $0 < l < 210^\circ$; the remaining 33 were targeted randomly in the region $b < -30^\circ$. Metal-poor red giant branch (RGB) halo star candidates were identified photometrically using a combination of Washington System filters (Morrison et al. 2001). In a subset of 52 fields (amounting to $\sim 25\,{\rm sq.deg.}$) all candidates were followed up spectroscopically to distinguish true giants from nearby metal-poor dwarfs. Radial velocities and metallicities were determined from the spectra of 101 confirmed halo RGB stars with errors of 10-15 ${\rm km\,s^{-1}}$ and 0.25–0.3 dex respectively. Distances for these stars were determined from spectroscopic luminosity estimates as described by Morrison et al. (2000, 2003). Errors associated with the spectroscopic metallicity measurements (used to select fiducial globular cluster colour-magnitude tracks) are the most significant contribution to a typical distance error of $\sim 15$ per cent.

Xue et al. (2008) have published a catalogue of 2558 stars from SDSS DR6 which they identify as halo BHBs with high confidence (contamination $< 10\%$), using a combination of colour cuts and Balmer line diagnostics. This sample ranges in distance from

[2] Separating the components of distance in the metric in this way is natural where radial distance error dominates the uncertainty. One simple way to proceed may be to include an '$r\,d\theta$' term in the angular separation, where $r$ is defined (for example) as the mean distance of the two stars. In a modified Eqn. 3 this could be achieved by setting $w_{\phi,ij} \propto \sqrt{d_i d_j}$. With this definition, larger values of $\Delta$ result for pairs of the same angular separation lying at larger mean absolute distances (note that this is not the case in the Starkenburg et al. formulation of Eqn. 3). However, this treatment is not sensible at large angular separations.

[3] That is, a value large enough to confidently link together many stars, without creating spurious groups.

[4] In the case of the Xue et al. BHB stars we have set $\tilde{w}_d = 130$ kpc in Eqn. 3. Starkenburg et al. define $\tilde{w}_d$ as the maximum distance probed by the survey – the Xue et al. high-confidence sample extends to $w_d = 60$ kpc, and the most distant BHB found in the full SDSS sample lies at $\sim 90$ kpc. However, adjusting $\tilde{w}_d$ to these 'limits' makes a negligible difference to the correlation function.

$4 - 60\,\mathrm{kpc}$; a cut on distance error excluded more distant stars observed in SDSS. The errors on distance ($\sim 5\%$) and radial velocity ($5 - 20\,\mathrm{km\,s^{-1}}$) for stars in the Xue et al. catalogue are comparable to or better than those of Spaghetti. The Xue et al. BHBs are not a complete sample in any sense. In particular, as described by Xue et al., the targeting of SDSS spectroscopy disfavours follow-up of more distant BHBs. Completeness estimates are difficult for both surveys, particularly for SDSS, because of the complex procedures by which candidates are identified and confirmed.

In the study of the Galactic escape velocity profile for which the Xue et al. sample was obtained, the authors further restricted the data to 2401 stars by selecting only stars with a height $|z| > 4\,\mathrm{kpc}$ above the Galactic plane. This cut was designed to exclude thick disc BHB stars. In our analysis of the data we retain the full high-confidence sample of 2558 BHBs and do not impose any restriction on $|z|$ in our mock observations, beyond that of the SDSS footprint.

Xue et al. (2009) studied the pairwise radial velocity distribution of the Xue et al. (2008) BHB sample as a function of distance separation, $\langle|\Delta v_r|\rangle(\Delta r)$. They found no significant deviation from a constant $|\Delta v_r|$ at any scale in $\Delta r$. From comparisons to the simulations of Bullock & Johnston (2005), Xue et al. concluded that this statistic is not capable of detecting structure against a more smoothly distributed background in phase space made up from well-mixed streams. However, the observed signal was not compared to the expected signal from random realisations.

## 4 STELLAR HALO SIMULATIONS

### 4.1 N-body and galaxy formation model

The mock observations that we use to test the $\xi(< \Delta)$ correlation function are derived from simulations of the accreted stellar halo presented in Cooper et al. (2010). These simulations approximate the dynamics of stars in dwarf satellites of Milky Way-like galaxies by 'tagging' appropriate particles (i.e. those strongly bound within subhaloes) in the Aquarius suite of high-resolution N-body simulations (Springel et al. 2008). Each 'tag' associates a dark matter (DM) particle with a stellar population. This technique is valid in the regime of high mass-to-light ratios, which is supported in this case by observations of stellar kinematics in dwarf galaxies (e.g. Walker et al. 2009).

The tagging method has a single free parameter, the fraction of most bound particles chosen in each DM halo for each assignment of newly-formed stars (see Cooper et al. 2010, for further details). This parameter was fixed (to a value of 1 per cent) by requiring the population of *surviving* satellites at the present day to have a distribution of half-light radius as a function of luminosity consistent with Milky Way and M31 observations[5]. The Cooper et al. models differ from the earlier models of Bullock & Johnston (2005) in that they treat the full cosmological evolution of all satellites self-consistently in a single N-body simulation, and use a comprehensive semi-analytic model of galaxy formation (Bower et al. 2006) constrained by data on large scales and compatible with the observed MW satellite luminosity function. Both the Cooper et al. and the Bullock & Johnston simulations produce highly structured stellar haloes built from the debris of disrupted dwarf galaxies. Other halo components formed *in situ* may be present in real galaxies (e.g. Abadi, Navarro & Steinmetz 2006; Zolotov et al. 2009) but these are likely to be more smoothly distributed than the accreted component (Helmi et al. in preparation).

As in Cooper et al. (2010), we refer to our six simulations as haloes Aq-A, Aq-B, Aq-C, Aq-D, Aq-E and Aq-F. From these simulations, we construct catalogues of tracer stars (representing RGB or BHB stars) by converting the stellar mass assigned to each dark matter particle into an appropriate number of stars. Each DM particle can give rise to many tracer stars if it is tagged with sufficient stellar mass.

The positions and velocities of these tracer stars are interpolated between nearby tagged DM particles in phase space. To accomplish this, the 32 nearest phase space neighbours of each tagged particle are identified using the procedure described below. The mean dispersion in each of the six phase-space coordinates is then calculated for each particle by averaging over these neighbours. These dispersions define a 6D ellipsoidal Gaussian kernel centred on the particle, from which the positions and velocities of its tracer stars are drawn randomly. Each progenitor object (set of tagged DM particles accreted as members of a single subhalo) is treated individually in this smoothing operation, i.e. particles are smoothed using only neighbours from the same progenitor (so there is no 'cross talk' between streams from different progenitors). This procedure can be thought of as a crude approximation to running our original simulation again including each tracer star as a test particle.

The 'distance in phase space' used to identify neighbours in the interpolation scheme is defined by a scaling relation between distances in configuration space and velocity space[6]. For each progenitor, we adopt an individual scaling which corresponds to making the median pairwise interparticle separation of its particles in configuration space (at $z = 0$) equal to their median separation in velocity space. In practice, the value of this scaling parameter makes very little difference to the results we present, when compared to the extreme choice of selecting only 32 velocity or position neighbours (disregarding the other three coordinates in each case). Giving more weight to configuration-space neighbours smears out velocity substructure within the debris of a progenitor (for example, where two wraps of a stream pass near one another). Giving more weight to velocity neighbours has the opposite effect – stars can be interpolated over arbitrarily large separations in configuration space, but coherent velocity structures are preserved. Therefore, the 'optimal' choice is the scaling which balances smoothing in configuration space against smoothing in velocity space.

To quantify this balance between smoothing in configuration and velocity space, we compute six smoothing lengths for each particle, $\epsilon_{x,i}$ and $\epsilon_{v,i}$, where $i$ represents a single dimension in space or velocity. We compute these as the spherically averaged dispersion in position and velocity, respectively, taken over the 32 phase-space neighbours of the particle. We define the 'optimum' choice of scaling for *each* progenitor galaxy as that which minimises the quantity

$$\sigma_\epsilon^2 = \left( \frac{1}{\bar{\epsilon}_{x,\mathrm{min}}} \sum_{i=0}^{3} \epsilon_{x,i} \right)^2 + \left( \frac{1}{\bar{\epsilon}_{v,\mathrm{min}}} \sum_{i=0}^{3} \epsilon_{v,i} \right)^2 . \quad (6)$$

This is the sum in quadrature of the mean smoothing lengths in configuration and velocity space, normalized respectively by $\bar{\epsilon}_{x,\mathrm{min}}$,

[5] The luminosity function of surviving satellites in these models also agrees with MW and M31 data. This agreement is mostly due to the underlying galaxy formation model.

[6] In this part of the calculation, we are only interested in finding neighbours, so the absolute values of these distances are not important. This scaling of velocity space to configuration space for the purpose of resampling the simulations should not be confused with the $\Delta$ metric we define for our analysis of clustering.

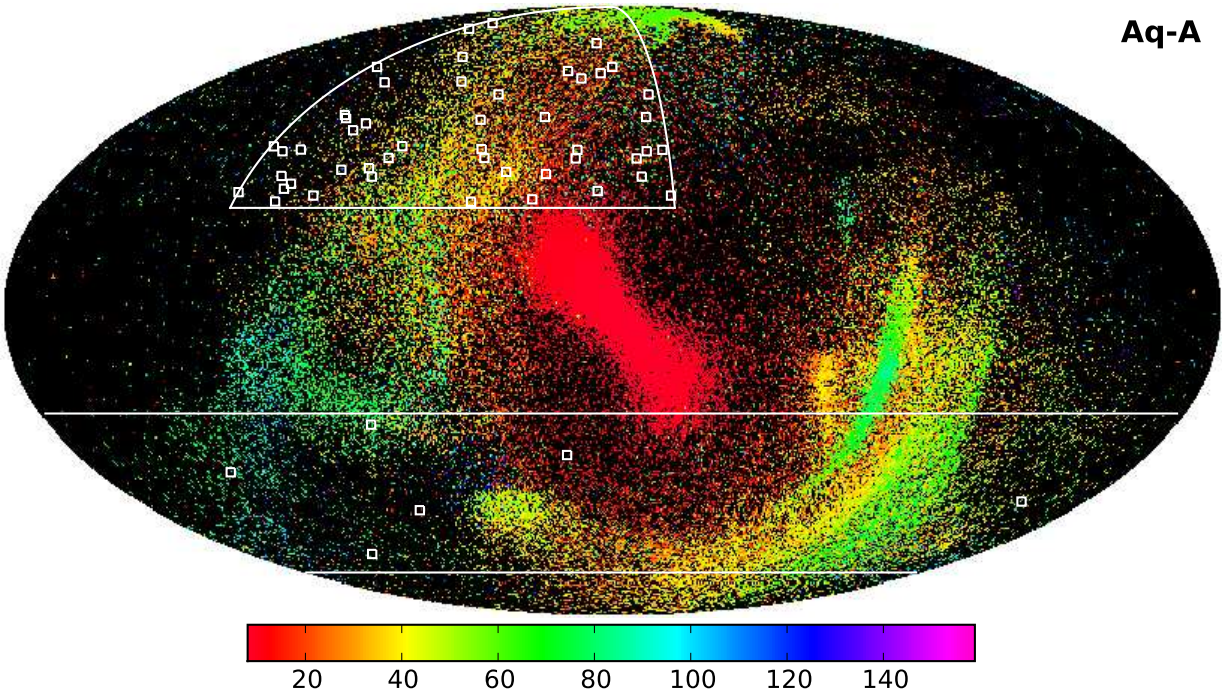


**Figure 2.** Sky distribution of halo RGB stars in simulation Aq-A; colours indicate mean distance from the observer in kiloparsecs. This is a Mollwiede projection in Galactic coordinates, centred on $(l, b) = (0, 0)$, taking $r_\odot = 8\,\mathrm{kpc}$. White lines delineate the Spaghetti survey target areas and white squares a fiducial set of 52 randomly located $1^\circ$ fields (not to scale). The accreted component of the galactic bulge produces the elongated central feature at small distances. Here the major axis of this component is orientated from bottom right (further from the observer) to top left (closer to the observer).

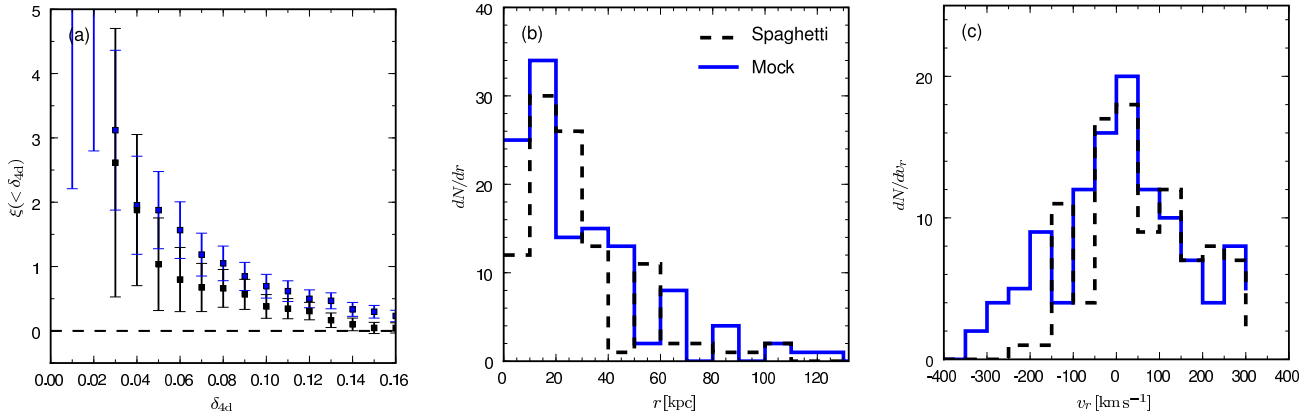


**Figure 3.** Fiducial realisation of a Spaghetti-like survey in halo Aq-A resembling the Spaghetti data (black dashed lines and points). The fiducial fields are those shown in Fig. 2. From left to right the panels show (a) the correlation function in the metric of Starkenburg et al. (2009) and distributions of (b) radial distance and (c) line-of-sight velocity (without correcting for the motion of the local standard of rest). Blue lines and points correspond to mock observations with our standard mock catalogue, sampled at a rate defined by $f_{\mathrm{RGB}}^{-1} = 1000\,M_\odot/\mathrm{star}$.

the 'minimal' mean smoothing length in configuration space (obtained from the 32 nearest configuration space neighbours) and $\bar{\epsilon}_{v,\mathrm{min}}$, the 'minimal' mean smoothing length in velocity space (obtained from the 32 nearest velocity space neighbours). We find that the scaling obtained by matching the median interparticle separations in position and velocity as described above is typically a good approximation to this optimal value – a similar result is discussed in more detail by Maciejewski et al. (2009).

In the Cooper et al. model, the most bound 1% of DM particles in a halo at the time when a given stellar population forms in a satellite are chosen as tracers of that population. Hence, each DM *particle* to which stars are assigned has an individual mass-to-light ratio, M/L, which can be as high as $\sim 1$ (i.e. $M_{\mathrm{stellar}} \sim 10^4\,M_\odot$)

and as low as $\sim 10^{-6}$. This will affect the density of stars seeded by a DM particle independently of the density of its neighbours in phase space (i.e. a low M/L particle will create a denser 'cluster' of tracers relative to a high M/L particle with the same neighbouring positions and velocities). We have tested an alternative approach in which the M/L of each particle in a given progenitor is resampled by distributing the total stellar mass of the progenitor evenly amongst its tagged particles[7]. We find that the extra clustering due to a few 'hot' particles in our default approach makes no difference to our results. Many of the lowest M/L particles are in the bulge component of the halo which is largely excluded (at least from our final SDSS mock observations) by cuts in distance and Galactic latitude.

### 4.2 Tracer Stars

Each N-body dark matter particle in our simulation contributes a number of tracer stars to our mock observations, based on the stellar population with which it has been 'tagged'. In the case of Spaghetti, the tracers are RGB stars meeting the (complex) selection criteria of the survey. Here we assume a global scaling between the stellar mass associated with each N-body particle, $M_\star$, and the number of Spaghetti red giants it contributes to our mock catalogues, i.e. $N_{\rm RGB} = f_{\rm RGB} M_\star$ where $f_{\rm RGB}$ is the number of tracer stars per unit mass of the original stellar population[8]. For each N-body particle, the actual number of RGB stars generated is drawn from a Poisson distribution with mean $N_{\rm RGB}$.

The correlation function results described below are independent of the choice of $f_{\rm RGB}$, provided that the underlying distribution is well sampled at a given scale. We have therefore selected a fiducial value of $f_{\rm RGB}^{-1} = 1000\, M_\odot/{\rm star}$. Fig. 3 shows that with this value, the normalisations of the radial velocity and distance histograms in a Spaghetti-like field set are similar to those actually observed in Spaghetti. Morrison (1993) determined a value of $f_{\rm RGB}^{-1} \sim 620\ L_{V,\odot}/{\rm star}$ for halo giants (RGB and AGB) in the Solar neighbourhood, corresponding to approximately the same ratio (assuming $M/L \sim 2$ and a minimal AGB contribution).

We obtain distances and radial velocities to each RGB star assuming a randomly oriented vector of length $r_\odot = 8$ kpc linking the observer to the Galactic centre. Each random placement of the observer on the 'Solar shell' is referred to below as one *random realisation* of the mock catalogue. Wherever the observer is placed on this spherical shell, Galactic longitude and latitude are defined in the same way with respect to the footprint of the survey, with $(l, b) = (0, 0)$ being the vector directed from the observer to the centre of the halo. As there is no Galactic plane in our simulations (which contain only the accreted component of the halo and the bulge), there is no direct constraint on the orientation of the 'rotation axis' of the galaxy seen by the observer (this is more significant in the context of the SDSS survey, so in Section 6 we use the shape of the halo to fix the orientation of the Galactic plane). We adopt a Solar motion of $U, V, W = (10, 5.2, 7.2)\ {\rm km\,s^{-1}}$ and a velocity of the local standard of rest about the Galactic centre $v_{\rm LSR} = 220\ {\rm km\,s^{-1}}$.

[7] This is almost equivalent to choosing M/L only once, at the time in the simulation when the progenitor falls into the main halo (similar to the lower-resolution model of De Lucia & Helmi 2008).

[8] We do this as we prefer to make a straightforward comparison with the observational data in this paper. In principle, the age and metallicity information associated with each stellar population in our model could be used to populate an individual colour-magnitude diagram for each N-body particle, and make a detailed prediction for the appropriate number of tracers.

### 4.3 Spaghetti survey mock catalogues

We use our stellar halo models to create mock catalogues of individual RGB stars, matched to the parameters of the Spaghetti survey described above (including a cut in distance to select stars between 7 kpc and 130 kpc, the approximate range of the Spaghetti sample). Four different instruments with three different field sizes were used in the actual survey. Here we adopt the mean field size of $\sim 0.5\,{\rm sq.deg.}$ (Dohm-Palmer et al. 2000) for all 52 fields corresponding to the spectroscopic sample. We locate these fields randomly within the constraints on sky area given above. We average over many different random sets of field positions when making comparisons based on a fixed observer position. In both model and data, the random sample $\langle RR \rangle$ is generated by reshuffling distances and velocities in the mock catalogue 1000 times, following Starkenburg et al. (2009).

Fig. 3 verifies that our Aq-A model halo contains structures quantitatively similar to those seen by the Spaghetti survey. To illustrate this we have specifically chosen a set of fields for which the distributions of distance and velocity are well-matched. As we will show in the following section, however, there is a large amount of variance in the recovered signal among mock Spaghetti surveys adopting different placements of fields (for a fixed observer), and among different observer positions on the Solar shell.

## 5 APPLICATION OF THE $\Delta$ METRIC

In this section we describe our choice of the weight factor $w_v$ in the $\Delta$ metric (Eqn. 1). We then analyse the clustering of the Spaghetti and SDSS using our mock catalogues, and discuss the limitations of Spaghetti.

### 5.1 Distance - velocity scaling

There is no clearly well-motivated way to choose a value of the velocity-to-distance scaling $w_v$; without a physical justification, it must be treated as a free parameter. The choice of $w_v$ determines the scale of substructure to which the correlation function is most sensitive. Naively we expect this to be the typical width and transverse velocity dispersion of a 'stream'. It is preferable to fix this parameter in a universal manner that does not depend on any particular survey parameters or geometry. We make a fiducial choice of $w_v$ as follows.

In each simulated halo we adopt the SDSS-like survey configuration discussed below (without observational errors or assumptions about the location of the Sun). We construct (separately) one dimensional distributions of the separation in radial distance and velocity between stars. We generate many random realisations of these distributions by first convolving each simulated star with Gaussian kernels of width $8\,{\rm kpc}$ (distance) and $80\,{\rm km\,s^{-1}}$ (velocity), and then drawing from the equivalent 'smoothed' distributions. The kernel sizes were chosen as a compromise between signal (diminished by undersmoothing) and noise (increased by oversmoothing). Using these random realisations we construct correlation functions for each distribution. These two correlation functions are shown for halo Aq-A in Fig. 4. Although the signals are intrinsically weak, they have a very similar shape for both distributions, each with a characteristic 'turnover' scale. Matching

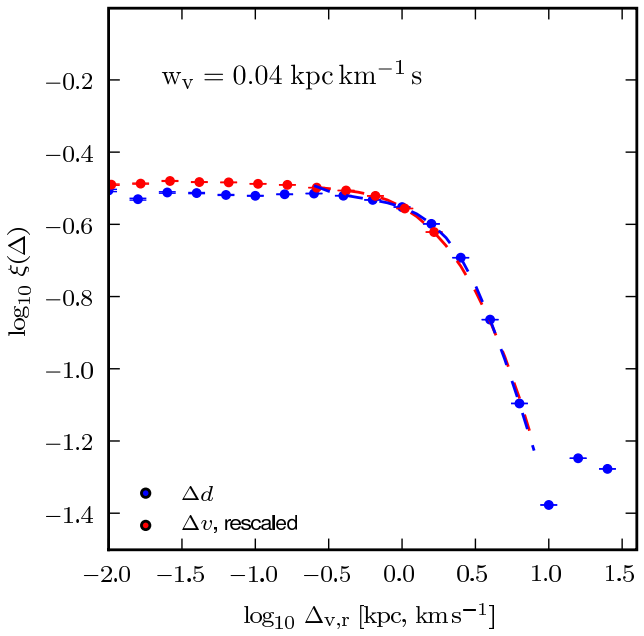


**Figure 4.** Correlation functions in space separation (blue) and velocity separation (red) for stars in halo Aq-A. The velocity separation correlation function has been scaled to match the turnover in the configuration space separation correlation function by a factor $w_v = 0.04\,\mathrm{kpc\,km^{-1}\,s}$.

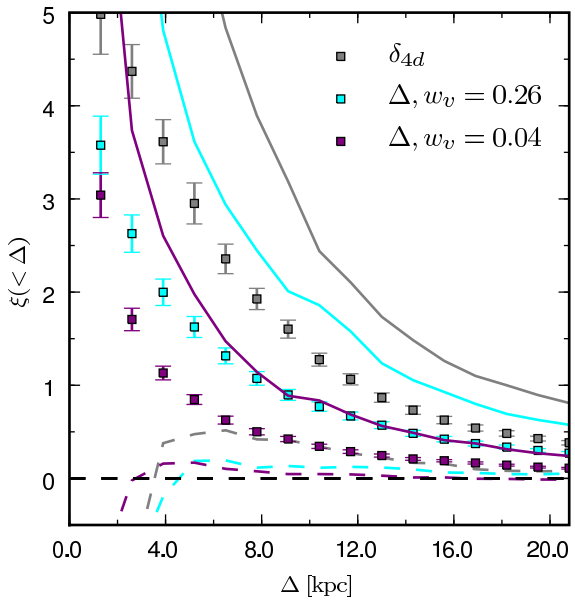


**Figure 5.** The $\Delta$ metric cumulative correlation function $\xi(<\Delta)$ for mock Spaghetti surveys in halo Aq-A. Each set of points with error bars shows the mean and standard error of 200 random realisations; dashed and solid lines indicate $10^{\mathrm{th}}$ and $90^{\mathrm{th}}$ percentiles of the distribution in each bin, respectively. Purple lines/points correspond to $w_v = 0.04$ and cyan lines/points to $w_v = 0.26$. The $\delta_{4\mathrm{d}}$ metric of Starkenburg et al. (without $\eta$ terms, see text) is also shown (in grey). These correlation functions are those of mock catalogues with no observational errors.

this scale in the two correlation functions corresponds to $w_v \sim 0.04 \pm 0.01\,\mathrm{kpc\,km^{-1}\,s}$ for the six haloes, which we adopt as a fiducial value. We caution that although the scales on which we match the one-dimensional correlation functions are somewhat smaller than the smoothing scales we adopt to create the random distributions, this does not guarantee that our choice of $w_v$ is unaffected by our choice of smoothing.

This is not a very satisfactory way of fixing $w_v$. However, in practice our conclusions are not strongly sensitive to the precise value we adopt. Values of $w_v$ of the order of 0.01–1.0 $\mathrm{kpc\,km^{-1}\,s}$ result in very similar $\xi(\Delta)$ correlation functions. Values lower than $0.01\,\mathrm{kpc\,km^{-1}\,s}$ recover very little signal. Values above $1\,\mathrm{kpc\,km^{-1}\,s}$ treat $1\,\mathrm{km\,s^{-1}}$ velocity differences as equivalent to $> 1$ kpc separations in space, and so make the cumulative correlation function very noisy on small scales for only a marginal increase in the overall signal (this noise in turn results in more scatter between the signals measured by different observers). We find that our choice of $w_v \sim 0.04\,\mathrm{kpc\,km^{-1}\,s}$ is a reasonable compromise. The above choices can be compared with the approach of Starkenburg et al. (2009), who take the ratio of the Spaghetti survey limits in radial distance and velocity to obtain $w_v = 0.26\,\mathrm{kpc\,km^{-1}\,s}$. Either value is acceptable to illustrate the utility of our approach, our intention in this paper. We therefore adopt $w_v \sim 0.04\,\mathrm{kpc\,km^{-1}\,s}$.

Fig. 5 shows the cumulative correlation function with the metric of Eqn. 1 averaged over many randomly placed observers for a mock Spaghetti survey[9]. We show results for $w_v = 0.26\,\mathrm{kpc\,km^{-1}\,s}$ and $w_v = 0.04\,\mathrm{kpc\,km^{-1}\,s}$. The scatter between observers is much larger for a Spaghetti-like survey than for the SDSS-like surveys we focus on below. Overall the choice of $w_v$ makes little difference. A large value of $w_v$ increases the amplitude of the mean signal but also increases the scatter between observers. We also show in Fig. 5 the equivalent correlation function using the $\delta_{4\mathrm{d}}$ metric of Starkenburg et al. (2009) (rescaled to units of kiloparsecs, and neglecting the $\eta$ terms discussed above). These two metric definitions give similar results. From Fig. 5 we conclude that with either metric, the signal measured by any single observer in this halo using a Spaghetti-like configuration may not be representative of the halo overall[10].

By applying their method to the simulations of Harding et al. (2001), Starkenburg et al. (2009) find that for $\delta_{4\mathrm{d}} < 0.05$ over 80 per cent of pairs in their correlation function are made up of stars originating in the same progenitor satellite (so called 'true' pairs). We find that our choice of $w_v = 0.04$ recovers a similar fraction ($\sim 70$ per cent) of true pairs using the same method. This supports the claim of Starkenburg et al. (2009) that the majority of pairs detected in the Spaghetti data are likely to be genuine. In the largest-scale bin used in our cumulative correlation function, $\Delta < 21$ kpc, the fraction of true pairs is $\sim 40$ per cent. The efficiency with which true pairs are recovered is most relevant to structure-finding applications and is not important for our global statistic. However, such high fractions do indicate that the clustering of stars from individual progenitor galaxies makes a substantial contribution to the signals we recover.

We conclude that, as expected, the $\Delta$ metric is very similar to the $\delta_{4\mathrm{d}}$ metric of Starkenburg et al. in the limit of small angular separations and with $\eta = 1$ in Eqns. 4 and 5. The $\Delta$ metric has the advantage of a more straightforward definition. Furthermore, from

[9] When comparing our results with those of Starkenburg et al. (2009), it is important to note that the $\Delta$ metric distance has units of kiloparsecs, whereas $\delta_{4\mathrm{d}}$ defined by Eqn. 3 has units of [130 kpc].

[10] This conclusion does not invalidate the *detection* of significant substructure in the Spaghetti data with this approach by Starkenburg et al. (2009)

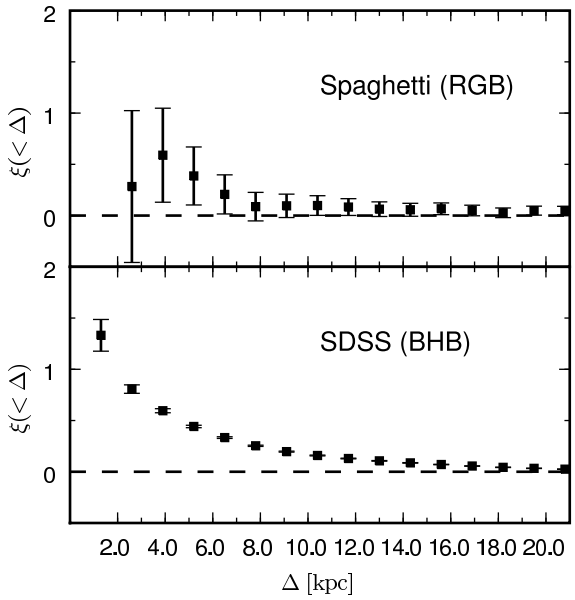


**Figure 6.** Upper panel: The $\Delta$ metric cumulative correlation function $\xi(<\Delta)$ of the Spaghetti survey (Poisson error bars). Lower panel: $\xi(<\Delta)$ for the SDSS DR6 sample of Xue et al. (2008).

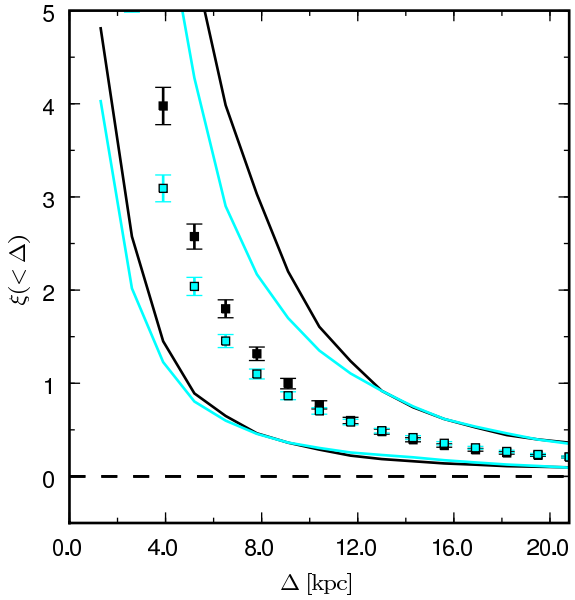


**Figure 7.** $\xi(<\Delta)$ for an 'enhanced' Spaghetti-like survey (in halo Aq-A) having 200 pencil beams at high latitude in each Galactic hemisphere. Black points show the mean signal (with its standard error), and black lines the 10–90 per cent range of the mean signal averaged over many observers. Cyan points and lines correspond the distribution of signals for many surveys (with randomly placed fields) carried out by a single randomly chosen observer.

the wide scatter around the mean signal in Fig. 5 it is clear that an individual Spaghetti-like survey cannot place a strong constraint on the structure of the halo overall using either statistic.

### 5.2 Observations of $\xi(<\Delta)$ Spaghetti and SDSS

In Fig. 6 we show $\xi(<\Delta)$ computed with the same observational data used in Fig. 1. In the case of the Spaghetti data, there is a clear reduction in the clustering signal relative to Fig. 1. This can be most easily understood in the limit of small angular separations as discussed above. In this limit, the transverse and perpendicular components of distance are coupled in the $\Delta$ metric. In effect, angular separations are translated to transverse distances, with a fixed angular scale corresponding to a larger '$r_\perp$' at larger $r_\parallel$. Pairs of a given $\phi_{ij}$ in the $\delta_{4\mathrm{d}}$ metric can therefore be displaced to a relatively higher- or lower-separation bin in the $\Delta$ metric, depending on their radial distance. This effect does not significantly reduce the number of $\langle DD \rangle$ pairs. However, the number of random pairs in small-separation bins increases on average in the $\Delta$ metric, where pairs physically close to the observer but separated by large angular distances can be assigned small transverse separations. This reduces the significance of observed pairs in small separation bins. Most of the signal in the Spaghetti survey comes from only a few excess data pairs, so this dilution has a significant effect.

The cumulative correlation function of the much larger SDSS BHB sample is very similar in both the $\delta_{4\mathrm{d}}$ and $\Delta$ metrics. This suggests that the differences between the two metrics are small in practice. This is not surprising. To within an order of magnitude, the weights in $\delta_{4\mathrm{d}}$ maximise the signal in Spaghetti, which is similar in spirit to the empirical approach described in the previous section.

### 5.3 The halo-average signal and pencil-beam surveys

A useful survey should recover 'global' properties of the halo with high significance, i.e. properties that are insensitive to the observer's position on the surface of the Solar shell. We have shown that our metric can recover a clustering signal due to structure in the halo, using the data from a Spaghetti-like survey. However, Fig. 5 demonstrates that the signal from such a survey measured by any individual observer is extremely sensitive to the placement of its pencil beams. A corollary of this is that the 'halo average' signal of many observers also has a large scatter - in other words, the survey is limited by 'cosmic variance'. In this sense, 'blind' application of $\xi(<\Delta)$ to the data from Spaghetti cannot constrain the properties of the stellar halo, even if those data were complete in the surveyed fields.

The SDSS sample of BHB stars provides a much more significant measure of the global signal, as we demonstrate in the following section. However, SDSS is an expensive survey. Furthermore, in future it may be interesting to compare the correlation functions of different tracers that can be surveyed only in the manner of Spaghetti (in which costly spectroscopy of individual targets is required to construct the sample). In Fig. 7 we show a pencil beam survey covering both Galactic caps with 200 fields, using the distance limits of Spaghetti. Although the scatter remains large, the signal now deviates significantly from zero at small separations. The local observer's signal is also a reasonable measure of the halo average. This is encouraging, because it implies that only a relatively modest improvement is required over Spaghetti to provide useful constraints on halo structure (as suggested by Starkenburg et al. 2009). Halo simulations such as those of Cooper et al. (2010) could be used to optimise a particular survey to detect clustering due to accreted substructures, accounting for the effects of observational errors and incompleteness, which we have not addressed here.

## 6 CLUSTERING OF SDSS BHB STARS

We make mock SDSS surveys in a similar manner to the mock Spaghetti surveys described above using the same catalogues of tracer stars. Although the number density of BHB stars per halo is not equal to that of RGB stars even under our assumption of a uniform underlying population, the $\xi(< \Delta)$ statistic is not sensitive to the absolute number of tracers. Accounting for inhomogeneity in the stellar populations of the halo may introduce more significant differences between the distributions of BHB and RGB stars. The 'bias' that may be introduced by choosing particular tracer populations is beyond the scope of this paper.

For a given observer location, we select all tracer stars in our catalogue within the SDSS DR6 footprint having distances $6 - 60\,\mathrm{kpc}$. We do not include any stars gravitationally bound to satellites. However, we do include stars in their tidal tails (which by our definition are part of the stellar halo). We note that in their study of the pairwise velocity distribution of their BHB catalogue, Xue et al. (2009) exclude nine stars deemed to belong to globular clusters. We do not exclude these stars from our analysis of the SDSS data.

In Fig. 8 we compare all six haloes with the observations of SDSS BHBs shown in Fig. 6. It is clear from Fig. 8 that the distribution of signals around the mean of many observers is much narrower than that for the Spaghetti survey show in Fig. 5. Because of the more extensive sky coverage of SDSS, fewer random orientations are required for the global halo signal to converge than in the case of Spaghetti. We find 80 randomly placed observers to be sufficient. For each of these observers, we compute the correlation function using 20 random 'shufflings' of the data.

Panel (a) of Fig. 8 shows that all haloes have an average signal clearly deviating from zero at small separations. Furthermore, significant differences are apparent in the average clustering signal between each of our six simulated haloes. This demonstrates that this statistic can distinguish between plausible alternatives for the structure of the MW halo in a ΛCDM model. In particular, halo Aq-C (purple) shows considerably more 'substructure' at all distances. Visually, this halo is dominated by several massive, dynamically young streams from recently-accreted satellites (see figures 6 and 7 of Cooper et al. 2010). Halo Aq-B (cyan), which has a lower clustering signal, is centrally concentrated, and contains only a small number of coherent low-mass streams. Halo Aq-E (blue) is also centrally concentrated, but differs from Aq-B in having a substantial accreted 'thick disc' component. Haloes Aq-A (black) and Aq-D (green) are of intermediate mass and contain a variety of complex features in different stages of mixing. Halo Aq-A contains a Sagittarius-like stream and several widely dispersed clouds/shells, while Aq-D shows a highly elongated coherent complex of bright streams. Halo Aq-F is unusual – most of its stars are accreted in a late major merger and it is substantially brighter than the Milky Way at the Solar radius. For these reasons it more closely resembles the 'shell'-dominated haloes of some elliptical galaxies than those of Milky-Way like spirals. Its clustering signal is the lowest of our six models and is dominated by a much larger number of stars close to the observer. Other than the 'shell' system, very little coherent structure is apparent in this halo.

In summary, the clustering signal detected by our $\xi(< \Delta)$ statistic broadly reflects the visual impression of the amount of spatially coherent structure in each halo. However, to make a realistic comparison with the Milky Way data, it is important to account for the effects of observational errors (which become more significant at small scales). The consequences of including observational errors in the simulated data are shown in panel (b) of Fig. 8. Each distance and radial velocity is perturbed by representative Gaussian errors of $\sigma_d = 5\%$ and $\sigma_v = 15\ \mathrm{km\,s^{-1}}$ respectively (Xue et al. 2008). As expected, this smears out structure at small scales and suppresses the clustering signal in all haloes. Halo C in particular shows a much reduced signal, although the similarities and differences between the haloes are mostly preserved. An exception to this is the reduction in the signal from Aq-D relative to that of Aq-E, at small metric separations. This is likely to be because the stellar halo in Aq-D is more extensive than that of Aq-E (see Cooper et al. 2010). The small-scale clustering signal is presumably generated by stars further from the observer in Aq-D than in Aq-E. These stars have larger distance errors.

It is also relevant that the sky distribution of structure in our simulated haloes stellar halo is not random with respect to the orientation of the Galactic plane (and hence, with respect to the coverage of the Sloan survey). It is a strong prediction of our models that a planar alignment of halo debris will be observed in the Milky Way. This correlated alignment of tidal features is a natural consequence of structure formation in ΛCDM, because the direction from which most massive substructures are accreted is determined by the filamentary nature of the large-scale structure. These 'accretion axes' are correlated with the spin vector and shape of the main halo, and hence with the likely orientation of the Galactic plane (Libeskind et al. 2005; Li & Helmi 2008; Lovell et al. 2010).

In panel (c) of Fig. 8 we show the clustering signals for the case in which we restrict the observer's position to a 'Galactic plane'. We define this as the plane perpendicular to the minor principal axis of the inertia tensor for halo stars with galactocentric radii less than 3 kpc (see Cooper et al. 2010). The observer is randomly located on a circle of radius 8 kpc in this plane, and the 'parity' of the Galactic poles is also randomized (this is important, because SDSS coverage is mostly concentrated around only one Galactic pole and the halo need not be symmetric).

As described in Cooper et al. (2010), an 'accreted bulge' component is present in all of our haloes. In all cases the shape of this component is triaxial (oblate in the case of halo Aq-E) and is similar to the shape of the dark halo in the same region. Our choice of alignment ensures that the major axes of this component lie in the Galactic plane. Hence, the practical effect of restricting the observer to the 'Galactic plane' is to prevent this (nearby) component from intruding into the SDSS footprint at high Galactic latitudes. There are other plausible choices of Galactic alignment (for example, relative to the shape or spin vectors of the entire dark halo). However, any choice is somewhat arbitrary in the absence of a self-consistent simulation of a galactic disc[11]. We choose to orient the Galactic plane relative to the accreted bulge because, of all the plausible choices, it has the most significant influence on the clustering signal. Even in this case, the overall effect is modest. In some cases (e.g. halo C) the amplitude of the mean signal increases slightly and the scatter between signals decreases.

Finally, in panel (d) of Fig. 8 we show the results of constraining the orientation of the Galactic plane and also including observational errors. This provides a more realistic comparison with the observational data. Haloes with a high degree of coherent substructure on large scales (represented by our halo Aq-C) appear to be incompatible with the Milky Way, as do those with very little sub-

[11] In a full hydrodynamic simulation the effects of feedback and adiabatic contraction may also make the halo more spherical (e.g. Tissera et al. 2010; Abadi et al. 2010).

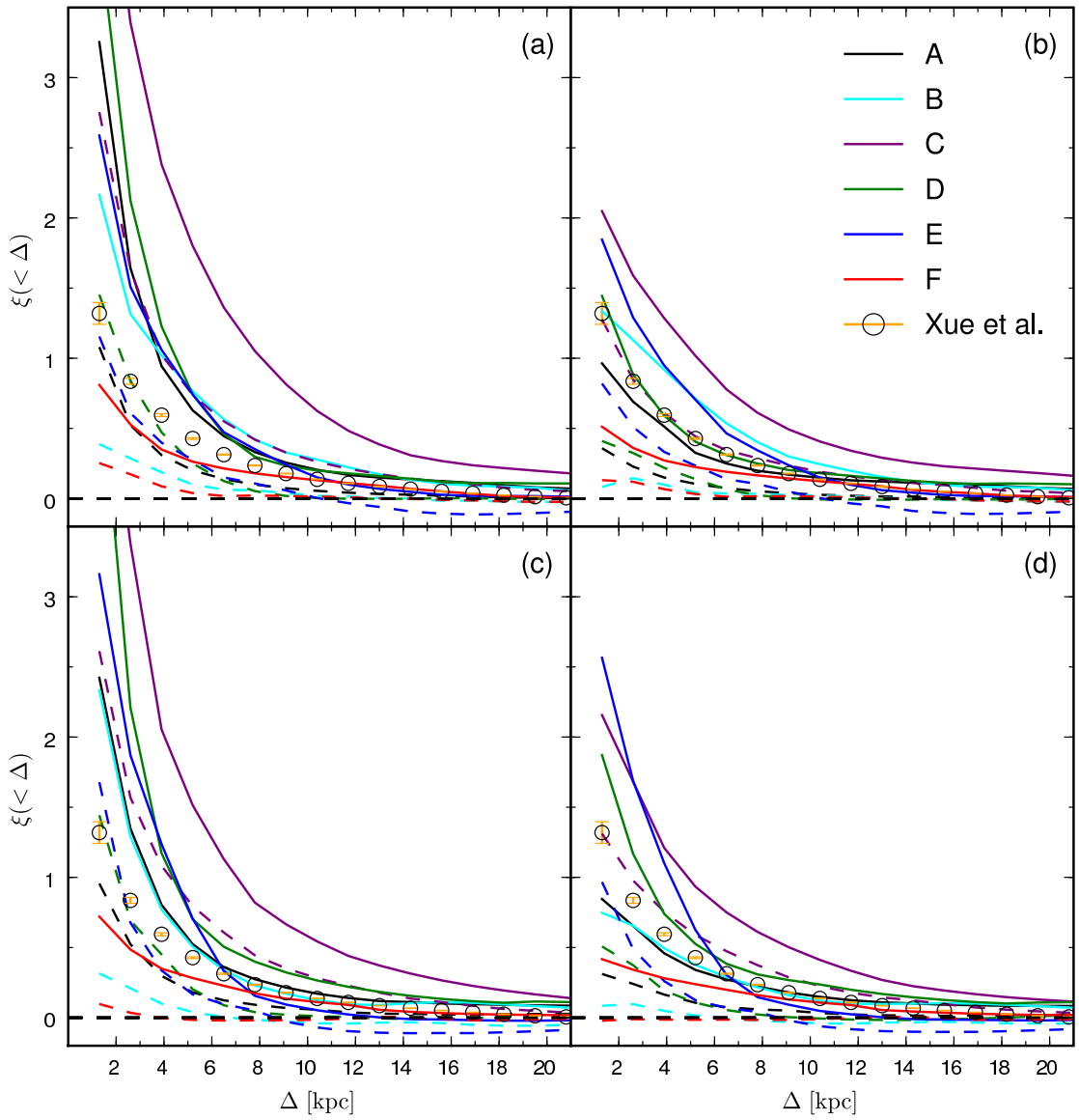


**Figure 8.** (a) The mean clustering signal $\xi(<\Delta)$ of 20 Solar-shell observers with a SDSS-like survey. Each halo simulation is represented by a pair of lines of the same colour (indicated in the legend). A solid line indicates the $90^{\rm th}$ percentile of the distribution of $\xi(<\Delta)$ in each radial bin, and a dashed line the $10^{\rm th}$ percentile. $\xi(<\Delta)$ for MW BHB stars from the catalogue of Xue et al. (2008) is shown with open circles and orange error bars (identical in all four panels). (b) Mock observations convolved with observational errors of $\sigma_d/d = 0.05$ and $\sigma_v = 15\,{\rm km\,s^{-1}}$. (c) Mock observations oriented such that the Galactic 'Z' direction is aligned with the minor axis of the halo ellipsoid within 3 kpc, but not convolved with errors. (d) Mock observations with errors of (b) and constrained alignment of (c).

structure due to the dominance of stars from a single progenitor accreted on a radial orbit (Aq-F). Our four remaining haloes are all broadly consistent with the data, with the MW having slightly more than the typical amount of substructure on the smallest scales in our analysis ($\Delta \sim 1\,{\rm kpc}$).

## 7 CONCLUSIONS

We have analysed the correlation function $\xi(<\Delta)$ of halo stars, defining their separations in four dimensions of phase space using a metric (which we call $\Delta$) in readily-obtained observables (angular and radial separation and radial velocity difference). A statistic of this type usefully quantifies kinematic and spatial substructure in the halo, and can easily be applied to observational data and catalogues generated from theoretical models. This analysis is particularly well suited to analysing the distant halo – other methods for studying clustering in many dimensions may be more suitable for the 'fine grained' data on the nearby halo that will soon be obtained by the *Gaia* mission (e.g. Gomez et al. 2010). We aim to apply other clustering statistics to our haloes in future work.

We have computed $\xi(<\Delta)$ for mock observations of six stellar haloes in the simulations of Cooper et al. (2010). All of these haloes were formed from satellites disrupted within $\Lambda$CDM dark haloes selected as plausible hosts for the Milky Way. Our statistic

distinguishes quantitatively between these six qualitatively different scenarios. On average, all six haloes show statistically significant correlations on scales in our metric equivalent to $\sim 1-10$ kpc.

We find that small pencil-beam surveys such as Spaghetti sample too few stars and cover too small an area to be well suited to analysis with our proposed statistic. Instead we have analysed a much larger catalogue of BHB star observations from the SDSS (Xue et al. 2008). The current Milky Way data are consistent with those simulated haloes having a moderate degree of spatial and kinematic substructure. Haloes dominated by massive coherent structures and haloes with little or no substructure appear less consistent with the Milky Way.

Our comparison between our models of $\Lambda$CDM stellar haloes and the Milky Way data demonstrates the potential of the statistical approach suggested by Starkenburg et al. (2009) and Xue et al. (2009). However, our application can only serve as a basic consistency test for substructure kinematics in these models. Although the properties of our model stellar haloes may vary between different semi-analytic implementations, it is nevertheless encouraging that our fiducial model passes this basic test.

A number of aspects of this approach could be improved with further work. It seems desirable to use well-measured radial velocity data to boost a clustering signal (such as our correlation function) above that obtained from configuration space data alone (as demonstrated by Starkenburg et al. 2009). However, no means of including these velocity data is, as yet, well-supported by theory (including the approach we adopt here). The parametrised metric we have used to illustrate the concept of scaling radial velocity separations to 'equivalent' configuration space separations is straightforward choice, and it is empirically useful in recovering a measurable signal. Nevertheless, we have not found any compelling or 'universal' way to select (or even justify the assumption of) the scaling parameter ($w_v$). Improving either the definition of the metric itself or the means of fixing this parameter is a clear priority for extensions of this approach. A similar issue affects the weighting of velocity information in clustering algorithms (e.g. Sharma et al. 2010).

A more extensive comparison between a stellar halo model and the observational data should also take account of effects such spectroscopic incompleteness. In addition, the fraction of the stellar halo expected to be made up from components formed 'in situ' (i.e. within Milky Way-like galaxies themselves, rather than in accreted systems) is not well constrained (Abadi et al. 2006; Zolotov et al. 2009, see also Helmi et al., in preparation). It seems reasonable to expect in situ halo components to be distributed smoothly, with spherical or axial symmetry. The absence of these components in our models may therefore lead to an artificially high clustering signal. It is possible to place crude limits on the fraction of stars in a 'missing' smooth component, for example by comparing the RMS variation of projected star counts in our models (Helmi et al., in preparation) to the Milky Way (Bell et al. 2008). However, the uncertainties involved are substantial. Including an ad hoc smooth component in the clustering analysis above would also require assumptions about its velocity distribution, which are even less straightforward. There is a pressing requirement for complete observational samples, even if they do not probe the most distant halo. The LAMOST Galactic survey is likely to be the first to approach this goal.

In summary, we have taken a first step in adapting a well-studied cosmological statistic, the two-point correlation function, to the study of the Milky Way halo. We have presented an application making few modelling assumptions, using high-resolution N-body simulations of stellar halo assembly. Our statistic can distinguish between plausible alternatives for the structure of Milky Way-like stellar haloes. At least one of our six simulations is consistent with currently available data for the Milky Way. With further refinements and more data, this approach can provide a practical and productive way to quantify the structure of the Milky Way halo for comparison with numerical models.

## ACKNOWLEDGEMENTS

The authors thank Heather Morrison and the Spaghetti Survey team for making their data available prior to publication. APC acknowledges an STFC studentship and thanks Else Starkenburg for useful discussions. SMC acknowledges the support of a Leverhulme Research Fellowship. CSF acknowledges a Royal Society Wolfson Research Merit Award. AH acknowledges funding support from the European Research Council under ERC-StG grant GALACTICA-24027. Fig. 2 was produced with the `HEALPy` implementation of `HEALPix` [http://healpix.jpl.nasa.gov, Górski et al. 2005].

## REFERENCES

Abadi M. G., Navarro J. F., Fardal M., Babul A., Steinmetz M., 2010, MNRAS, 407, 435
Abadi M. G., Navarro J. F., Steinmetz M., 2006, MNRAS, 365, 747
Bell E. F., Zucker D. B., Belokurov V., Sharma S., Johnston K. V., Bullock J. S., Hogg D. W., Jahnke K., et al., 2008, ApJ, 680, 295
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Brown A. G. A., Velázquez H. M., Aguilar L. A., 2005, MNRAS, 359, 1287
Brown W. R., Geller M. J., Kenyon S. J., Beers T. C., Kurtz M. J., Roll J. B., 2004, AJ, 127, 1555
Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931
Cooper A. P., Cole S., Frenk C. S., White S. D. M., Helly J., Benson A. J., De Lucia G., Helmi A., et al., 2010, MNRAS, 406, 744
De Lucia G., Helmi A., 2008, MNRAS, 391, 14
Dohm-Palmer R. C., Mateo M., Olszewski E., Morrison H., Harding P., Freeman K. C., Norris J., 2000, AJ, 120, 2496
Doinidis S. P., Beers T. C., 1989, ApJ, 340,
Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487
Gómez F. A., Helmi A., 2010, MNRAS, 401, 2285
Gómez F. A., Helmi A., Brown A. G. A., Li Y. S., 2010, MNRAS, 1196
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Harding P., Morrison H. L., Olszewski E. W., Arabadjis J., Mateo M., Dohm-Palmer R. C., Freeman K. C., Norris J. E., 2001, AJ, 122, 1397
Helmi A., de Zeeuw P. T., 2000, MNRAS, 319, 657
Johnston K. V., Bullock J. S., Sharma S., Font A., Robertson B. E., Leitner S. N., 2008, ApJ, 689, 936
Lemon D. J., Wyse R. F. G., Liske J., Driver S. P., Horne K., 2004, MNRAS, 347, 1043
Li Y.-S., Helmi A., 2008, MNRAS, 385, 1365
Libeskind N. I., Frenk C. S., Cole S., Helly J. C., Jenkins A., Navarro J. F., Power C., 2005, MNRAS, 363, 146
Lovell M., Eke V., Frenk C., Jenkins A., 2010, ArXiv e-prints, astro-ph/1008.0484
Maciejewski M., Colombi S., Alard C., Bouchet F., Pichon C., 2009, MNRAS, 393, 703
Morrison H. L., 1993, AJ, 106, 578
Morrison H. L., Mateo M., Olszewski E. W., Harding P., Dohm-Palmer R. C., Freeman K. C., Norris J. E., Morita M., 2000, AJ, 119, 2254
Morrison H. L., Norris J., Mateo M., Harding P., Olszewski E. W., Shectman S. A., Dohm-Palmer R. C., Helmi A., et al., 2003, AJ, 125, 2502

Morrison H. L., Olszewski E. W., Mateo M., Norris J. E., Harding P., Dohm-Palmer R. C., Freeman K. C., 2001, AJ, 121, 283
Re Fiorentin P., Helmi A., Lattanzi M. G., Spagna A., 2005, A&A, 439, 551
Schlaufman K. C., Rockosi C. M., Allende Prieto C., Beers T. C., Bizyaev D., Brewington H., Lee Y. S., Malanushenko V., et al., 2009, ApJ, 703, 2177
Searle L., Zinn R., 1978, ApJ, 225, 357
Sharma S., Johnston K. V., 2009, ApJ, 703, 1061
Sharma S., Johnston K. V., Majewski S. R., Muñoz R. R., Carlberg J. K., Bullock J., 2010, ArXiv e-prints, astro-ph/1009.0924
Skrutskie M. F., Cutri R. M., Stiening R., Weinberg M. D., Schneider S., Carpenter J. M., Beichman C., Capps R., et al., 2006, AJ, 131, 1163
Springel V., Wang J., Vogelsberger M., Ludlow A., Jenkins A., Helmi A., Navarro J. F., Frenk C. S., et al., 2008, MNRAS, 391, 1685
Starkenburg E., Helmi A., Morrison H. L., Harding P., van Woerden H., Mateo M., Olszewski E. W., Sivarani T., et al., 2009, ApJ, 698, 567
Tissera P. B., White S. D. M., Pedrosa S., Scannapieco C., 2010, MNRAS, 406, 922
Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Wyn Evans N., Gilmore G., 2009, ApJ, 704, 1274
White S. D. M., Springel V., 2000, in The First Stars, Weiss A., Abel T. G., Hill V., eds., p. 327
Xue X. X., Rix H. W., Zhao G., 2009, Research in Astronomy and Astrophysics, 9, 1230
Xue X. X., Rix H. W., Zhao G., Re Fiorentin P., Naab T., Steinmetz M., van den Bosch F. C., Beers T. C., et al., 2008, ApJ, 684, 1143
Zolotov A., Willman B., Brooks A. M., Governato F., Brook C. B., Hogg D. W., Quinn T., Stinson G., 2009, ApJ, 702, 1058

This paper has been typeset from a TEX/ LATEX file prepared by the author.