The radial profile and flattening of the Milky Way's stellar halo to 80 kpc from the SEGUE K-giant Survey

Xiang-Xiang Xue[1], Hans-Walter Rix[1], Zhibo Ma[2], Heather Morrison[2], Jo Bovy[3*], Branimir Sesar[1], William Janesh[2]

ABSTRACT

We characterize the radial density, metallicity and flattening profile of the Milky Way's stellar halo, based on the large sample of spectroscopically confirmed giant stars from SDSS/SEGUE-2 (Xue et al. 2014), spanning Galactocentric radii 10 kpc $\leq r_{GC} \leq$ 80 kpc. After excising stars that were algorithmically attributed to apparent halo substructure (including the Sagittarius stream) the sample has 1757 K giants , with metallicities precise to ~ 0.2 dex and distances to $\sim 16\%$. Compared to BHB stars or RR Lyrae, giants are more readily understood tracers of the overall halo star population, with less bias in age or metallicity. The well-characterized selection function of the sample enables forward modelling of those data, based on ellipsoidal stellar density models, $\nu_*(R,z)$, with Einasto profiles and (broken) power laws for their radial dependence, combined with a model for the metallicity gradient and the flattening profile. Among models with constant flattening, these data are reasonably well fit by an Einasto profile of $n = 3.1 \pm 0.5$ with an effective radius $\mathrm{r_{eff}} = 15 \pm 2$ kpc and a flattening of $q = 0.7 \pm 0.02$; or comparably well by an equally flattened broken power-law, with radial slopes of $\alpha_{in} = 2.1 \pm 0.3$ and $\alpha_{out} = 3.8 \pm 0.1$, with a break-radius of $r_{break} = 18 \pm 1$ kpc; this is largely consistent with earlier work. We find a modest, but significant metallicity gradient within the 'outer' stellar halo, [Fe/H] decreasing outward. If we allow for a variable flattening $q = f(\mathrm{r}_{GC})$, we find the distribution of halo giants to be considerably more flattened at small radii, $q(10\ \mathrm{kpc}) \sim 0.57$, compared to $q(> 30\mathrm{kpc}) \sim 0.8$. Remarkably, the data are then very well fit by a single power-law of index $\sim 4.2 \pm 0.1$ of the variable $r_q \equiv \sqrt{R^2 + (z/q(r))^2}$. In this simple and better fitting model, there is a break

[1]Max-Planck-Institute for Astronomy Königstuhl 17, D-69117, Heidelberg, Germany

[2]Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106, USA

[3]Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA

[3*]John Bahcall Fellow

in flattening at ~ 20 kpc, instead of a break in the radial density function. While different parameterizations of the radial profile vary in their parameters, their implied density gradient, $\partial \ln \nu_* / \partial \ln r$, is stable along a direction intermediate between major and minor axis; this gradient is crucial in any dynamical modelling that uses halo stars as tracers.

Subject headings: galaxies: individual(Milky Way) – Galaxy: halo – Galaxy: stellar content – stars: K giants

1. Introduction

The Milky Way's extended stellar halo contains only a small fraction ($\lesssim 1\%$) of the Galaxy's stars, but is an important diagnostic of its formation and dark matter distribution. The position-kinematics-abundance substructure in the stellar halo reflects the Galaxy's formation history, whether halo stars were born *in situ* or are disrupted satellite debris. By now, individual stars are the by far largest sample of kinematic tracers with 10 kpc $\lesssim$ $\mathrm{r}_{GC} \lesssim 100$ kpc (as opposed to globular clusters or satellite galaxies), and are hence the best tracers to determine the mass profile of Milky Way's dark matter halo in this radial range. It is obvious that good kinematic tracer samples should be large, and cover a wide radial range with accurate individual distances. However, beyond this, the spatial distribution of the tracers — in particular their radial profile — must be well understood to use such tracers for dynamical inferences. This point is perhaps clearest when considering the Jeans (1915) Equation, even in its simplest, spherical and isotropic version: the tracer density profile, $\nu_*(r)$, in particular its logarithmic radial derivative, $\partial \ln \nu_* / \partial \ln r$, affects the inferred enclosed mass, $M(< r)$, almost linearly. If we do not know the local power-law exponent to better than, say, 25%, we cannot infer the mass to better than 25% irrespective of the size and quality of the kinematic sample. To a somewhat lesser extent, the inferred mass also depends on the flattening of the tracer population. Yet, at present there is little consensus on the shape and the radial profile of the stellar halo.

The most straightforward way to quantify the stellar halo distribution is via star counts. Because this method requires large samples of well-understood completeness, it is often applied to photometric catalogs. Early studies adopted star counts to analyze globular clusters (Harris 1976), RR Lyrae variables (Hawkins 1984; Wetterer & McGraw 1996; Vivas & Zinn 2006), blue horizontal branch (BHB) stars (Sommer-Larsen 1987), a combination of BHBs and RR Lyraes (Preston et al. 1991), a star sample near the north galactic pole (Soubiran 1993), or K dwarfs (Gould et al. 1998), and found the stellar halo is well fitted by a single power-law ($\nu \approx (\mathrm{r}_{GC})^{-\alpha}$) with index $\alpha = 3 - 3.5$ and flattening of $q = 0.5 - 1$.

However, Saha (1985) found that RR Lyrae are well described by a broken power-law with $\alpha \sim 3$ out to 25 kpc, and $\alpha \sim 5$ beyond 25 kpc.

These earlier studies were based on a few hundred objects at most. In recent years, with the development of large sky surveys, the sizes of the photometric samples have expanded by a factor of more than 10. Robin et al. (2000) used a wide set of deep star counts in a pencil-beam survey at high and intermediate Galactic latitudes to model the density profile and found the best-fit density profile with a flattening of 0.76 and a power index of 2.44. Siegel et al. (2002) found that $70,000$ stars in seven Kapteyn selected areas are consistent with a power-law density with index of 2.75 and flattening of 0.6. However, both Robin et al. (2000) and Siegel et al. (2002) used galaxy models to work out the contamination of star counts by the disc population. This technique is not ideal for the lowest density stellar component in the Galaxy because any errors in the disk or thick disk will overwhelm the halo results. Morrison et al. (2000) used halo stars near the main-sequence turnoff from the "SPAGHETTI" survey to map the Galactic halo and found the halo density law over Galactocentric radii of 5-20 kpc and z-heights of 2-15 kpc followed a flattened power-law halo with the flattening of 0.6 and a power index of 3. Bell et al. (2008) used ~ 4 million color-selected main-sequence turnoff stars from SDSS DR5 up to 40 kpc to find a "best-fit" oblateness of the stellar halo of $0.5 - 0.8$, and the density profile of the stellar halo is approximately described by a power-law with index of $2-4$. Subsequently, Sesar et al. (2011) used $27,544$ near-turnoff main-sequence stars out to ~ 35 kpc selected from the Canada-France-Hawaii Telescope Legacy Survey to find a flattening of the stellar halo of 0.7 and the density distribution to be consistent with a broken power low with an inner slope of 2.62 and an outer slope of 3.8 at the break radius of 28 kpc, or an equally good Einasto profile (Einasto 1965) with a concentration index of 2.2 and effective radius of 22.2 kpc. Deason et al. (2011) analyzed $\sim 20,000$ A-type photometric stars selected from the Sloan Digital Sky Survey data release 8 (Ahn et al. 2012) and obtained a best-fitting broken power-law density with an inner slope of 2.3, and an outer slope of 4.6, with a break radius at 27 kpc and a constant flattening of 0.6. Subsequently Deason et al. (2014) found a very steep outer halo profile a power law of r^{-6} beyond 50 kpc, and yet steeper slopes of $\alpha = 6 - 10$ at larger radii.

In addition, several pieces of work point to variations of the stellar halo flattening with radius. Preston et al. (1991) found that the density distribution of RR Lyrae follows a power-law with $\alpha \sim 3.2$, together with a variable flattening changing linearly from 0.54 at center to 1 at 20kpc. Subsequent work (Sesar et al. 2011; Deason et al. 2011) found evidence for flattening, but no evidence for a change with radius.

Spectroscopic maps of the stellar halo beyond ~ 20 kpc in practice require luminous

post main-sequence (MS) stars, as turn-off or other MS stars are too faint for current wide-field spectrographs. RR Lyrae and BHB stars have repeatedly been used as tracers to study the halo density profile, because they have precise distances and are bright enough to be observed at radii out to ~ 100 kpc (Sesar et al. 2010; Xue et al. 2008, 2011; Deason et al. 2011, 2014). Yet, such stars, most prevalent in particularly old and metal-poor populations (Bell et al. 2010; Xue et al. 2011) are known to have a different structure and profile from the metal-poor red giant branch stars (RGB, K-giants). Moreover, Preston et al. (1991); Morrison et al. (2009) claimed that BHB stars in the inner halo had a different radial profile from RR Lyraes. To this end, it is crucial to construct the halo shape and radial profile of the stellar halo in K giants. The SEGUE K giant survey (Xue et al. 2014) presented a catalog of K giants with unbiased distance estimates good to 16% , metallicities, velocities and photometric information, drawn from the Sloan Extension for Galactic Understanding and Exploration (Yanny et al. 2009, SEGUE), which contains ~ 150 stars beyond 60 kpc. SEGUE was focused on the halo in its targeting, and has two phases SEGUE-1 and SEGUE-2. The target-selection of K giants changed a lot during SEGUE-1, but SEGUE-2 adopted a unified K-giant selection, so one can understand and model their selection function well for K giants observed in SEGUE-2. Due to the well-understood selection function, it is possible to determine the halo profile and shape of these tracers, which is the main goal of the present paper. Specifically, we set out to describe the stellar halo distribution, presuming that the density is stratified on (oblate) spheroids, with a radial profile from $10 - 80$ kpc that can be characterized by simple functional forms (either an Einasto profile or broken power-law). We also explore the metallicity dependence of the shape and radial profile of the stellar halo.

In the next section, we lay out the properties and the selection function of the SEGUE K giants. In §3, we present the method of fitting a series of parameterized models to SEGUE-2 K giants, explicitly and rigorously considering the selection function. This step is key in obtaining accurate radial profiles and metallicity gradient model at the meantime. The results for the stellar halo's radial profile, flattening are presented in §4, along with the metallicity gradient in the stellar halo. Finally, §5 discusses the comparison between our results and previous work, and implications for dynamical models.

2. SEGUE-2 K giants and their selection function

The Sloan Digital Sky Survey (SDSS; York et al. 2000) is an imaging and spectroscopic survey covering roughly a quarter of the sky, which has both *ugriz* imaging (Fukugita et al. 1996; Gunn et al. 1998; Stoughton et al. 2002; Pier et al. 2003; Eisenstein et al. 2011) and low resolution spectra ($\lambda/\Delta\lambda \sim 2000$). SEGUE is one of the key projects and has two phases,

SEGUE-1 and SEGUE-2, which aim to explore the nature of stellar populations from 0.5 kpc to 100 kpc (Yanny et al. 2009, and Rockosi et al. in prep.). SEGUE-2 spectroscopically observed around 120,000 stars, focusing on the stellar halo of the Galaxy.

To understand the underlying spatial distribution of the K giants on the basis of this spatially incomplete sample, we need to understand (and account for) the probability that a star of a given luminosity, color and metallicity ends up in the sample, given its direction and distance. Spectroscopic surveys of the Milky Way are inevitably affected by such selection effects (cf. Rix & Bovy 2013), often referred to as "selection biases". They arise from a set of objective and repeatable decisions of what to observe, necessitated by the survey design. In particular, only a small fraction of the sky was covered by SEGUE plates, and for most plates only a fraction of stars that satisfy the photometric selection criteria could be targeted with fibers. Finally, not all targeted stars yield spectra good enough to result in a catalog entry, i.e. had signal-to-noise ratio (S/N) high enough to verify that they are giants and yield a metallicity measurement. Bovy et al. (2012) and Rix & Bovy (2013) spelled out how to incorporate this selection function in fitting a parameterized model for the stellar density and we follow their approach in this and the next Section.

Both the SEGUE-1 and SEGUE-2 surveys targeted halo giant star candidates, using a variety of photometric and proper motions cuts. About 90% of the final K-giant sample came from objects observed as l -color K-giant targets. The l-color (Lenz et al. 1998) is a photometric metallicity indicator for stars in the color range $0.5 < (g-r)_0 < 0.8$, designed to select metal-poor K giants[1]. As mentioned in §1, only the selection function for K giants observed in SEGUE-2 can be understood well because SEGUE-2 adopted a consistent color-magnitude cut to select K giants throughout its entire survey: $15.5 < g_0 < 18.5$, $r_0 > 15$, $0.7 < (u-g)_0 < 3$, $0.5 < (g-r)_0 < 0.8$, $0.1 < (r-i)_0 < 0.6$, and $l - color > 0.09$. Therefore, we restrict our analysis to this category. We also require that l-color K giant candidates have good proper motion measurements ≤ 11 mas yr^{-1}. Since broadband photometry is a poor main-sequence *vs.* giant discriminator, not all stars targeted under the above criteria will be giants. The subsequent identification of K-giants is based solely on their spectroscopic properties. As described in Xue et al. (2014), this requires spectra that have good Mg index and stellar atmospheric parameters determined by the SEGUE Stellar Parameter Pipeline (SSPP; Lee et al. 2008a,b, 2011), but have no strong G band.

To understand the selection function, we must compare the color-magnitude distribution of these spectroscopically confirmed K giants to the analogous distribution of all possible

[1]https://www.sdss3.org/dr9/algorithms/segue_target_selection.php#S2_table.

photometric l -color K giant candidates. Figure 1 shows these two distributions (as contours and gray-scale, respectively), summed over all SEGUE-2 plates. As the marginalized histograms at the sides of the panels show, these two distributions are nearly indistinguishable: the chance of a photometric candidate being confirmed as a K-giant is independent of color and magnitude. This simplifies the subsequent analysis and is testament to SEGUE-2's consistency of target selection, targeting and spectral analysis.

However, while the selection function is constant with apparent magnitudes and dereddened colors, it varies from plate to plate, in particular with the Galactic latitude of the plate. Given the pencil-beam nature of the SEGUE survey, it makes sense to specify the selection fraction plate by plate. For each plate we define the number of spectroscopically confirmed K giant as N_{spec}, and the number of l-color K giant candidates in the plate (both those that were targeted, and those that were not) as N_{phot}. Thus, the plate-dependent selection function (shown as Figure 2) is given by

$$S(plate) \;=\; \frac{N_{spec}}{N_{phot}} \tag{1}$$

As we want to analyze the spatial distribution, we also restrict our sample to SEGUE-2 l -color K giants that have reliable distance estimates from Xue et al. (2014). To eliminate the contamination from the disk component, we cull K giants with [Fe/H] > -1.2 and $|z| < 4$ kpc, which leads to a final sample of 2413 l-color K giants. Figure 3 illustrates the basic sample properties: its sky coverage and spatial distribution (without accounting for the selection function) and the distribution of metallicity against distance from Xue et al. (2014). The stars' distribution reflects the pencil-beam pattern of the SEGUE survey; their Galactocentric distances range from 7 kpc to 85 kpc; their mean metallicity is -1.75 dex, with some being as metal-poor as -3.5. The reader should note that these metallicities are not typical of the halo as a whole because of our choice to exclude all stars with [Fe/H] > -1.2.

3. Modelling the Stellar Density and Metallicity Distribution in the Halo

In this Section we lay out a forward modeling approach to describe the spatial and metallicity distribution of the stellar halo with a set of flexible, but ultimately smooth and symmetric functions. Both the modelling practicalities and the astrophysics require that we fit the spatial distribution and the abundance distribution simultaneously. We defer to the §4 the question of how sensitively such a "smooth" description depends on the question of including or excluding stars that are presumably members of recognizable substructure (cf. Belokurov et al. 2006; Bell et al. 2008).

3.1. A Parameterized Model for the Observables

We presume that the stellar halo distribution can be sensibly approximated by a spheroidal distribution with a parameterized radial profile, allowing for radial variations in the metallicity distribution, $\nu_\star\big(\mathrm{x,y,z,[Fe/H]}\big) = \rho(\mathrm{[Fe/H]}\,|\,\mathrm{r_q,p_{Fe}}) \times \nu_\star(\mathrm{r_q}\,|\,\mathrm{p_H})$. The combination of $\mathrm{p_H}$ and $\mathrm{p_{Fe}}$ denotes the *halo parameters*, and $r_q \equiv \sqrt{R^2 + (z/q(r))^2}$ is the basic Galactocentric radial coordinate, given the flattening $q(r)$.

In this section, we spell out a straightforward and rigorous approach to determine the posterior probability distributions for *halo parameters* in light of the above data, our knowledge of the SEGUE-2 selection function, and well-established astrophysical prior on the luminosity function of giant stars. The number of *halo parameters* depends on the complexity of the model stellar halo distribution. This approach essentially follows Bovy et al. (2012) and Rix & Bovy (2013).

Since we already have good estimates of the distance modulus $\mathcal{DM}$ and the r-band absolute magnitude M_r for all objects in the sample (Xue et al. 2014), we treat $(\mathcal{D}_\odot, M_r, \mathrm{[Fe/H]})$ as the observables defining $\mathcal{D}_\odot \equiv (\mathcal{DM}, l, b)$, rather than $(m, c, \mathrm{[Fe/H]}, l, b)$, as it makes the fitting formalism more intuitive. However, the r-band apparent magnitude m and (dereddened) color c (referring in particular to $g-r$), will appear explicitly in the observational selection function. We denote the angular selection function as $S(l,b) \equiv S\big(\mathrm{plate(l,b)}\big)$ (see Eq.(1)), the magnitude-color selection function by $S(m(\mathcal{DM}, M_r), c(M_r, \mathrm{[Fe/H]}))$, the metallicity selection function by S([Fe/H]), additional spatial cuts by $S(\mathcal{D}_\odot)$, and the luminosity selection function as $S(M_r \mid \mathrm{[Fe/H]})$, expressed in terms of the "observables" above. We denote the prior external information on the distribution of absolute magnitude on the giant branch as $p(M_r)$; for old and metal-poor populations this is well established through cluster luminosity functions (see, for example, Xue et al. 2014). The luminosity selection function $S(M_r \mid \mathrm{[Fe/H]})$ will appear explicitly, because stars low on the giant branch were removed (depending on [Fe/H]) in Xue et al. (2014) to avoid confusion between RGB and red clump (RC) stars. Figure 4 shows the limits of $\mathrm{M_r}$ for a given [Fe/H]. In short, we need to spell out and then quantify all terms that matter for predicting the *rate function*, i.e. the expected frequency or probability of finding a sample member with a given $\mathcal{DM}, l, b, M_r, \mathrm{[Fe/H]}$ if a halo with p_H and p_{Fe} were true.

Given p_H and p_{Fe}, the expected rate function for finding a star with $(\mathcal{D}_\odot, M_r, \mathrm{[Fe/H]})$ is then

$$\lambda(\mathcal{D}_\odot, \mathrm{M_r}, [\mathrm{Fe/H}] \mid \mathrm{p_H}, \mathrm{p_{Fe}}) =$$
$$|\mathrm{J(x,y,z;}\mathcal{D}_\odot)| \times \nu_\star\big(\mathcal{D}_\odot \mid \mathrm{p_H}\big) \times \rho([\mathrm{Fe/H}] \mid \mathcal{D}_\odot, \mathrm{p_{Fe}}) \times \mathrm{p(M_r)} \times$$
$$S(\mathcal{D}_\odot) \times S(l,b) \times S([\mathrm{Fe/H}]) \times S(M_r \mid [\mathrm{Fe/H}]) \times S\Big(m(\mathcal{DM}, M_r), c(M_r, [\mathrm{Fe/H}])\Big). \quad (2)$$

This rate function is simply a concise way of specifying what set of observations we expect, given an assumed halo model, including all observational selection effects and pertinent prior information. The price for putting this is one place, is that the rate function has quite a number of distinct terms. In the next two subsections, we spell out and discuss these terms. To make the rate function a probability, it must be normalized for every new set of trial model parameters ($p_H,\ p_{Fe}$); this is the only time-consuming step in such modelling.

3.2. Models for the Spatial and [Fe/H] Distribution of the Stellar Halo

Following a number of previous studies, we presume that the overall radial density profile of the halo can be described by an Einasto profile or by (multiply-)broken power-law, with the density stratified on surfaces of constant r_q in all cases; in addition, we devise a simple model for radial [Fe/H] variations.

3.2.1. Einasto profile

The Einasto profile (Einasto 1965) is the 3D analog to the Sérsic profile (Sérsic 1963) for surface brightnesses, and has been used to describe the halo density distribution (Sesar et al. 2011; Deason et al. 2011; Merritt et al. 2006):

$$\nu_\star(r_q) \equiv \nu_0 \exp\left\{-d_n\left[(r_q/\mathrm{r_{eff}})^{1/n} - 1\right]\right\}, \quad (3)$$

where ν_0 is the (here irrelevant) normalization, $\mathrm{r_{eff}}$ is the effective radius , n is the concentration index, and $\mathrm{d_n} \approx 3\mathrm{n} - 1/3 + 0.0079/\mathrm{n}$, for $\mathrm{n} \geq 0.5$. The free parameters of an Einasto profile with a constant flattening are $\mathrm{p_H} = (\mathrm{r_{eff}}, \mathrm{n}, \mathrm{q})$.

3.2.2. Broken Power-Law Profiles

Broken (or even multiply-broken) power-laws (BPL) are another family of functional forms that has been used extensively to described the radial profile of the Galactic stellar halo (Saha 1985; Sesar et al. 2011; Deason et al. 2011, 2014). In most cases the change in the power-law index, $d \ln \nu_* / d \ln r$ has been taken to be a step function. We adopt:

$$\nu_\star \left(\mathrm{r_q} \right) = \begin{cases} \nu_0 \ \mathrm{r_q}^{-\alpha_{\mathrm{in}}}, & \mathrm{r_q} \leq \mathrm{r_{break}} \\ \nu_0 \times \mathrm{r_{break}}^{(\alpha_{\mathrm{out}} - \alpha_{\mathrm{in}})} \times \mathrm{r_q}^{-\alpha_{\mathrm{out}}}, & \mathrm{r_q} > \mathrm{r_{break}} \end{cases} \tag{4}$$

In addition to the flattening and the (irrelevant) normalization, a (singly) broken power-law has three parameters: $\alpha_{in},\ \alpha_{out},\ r_{break}$. Of course, this profile family encompasses a single power law (SPL), and it can be generalized to a multiple-broken power-law (triple power-law, TPL), by introducing an additional pair of $\left(\alpha, r_{break}\right)$ (cf. Deason et al. 2014).

3.2.3. Halo Flattening

While Preston et al. (1991) found evidence for decrease of flattening with increasing radius, others did not (Sesar et al. 2011; Deason et al. 2011). As there is evidence that at least the innermost part of the halo is quite flattened (Carollo et al. 2010), we explore how our sample can inform us about radial variations in the flattening of the stellar halo beyond r_{GC} =10 kpc. To date no particular functional form to parameterize the possible variation of halo flattening has been established; therefore, we consider the functional form for q(r) as:

$$\mathrm{q}\left(\mathrm{r}\right) = \mathrm{q}_\infty - \left(\mathrm{q}_\infty - \mathrm{q}_0\right) \exp\left(1 - \frac{\sqrt{\mathrm{r}^2 + \mathrm{r}_0^2}}{\mathrm{r}_0}\right), \tag{5}$$

where q_0 is the flattening at center, changing to q_∞ at large radii, with r_0 is the (here, exponential) scale radius, over which the change of flattening occurs.

3.2.4. Radial Variations in the Metallicity Distribution

The existence of a halo metallicity gradient is currently controversial (see, for example Carollo et al. 2007; Schönrich et al. 2011; Fernández-Alvar et al. 2015). This is partially due to the lack of in situ tracers for which accurate abundance measures are possible with low resolution spectroscopy, and partially due to the lack of attention to the effects of the complex SEGUE selection function. This is the first paper which has applied a forward

modeling approach to the K giants in the halo, addressing both of these concerns. We choose a mixture model for the halo metallicity distribution as follows. The metal-rich component can be described by a Gaussian with a mean at -1.4 dex and a sigma of 0.2 dex, and the metal-poor component follows a Gaussian with a mean at -2.1 dex and a sigma of 0.35 dex. Therefore, we suppose the metallicity *distribution*, $\rho([\mathrm{Fe/H}] \mid \mathcal{D}_\odot, p_{Fe})$ as a radially varying combination of these two metallicity components. The metallicity distribution model is also expressed as a function of $\mathrm{r_q}$. We choose:

$$\rho([\mathrm{Fe/H}]|\mathrm{r_q}, \mathrm{p_{Fe}}) = \mathrm{f_0}\times(\mathrm{r_q}/20\mathrm{kpc})^\gamma \times \mathcal{G}(-1.4, 0.2) + (1-\mathrm{f_0}\times(\mathrm{r_q}/20\mathrm{kpc})^\gamma)\times\mathcal{G}(-2.1, 0.35) \quad (6)$$

where $\mathrm{r_q}$ is the same as defined in Eq. 3; $\mathcal{G}$(mean, dispersion) are Gaussian functions. Besides the common parameter q, the metallicity gradient has two other parameters, $\mathrm{f_0}$ and γ (we mark them as $\mathrm{p_{Fe}}$).

3.3. Selection Effects and Prior Information

The remaining terms in Eq. 2 incorporate the various selection effects and pieces of prior information into the prediction of the rate function, and are given in the following.

The Jacobian term is given by

$$|J(x, y, z; \mathcal{DM})|_{plate} = \Omega_{plate} \cdot \ln 10/5 \cdot \left(10^{\frac{\mathcal{DM}}{5}-2}\ \mathrm{kpc}\right)^3, \quad (7)$$

where $\Omega_{plate} = 7\ deg^2$ and we presume the Jacobian to be constant within a plate. The prior on the absolute magnitude distribution along the giant branch (Xue et al. 2014) is given by

$$\mathrm{p(M_r)} \propto \begin{cases} 10^{0.32\mathrm{M_r}}, & \text{if } \mathrm{M_{r\ min,obs}} < \mathrm{M_r} < \mathrm{M_{r\ max,obs}} \\ 0, & \text{otherwise,} \end{cases} , \quad (8)$$

The sample's metallicity cuts, aimed at eliminating bulge and thick disk contributions, as well as any spurious metallicity determinations, $-3.5 < [\mathrm{Fe/H}] < -1.2$, are reflected in

$$\mathrm{S([Fe/H])} \propto \begin{cases} 1, & \text{if } [\mathrm{Fe/H}]_{\mathrm{min,obs}} < [\mathrm{Fe/H}] < [\mathrm{Fe/H}]_{\mathrm{max,obs}} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

The spatial cuts to geometrically excise any bulge and thick disk stars are

$$S(\mathcal{D}_\odot)) \propto \begin{cases} 1, & \text{if } |\mathrm{z}(\mathcal{D}_\odot)| > 4\ \mathrm{kpc} \\ 0, & \text{otherwise} \end{cases}$$

(10)

The reliable distance determination requires minimizing the contamination of RGB stars by red clump stars, which translates into [Fe/H]-dependent absolute magnitude cuts (Figure 4):

$$S(M_r \mid [\mathrm{Fe/H}]) = \begin{cases} 1, & \text{if } \mathrm{M_{min}([Fe/H]) < M_r < M_{max}([Fe/H])} \text{ for } [\mathrm{Fe/H}] \in [-3.5, -1.2] \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

The set of observed SEGUE plates leads to a rather sparse angular selection function in (l, b), illustrated in Figures 2 and 3:

$$\mathrm{S(l_{plate}, b_{plate})} = \begin{cases} \frac{\mathrm{N_{spec}}}{\mathrm{N_{phot}}}, & \text{if in plate} \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

Finally, the apparent magnitude and color cuts (see Figure 1) for the SEGUE targeting are reflected in

$$\mathrm{S\Big(m_r(\mathcal{DM}, M_r), c(M_r, [Fe/H])\Big)} \propto \begin{cases} 1, & \text{if } \mathrm{m_{min} < m_r < m_{max}} \text{ and } \mathrm{c_{min} < c < c_{max}} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

Taken these terms together, and following Bovy et al. (2012), we can now directly calculate the likelihood of the data given (p_H, p_{Fe}) and the rate function:

$$\mathscr{L}\big(\mathrm{data_i|p_H, p_{Fe}}\big) \;=\; \mathrm{c_\lambda^{-N_{KG}}} \prod_{\mathrm{i=1}}^{\mathrm{N_{KG}}} \lambda(\mathrm{M_{ri}}, \mathcal{DM}_\mathrm{i}, \mathrm{[Fe/H]_i, l_i, b_i|p_H, p_{Fe}}), \tag{14}$$

where the normalization $\mathrm{c_\lambda}$ is the integral over the volume in the $(\mathcal{DM}, \mathrm{l, b, M_r, [Fe/H]})$ space,

$$\mathrm{c_\lambda} \;=\; \sum_{\mathrm{i=1}}^{\mathrm{N_{plate}}} \int\int\int \lambda(\mathrm{M_r}, \mathcal{DM}, \mathrm{[Fe/H], l_{plate}, b_{plate}|p_H, p_{Fe}}) \mathrm{dM_r} d\mathcal{DM} \mathrm{d[Fe/H]} \tag{15}$$

Here we have already performed the *dldb* integral for each plate. This normalization integral is the computationally most expensive part of the parameter estimates. But it can be

computed efficiently using Gaussian quadratures, where we adopt $48 \times 20 \times 20$ transformation points in $\mathcal{DM}$, M_r and [Fe/H] space, and where the parameter-independent parts such as the Jacobian term, the luminosity prior, and selection functions can be pre-computed on a dense grid. We assume the priors on all parameters $(\mathrm{p_H}, \mathrm{p_{Fe}})$ to be flat over the pertinent range; therefore the posterior distribution function (*pdf*) of the parameters, $\mathrm{p}(\mathrm{p_H}, \mathrm{p_{Fe}}|\{\mathrm{data}\})$, is proportionate to the likelihood (Eq. 14), differing only in its units ('1/parameters' *vs.* '1/data'). We then vary the (p_H, p_{Fe}) to sample the parameter *pdf* using emcee, which implements an efficient Markov chain Monte Carlo technique[2] (Foreman-Mackey et al. 2013).

4. Results

We now present the results of applying the modelling from §3 to the sample of §2. We illustrate these results in two ways: first, by showing the joint *pdf*s of the halo model parameters; second, we show what the radial and flattening profiles actually look like, by drawing samples from these parameter *pdf*s. We start out with the simplest model (the one with the fewest parameters), an Einasto profile of constant flattening, to illustrate the results, present the basic gradient in the [Fe/H] distribution and to explore the impact of excising stars that are manifestly in sub-structures. We then proceed to include variable flattening, and the broken power-laws as radial profiles. The results are summarized in Table 1.

Figure 5 illustrates the result of the simplest model we fit to the entire data set, by sampling the parameter *pdf* using Eq. 14. This model has three parameters for $\nu_*(r_q \,|\, p_H)$, $p_H = \{\mathrm{r_{eff}}, n, q\}$, and two parameters for the metallicity distribution $\rho([\mathrm{Fe/H}] \,|\, r_q, p_{Fe})$, namely $p_{Fe} = \{f_0, \gamma\}$. The projections in Figure 5 of the *pdf* involving only p_H are shown in grey, while those also involving p_{Fe} are shown in blue. This Figure illustrated that the sample size and data quality of the sample is sufficient to provide formally very well constrained parameters.

4.1. Radial Gradients in the Halo's Metallicity Distribution

Figure 5 shows that the mix of the intermediate ($\langle [Fe/H] \rangle = -1.4$) and metal-poor ($\langle [Fe/H] \rangle = -2.1$) [Fe/H]-components changes with radius (i.e. $\gamma \neq 0$): the relative importance of the metal poor component increases towards large radii, by a factor of 1.4 ± 0.1 from 10 kpc to 65 kpc. Our best-fit metallicity distribution predicts a mean metallicity of

[2]The user guide for emcee can be found at http://dan.iel.fm/emcee/current/

-1.8 in this radial range. The Figure also shows that the covariances between the fits to the stellar density and to the [Fe/H]-distribution are weak. The resulting [Fe/H]-distribution, projected into the space of the rate function is shown in the inset of Figure 5; this inset illustrates that the distribution of the actual sample members in the r_{GC} − [Fe/H] plane is plausible. Note that nearly all samples stars within $r_{GC} = 10$ kpc, much of the "inner halo" of Carollo et al. (2007), have been eliminated from the fit (see Figure 3); therefore, our result is the detection of an outward metallicity gradient within the "outer" halo.

4.2. The Impact of Halo Substructures on Fitting Smooth Models

As discussed in §1 , both cosmological models and observations imply that a good portion of halo stars, at least beyond 20 kpc, are in substructures. Especially the prominent ones, such as the Sgr stream and the Virgo overdensity can and will affect the fits of smooth models, as pointed out by Deason et al. (2011). Recently, Janesh et al. (2015) used a position-velocity clustering estimator (the 4distance) in combination with a friends-of-friends (FoF) algorithm to identify the stars in the Xue et al. (2014) K-giant sample belonging to substructures; they found that $\sim 27\%$ of the K giants are clearly associated with the Sgr Streams, the Orphan Streams, the Cetus Polar Stream, and other unknown substructures. The results of Janesh et al. (2015) provide a straightforward algorithmic way of excising much of the manifest substructure from the sample. For the Sgr stream (Belokurov et al. 2014) we know the expected position-velocity distribution quite well, which suggest that six of the most distant sample members are likely members of Sgr's trailing arm. These lie at $r_{GC} > 60$ kpc, $120^o < L_\odot < 140^o$, $|B_\odot| < 13^o$ and cluster tightly around $V_{gsr} = 100 \mathrm{kms}^{-1}$, and also eliminate them from the original sample of 2413 l-color K giants, leaving 1757 l-color K giants.

We then repeat the fitting of the same model as illustrated in Figure 5, and find – unsurprisingly – significant differences (Figure 6): an Einasto profile that is more concentrated, has a smaller effective radius and is more flattened, $n = 3.1 \pm 0.5$, $r_{eff} = 15 \pm 2$kpc, and $q = 0.7 \pm 0.02$. The metallicity gradient, however, remains very similar with $f_0 = 0.6 \pm 0.01$ and $\gamma = -0.2 \pm 0.04$. Such a model already provides quite a good fit to the data, as Figure 7 shows: the model prediction for the $\mathcal{DM}$ distribution of the sample members, averaged over all directions and metallicities and accounting for all selection effects, matches the actual $\mathcal{DM}$ distribution quite well. This Figure corresponds to the top-right inset in Figure 6, just marginalized over metallicities.

4.3. Different Functional Forms for the Radial Profile

Following previous approaches, we next explore different functional forms for the radial profile, assuming radially constant flattening.

Figure 8 shows the results analogous to Fig. 6, but for a broken power-law (BPL): at a best-fit, constant flattening of q = 0.7±0.02 we find an inner slope of $\alpha_{\rm in} = 2.1\pm0.2$, distinctly different from the outer slope of $\alpha_{\rm out} = 3.8 \pm 0.1$, with a break radius of $\rm r_{break} = 18 \pm 1$ kpc. The flattening q is consistent with that of the best-fit Einasto profile. Again, also with broken power-law profile p($\mathcal{DM}$, [Fe/H]) remains similar and is consistent with the observations. As the values of $\Delta \ln(\mathscr{L})$ in Table 1 indicate, the Einasto and BPL radial profiles fit the data comparably well. A twice broken power-law (TPL), with the break radii held fixed at the radii suggested by Deason et al. (2014) leads to comparably good fit (see Fig. 9 and Table 1). Compared to Deason et al. (2014), the fit is consistent within 65 kpc; it shows no evidence for a steep drop beyond, but our present sample is vary sparse at these distances. For reference, we also fit the data with a *single* power-law of constant flattening, but on the basis of its $\Delta \ln(\mathscr{L})$ it can be ruled out compared to the BPL and TPL models. Our results confirm that if constant flattening is assumed, the halo profile must be described by a radial profile break at ~ 20 kpc.

4.4. Radial variation of halo flattening

We now proceed to allow more model flexibility by allowing the flattening $q(r)$ to vary with radius, according to Eq. 5; we fit both and Einasto profile and power-laws (BPL and SPL) to the data. Allowing for flattening variations makes for distinctly better fits to the data, as quantified by $\Delta \ln(\mathscr{L})$ (Table 1). These fits imply a *strong* variation of the halo flattening with radius, as the *pdf*s for an Einasto profile in Fig. 10 illustrate: while at large radii $q_\infty \sim 0.78$, the flattening at asymptotically small radii becomes formally very strong, $q_0 \approx 0.2$. Given that the characteristic radius at which this flattening change occurs is $r_{q_0} \approx 6$ kpc, and hence smaller than the minimal radius of all data points at 10 kpc, the actual flattening changed across the radial range constrained by the data is more modest, as Fig. 12 illustrates: q changes from 0.57 at 10 kpc to 0.78 at large radii. Allowing for variable flattening also changes the Einasto parameters quite drastically: the formal effective radius becomes very small, and there is a strong covariance between $\rm r_{eff}$ and concentration n. However, the actual predictions for the slope of the radial profile, $\partial \ln \nu_* / \partial \ln r$, in the radial regime constrained by the data are quite similar between models with constant and variable flattening, as Figs. 13 and 14 show. This just illustrates to compare model fits drawn from different functional form families in their projection in the space of observables, not just in

the space of the parameters.

The analogous *pdf*s of the analogous fit for a broken power-law look at first sight bewildering. However, it reveals that the implied flattening profile is the same as for an Einasto fit. The unconstrained *pdf*s on the break radius simply reflect the fact that the inner and outer power-law slopes become indistinguishable, which is equivalent to a single power-law. Therefore, we turn to fit the data to a single power-law with a varying flattening (see Figure 11). Remarkably, we find a very good fit with a single power-law in r_q, when allowing for variable flattening, $q(r)$ (Table 1). The flattening profile is the same as for an Einasto fit, as also shown in Figure 12: there is effectively a break in the flattening profile at 15-20 kpc. Figs. 13 show that the actual local density slope along an intermediate axis for this single power-law is similar to the other fits. Taken together this implies in the context of these density models, that the stellar halo density is best and simplest described (10-70 kpc) by a single power law in the variable r_q, with the break in the radial profile at 20 kpc replaced by a break in the flattening at a comparable radius (see Figure 12).

5. Discussion and Summary

Using a large spectroscopically confirmed sample of K-giants, well-understood population tracers of the stellar halo, we have attempted to characterize the radial profile, the flattening profile and the metallicity profile of the 'outer halo' from 10 kpc to ~ 70 kpc; we have done so after excising algorithmically identified members of distinct halo substructures from the sample. On the one hand, this analysis has re-emphasised that the choice of the functional forms for the density models matters in casting the results; on the other hand we could show that three robust aspects emerge: the profile of the local density slope, $\frac{\mathrm{d}\ln\nu}{\mathrm{dlnr}}$, when taken along an axis intermediate between major and minor axis, is consistent among all functional profile forms we explored; the stellar halo distribution appears distinctly flatter within 20 kpc; there is a slight, but significant, outward metallicity decrease within in the 'outer halo'. We also find a break in the properties of the halo density distribution at ~ 20 kpc. This can be attributed to a break in the radial power law, as in a number of previous analyses; but we find that it foremost reflects a break in the halo flattening: the data are well fit by a single power law in the spheroidal coordinate r_q, with a distinct change in $q(r)$ at a radius of ~ 20 kpc.

We illustrate the first point about the local slope $\frac{\mathrm{d}\ln\nu}{\mathrm{dlnr}}$ in Figure 13. As mentioned in the introduction, this slope plays a particular role in dynamical modelling, perhaps most easily seen in the case of axisymmetric Jeans equation modelling. Figure 13 compares the radial profile functions between the models with constant or variable flattening: all models have

consistent radial profile slopes, $\frac{\mathrm{dln}\nu(\mathrm{R},\mathrm{R}/\sqrt{2})}{\mathrm{dlnR}}$ within 65 kpc. There is remarkable qualitative agreement, in the sense that these slopes vary consistently with radius, within the constraints of the functional form imposed. In all cases does the halo profile steepen between 10 kpc and 65 kpc. Clearly, the fits with fixed q differ more from one another and from the $q(r)$ fits, as they must through their more stringent functional form constraints. Allowing for a yet steeper drop beyond 65 kpc with a TPL profile, leads to indistinguishable results. The largest discrepancies are between power-laws and Einasto profiles beyond 65 kpc; we attribute that to the Einasto profile being constrained by the abundance of data at $r_{gc} \leq 65$ kpc (see bottom panel of Fig. 13), which inevitably leads to $\frac{\mathrm{dln}\nu}{\mathrm{dlnr}}$ at large radii. A look at Table 1 however shows that all of these models make the data appear comparably likely; the likelihood difference are significant, but not drastic (see also Fig. 7). In particular, the Einasto and broken power-law profiles lead to near-identical data likelihoods; the model fits with the flattening forced to be constant make the data significantly less likely. Remarkably, the single power-law with variable flattening has nearly the same $\frac{\mathrm{d\,ln}\,\nu}{\mathrm{dlnr}}$ along the intermediate axis, where the seeming radial change in power-law index is simply a reflection of the change in the radial coordinate r_q, attributable to $q(r)$.

We also compare these findings to other work, in particular that of Deason et al. (2011, 2014). Figure 14 shows that the halo profile, as traced in this analysis via K-giants, is consistent with the findings of Deason et al. (2011) within 65 kpc. To follow up the findings of Deason et al. (2014), a steep drop in the density of BHB stars beyond 65 kpc, we specifically fit a triple power-law with parameters (α_1, α_2, α_3, q). For comparison, we held the break radii fixed at $\mathrm{r_{break1}} = 18$ kpc and $\mathrm{r_{break2}} = 65$ kpc. Yet, as Figure 14 reveals the halo slopes in BHB stars and found here in K giants are formally inconsistent beyond ~ 65 kpc. Yet, there are only 7 stars beyond 65 kpc, so the paucity of distant K giants precludes a more stringent comparison. Besides, another two avenues are conceivable: first, we know that the substructure differs between BHB stars and giants and MS turn-off stars (Bell et al. 2008; Janesh et al. 2015; Xue et al. 2011). Second and related, it is conceivable than many of the stars at large radii in the present K-giant sample could be Sgr stream members, even though we made an attempt to remove such stars from the sample.

In summary, we believe that the present study has brought forward a number of new aspects: first, we have worked out a forward modelling of the spectroscopic data that has not been applied previously in this context. We believe that, in particular for tracers that are not standard candles like BHB stars, such an approach is warranted and powerful. Second, we were able to show on this basis 1) that there is an outward metallicity gradient in the halo beyond 10 kpc; 2) that the halo is distinctly flatter between 10-20 kpc, compared to larger radii. This distinct change in flattening suggests it more appropriate to think of the break in halo profile at 20 kpc as a break in flattening, rather than as a break in the radial profile at

forced constant flattening; 3) that there is overall consistency with previous analyses when it comes to the dynamically important quantity $\frac{\mathrm{d}\ln\nu}{\mathrm{dlnr}}$ in the radial range 10-65 kpc.

The research has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP 7) ERC Grant Agreement No. [321035].

X.-X.X. acknowledges the Alexander von Humboldt foundation for a fellowship that enabled this work and the support from NSFC grant 11103031, 11233004, 11390371 and 11003017.

HWR acknowledges funding by the Sonderforschungsbereich SFB 881 The Milky Way System (subproject A3) of the German Research Foundation (DFG).

J.B. acknowledges support from a John N. Bahcall Fellowship and the W.M. Keck Foundation.

HLM acknowledges support from NSf grant AST-1009886.

We thank Glenn van de Ven and Ling Zhu for helpful discussions, and Vasily Belokurov and Alis Deason for valuable input on an earlier version of the manuscript.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

REFERENCES

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21

Bell, E. F., Xue, X. X., Rix, H., Ruhland, C., & Hogg, D. W. 2010, AJ, 140, 1850

Bell, E. F., Zucker, D. B., Belokurov, V., et al. 2008, ApJ, 680, 295

Belokurov, V., Koposov, S. E., Evans, N. W., et al. 2014, MNRAS, 437, 116

Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJ, 642, L137

Bovy, J., Rix, H.-W., Liu, C., et al. 2012, ApJ, 753, 148

Carollo, D., Beers, T. C., Chiba, M., et al. 2010, ApJ, 712, 692

Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Nature, 450, 1020

Deason, A. J., Belokurov, V., & Evans, N. W. 2011, MNRAS, 416, 2903

Deason, A. J., Belokurov, V., Koposov, S. E., & Rockosi, C. M. 2014, ApJ, 787, 30

Einasto, J. 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72

Fernández-Alvar, E., Allende Prieto, C., Schlesinger, K. J., et al. 2015, A&A, 577, A81

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

Gould, A., Flynn, C., & Bahcall, J. N. 1998, ApJ, 503, 798

Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, AJ, 116, 3040

Harris, W. E. 1976, AJ, 81, 1095

Hawkins, M. R. S. 1984, MNRAS, 206, 433

Janesh, W., Morrison, H. L., Ma, Z., et al. 2015, ArXiv e-prints

Jeans, J. H. 1915, MNRAS, 76, 70

Lee, Y. S., Beers, T. C., Allende Prieto, C., et al. 2011, AJ, 141, 90

Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008a, AJ, 136, 2022

Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008b, AJ, 136, 2050

Lenz, D. D., Newberg, J., Rosner, R., Richards, G. T., & Stoughton, C. 1998, ApJS, 119, 121

Merritt, D., Graham, A. W., Moore, B., Diemand, J., & Terzić, B. 2006, AJ, 132, 2685

Morrison, H. L., Helmi, A., Sun, J., et al. 2009, ApJ, 694, 130

Morrison, H. L., Mateo, M., Olszewski, E. W., et al. 2000, AJ, 119, 2254

Pier, J. R., Munn, J. A., Hindsley, R. B., et al. 2003, AJ, 125, 1559

Preston, G. W., Shectman, S. A., & Beers, T. C. 1991, ApJ, 375, 121

Rix, H.-W. & Bovy, J. 2013, A&A Rev., 21, 61

Robin, A. C., Reylé, C., & Crézé, M. 2000, A&A, 359, 103

Saha, A. 1985, ApJ, 289, 310

Schönrich, R., Asplund, M., & Casagrande, L. 2011, MNRAS, 415, 3807

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Sesar, B., Ivezić, Ž., Grammer, S. H., et al. 2010, ApJ, 708, 717

Sesar, B., Jurić, M., & Ivezić, Ž. 2011, ApJ, 731, 4

Siegel, M. H., Majewski, S. R., Reid, I. N., & Thompson, I. B. 2002, ApJ, 578, 151

Sommer-Larsen, J. 1987, MNRAS, 227, 21P

Soubiran, C. 1993, A&A, 274, 181

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485

Vivas, A. K. & Zinn, R. 2006, AJ, 132, 714

Wetterer, C. J. & McGraw, J. T. 1996, AJ, 112, 1046

Xue, X.-X., Ma, Z., Rix, H.-W., et al. 2014, ApJ, 784, 170

Xue, X.-X., Rix, H.-W., Yanny, B., et al. 2011, ApJ, 738, 79

Xue, X. X., Rix, H. W., Zhao, G., et al. 2008, ApJ, 684, 1143 [X08]

Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

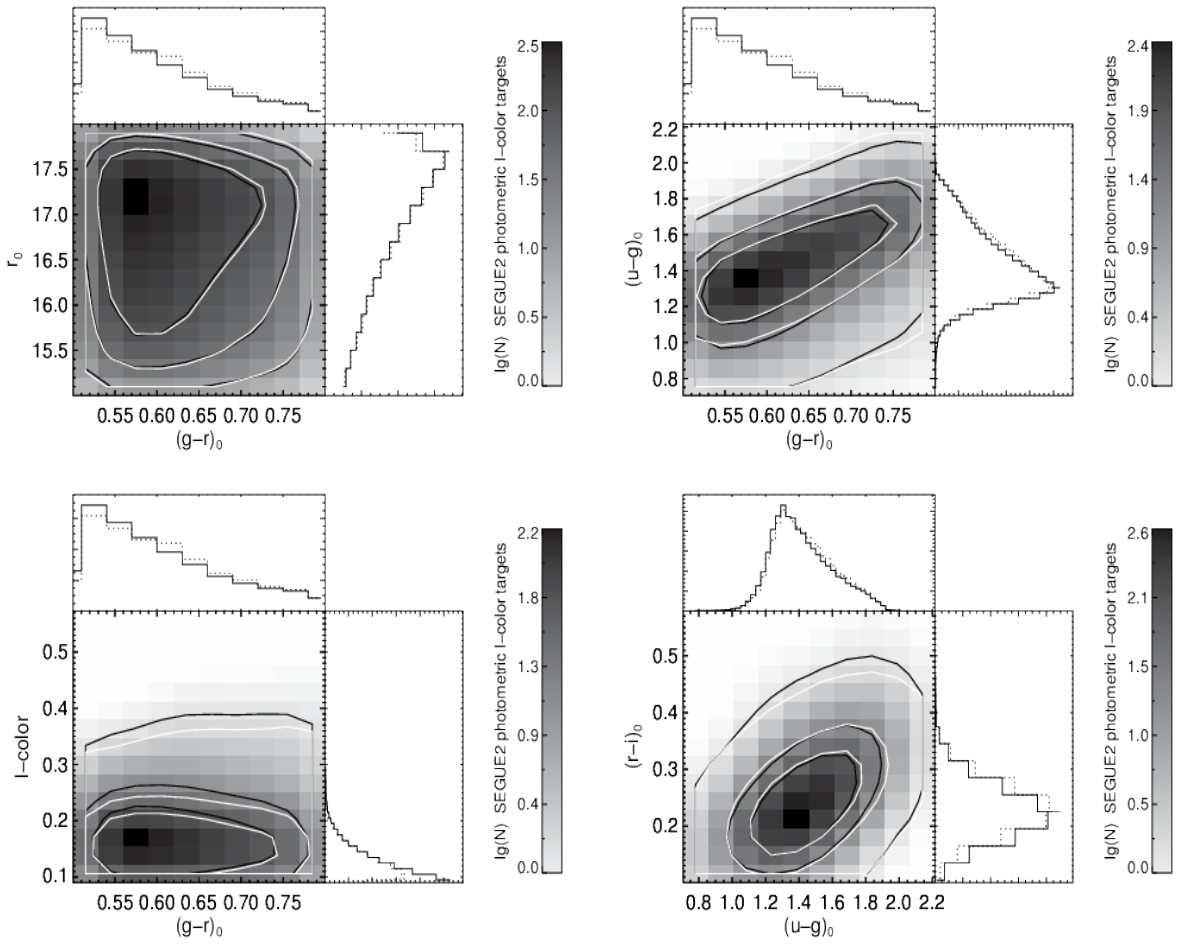

Fig. 1.—: Distribution of SEGUE-2 photometric l-color K-giant candidates in color-color and color-magnitude space (gray map, black contours, solid histogram) and the spectroscopically sample targetted and successfully analyzed (white contours, dotted histogram). The contours contain 68%, 95% and 99% of the distribution. The spectroscopic sample with parameters (including [Fe/H] and $\mathcal{DM}$) is a fair subset of the photometric targets with respect to colors and magnitudes. The color and magnitude limits enter the modelling through Eq. 13.

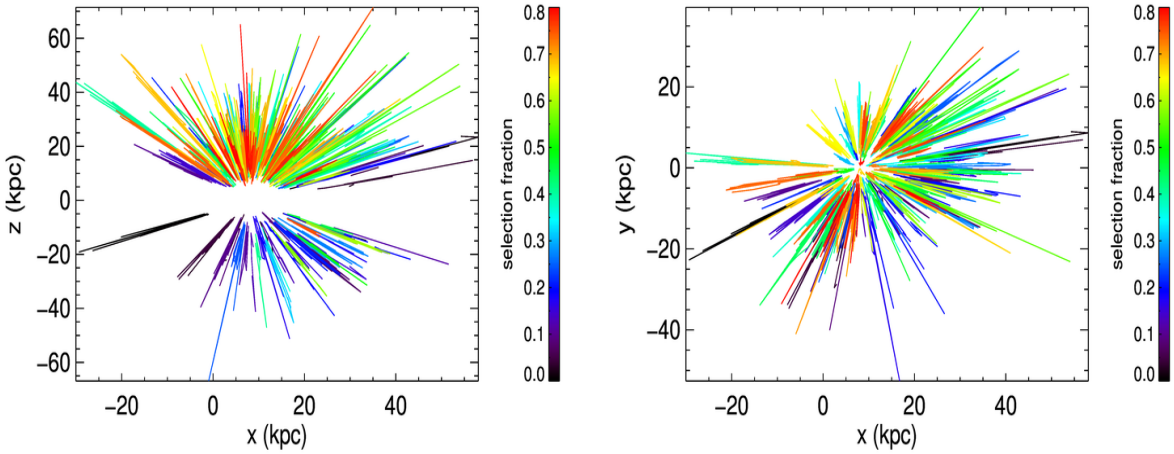

Fig. 2.—: The SEGUE-2 targetting fraction of photometrically selected of l-color K giants, as a function of Galactic coordinates x and y (left panel), and of Galactocentric coordinates x and vertical height z (right panel). Each line represent a spectroscopic plate, with the extent of the line the nearest and farthest object targetted. The color indicates the fraction of stars that photometrically pass the selection criteria that actually got targetted spectroscopically. At high galactic latitudes, most photometrically eligible targets had spectra taken; at low galactic latitude only a modest fraction, which is however known for each plate and accounted for in the modelling (Eq. 12).

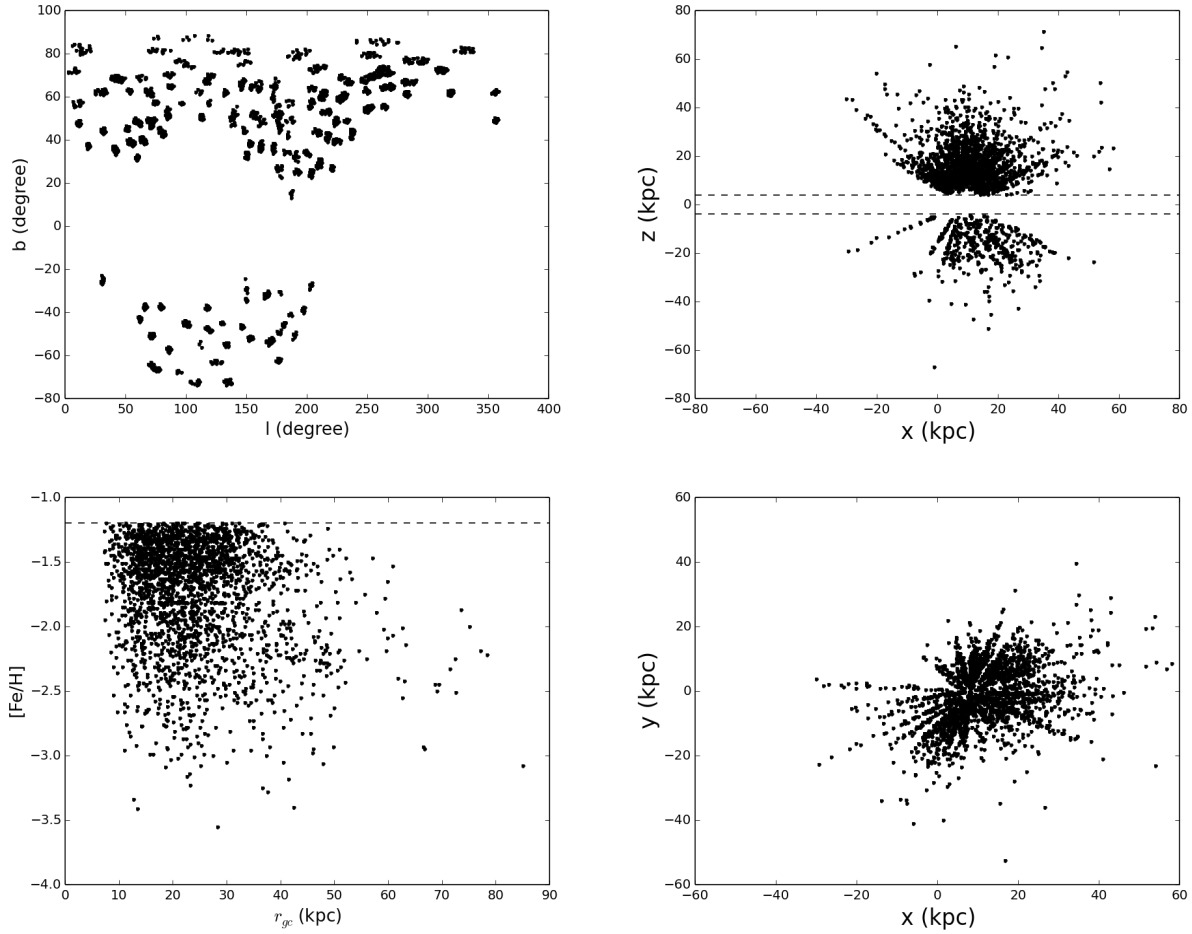

Fig. 3.—: (Upper left) The sky coverage and the spatial distributions (right panels) of SEGUE-2 l-color K giants in the sample; they appear as a pencil-beam pattern, due to the nature of SEGUE survey. (Lower left) The distribution of metallicities along with the Galactocentric radii shows that the mean metallicity is about -1.75 dex, and some K giants have metallicities of ~ -3.5. The stars with [Fe/H] > -1.2 and $|z| < 4$ kpc are culled because they could belong to the disk. The metallicity and distance cuts enter the modelling though Eqs. 9 and 10; the angular selection function through Eq. 12

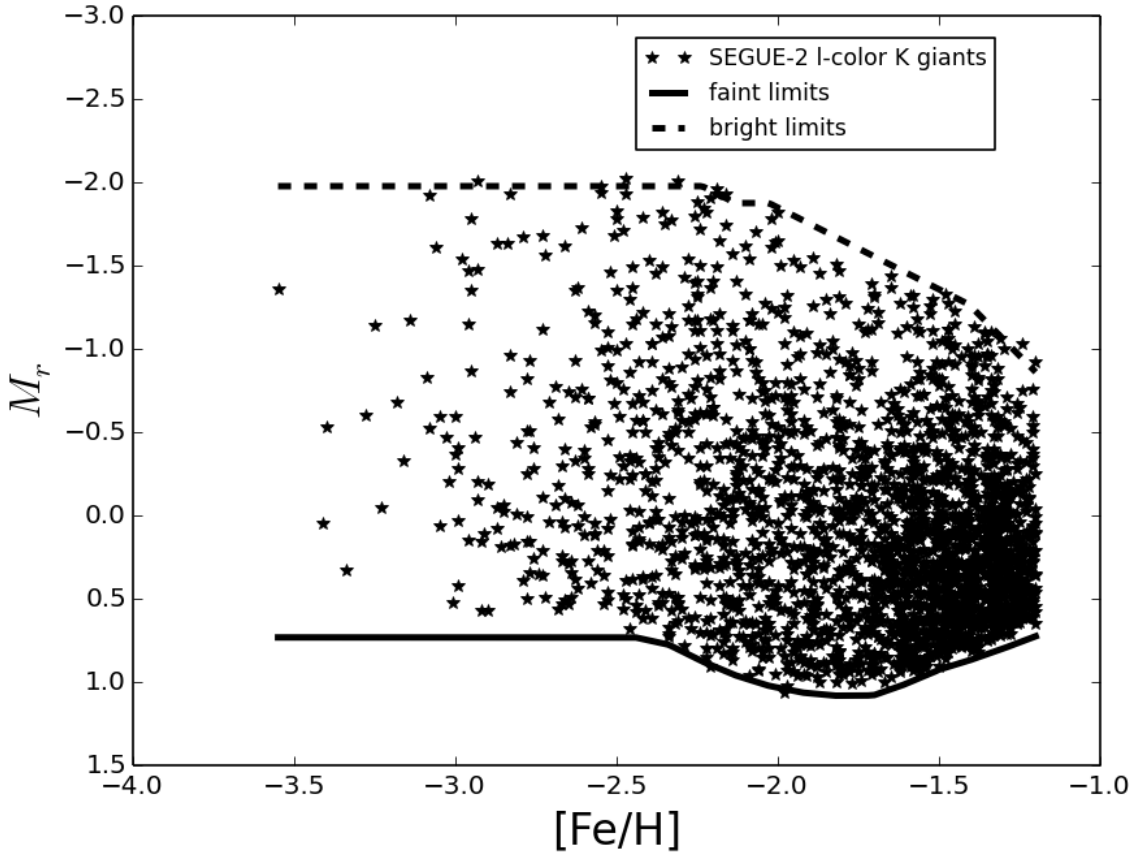

Fig. 4.—: The limits of the absolute magnitudes M_r changing with the metallicities [Fe/H]. The bright limits are caused by the metallicity-dependence of the fiducials' bright tips, while the faint limits are due to removal of possible red clump stars by Xue et al. (2014), and enter the modelling through Eq.11.

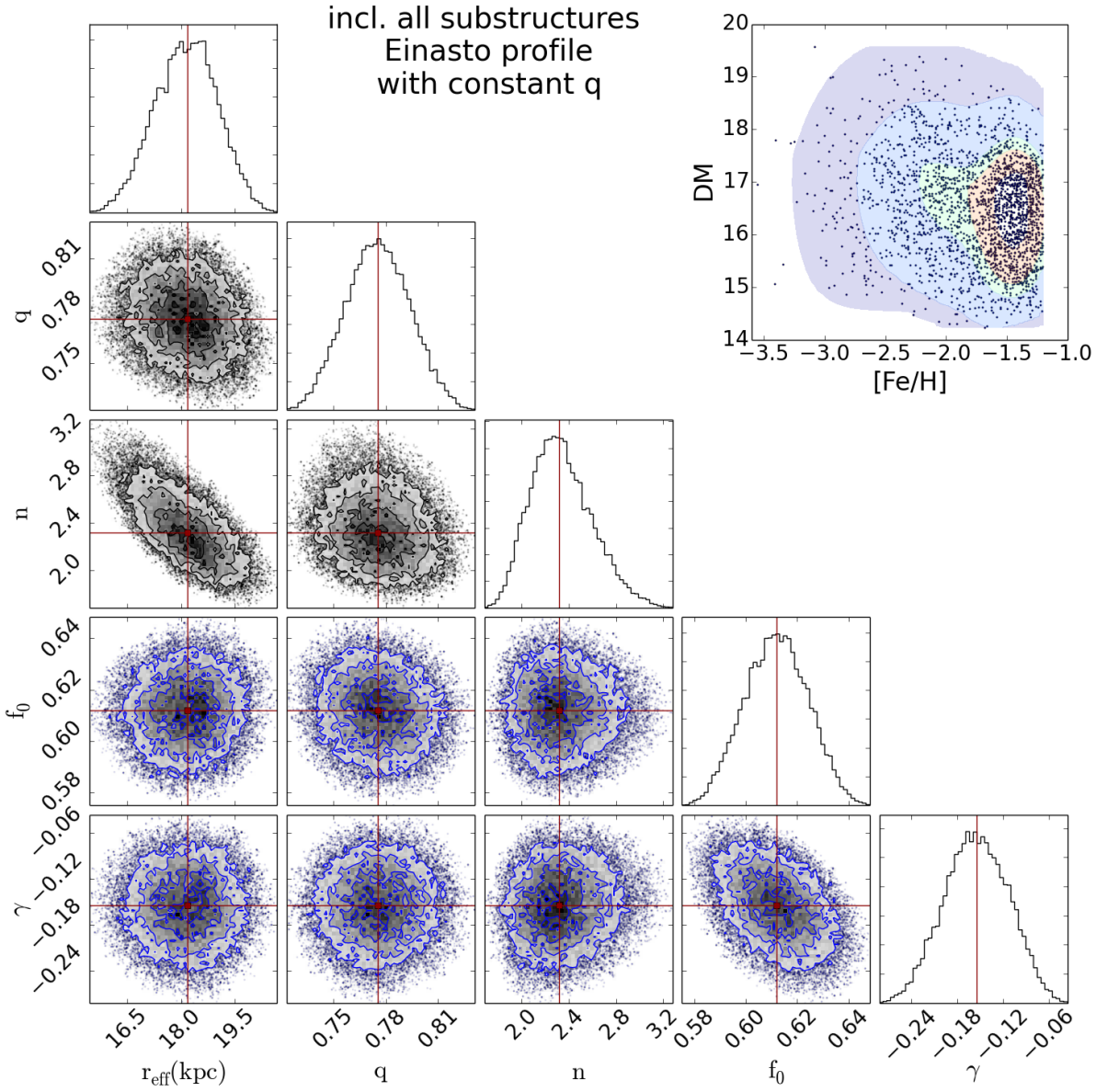

Fig. 5.—: All one and two dimensional projections of the posterior probability distributions for the parameters $(\mathrm{q}, \mathrm{n}, \mathrm{r_{eff}}, \mathrm{f_0}, \gamma)$ of the Einasto-profile and metallicity model, shown here for the case that *no* halo substructures are excised. The red lines and squares mark the median value of each parameter. The black sub-figures show the parameter *pdf* for the spatial density profiles, while the blue sub-figures show parameter *pdf*s for the metallicity distribution model. The top-right figure compares the data set (black dots) to the model prediction for $[\mathrm{Fe/H}] - \mathcal{DM}$ distribution , averaged over all directions and accounting for all selection effects. The colored model contours encompass 99.9%, 95%, 68%, 50%, and 16% of the normalized model probability, projected into the $[\mathrm{Fe/H}] - \mathcal{DM}$-plane. It matches the actual $[\mathrm{Fe/H}] - \mathcal{DM}$ distribution well. Of course, the model also makes a prediction for the (l, b) distribution of stars, but that prediction is dominated by the positions of the spectroscopic plates.

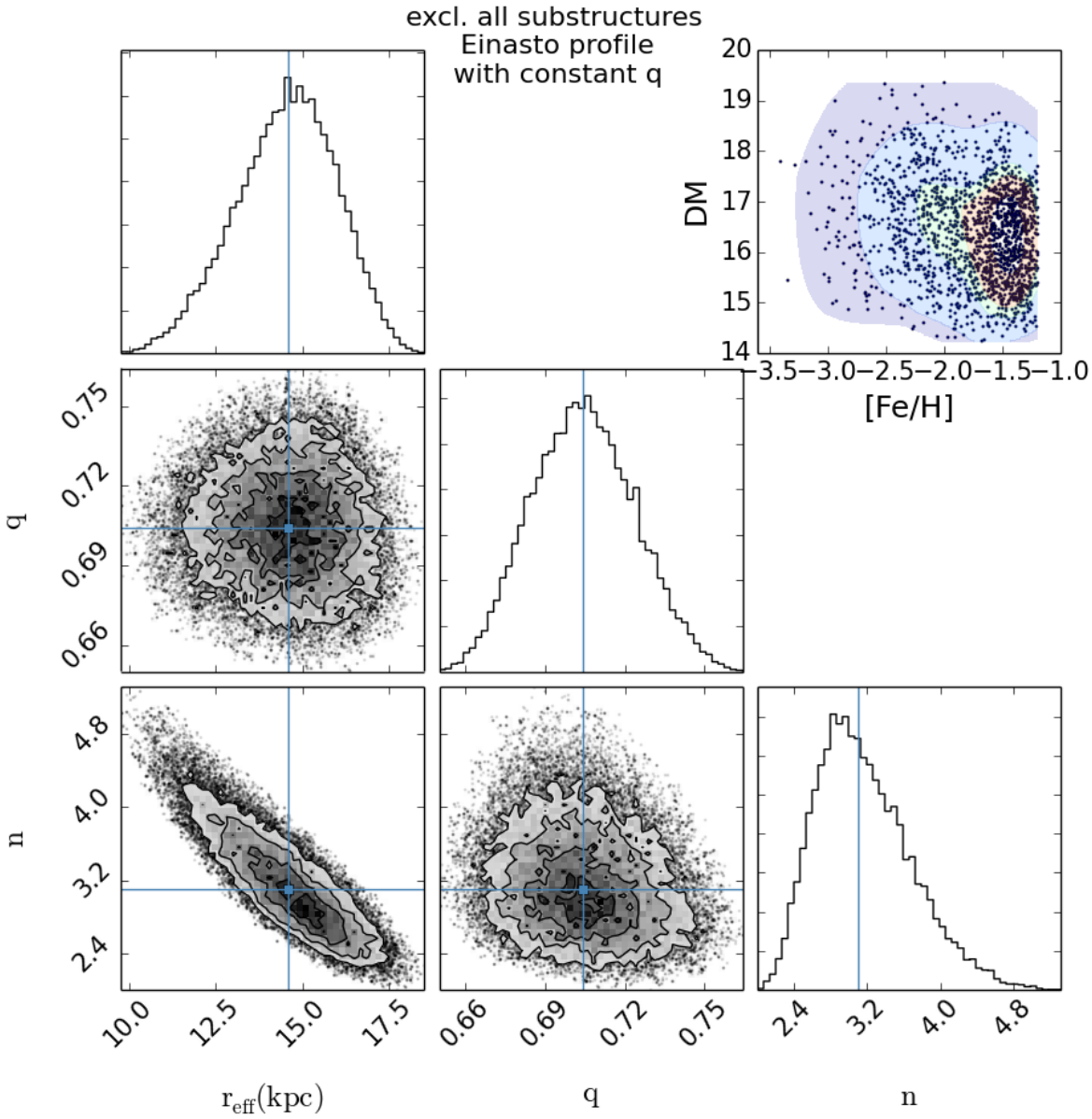

Fig. 6.—: The one and two dimensional projections of the posterior probability distributions of parameters (q, n, r_{eff}) of the Einasto-profile fit to the sample remaining after excluding stars identified by Janesh et al. (2015) as likely members of recognizable halo substructures. Analogous to Fig. 5, the basic data-model comparison is shown in the top-right panel. The comparison to Figure 5 shows that excluding substructures leads to a more concentrated and slightly more flattened Einasto profile; the profile's concentration parameter n is covariant with the effective radius parameter, as r_{eff} approaches the inner distance cut-off of the sample (10 kpc).

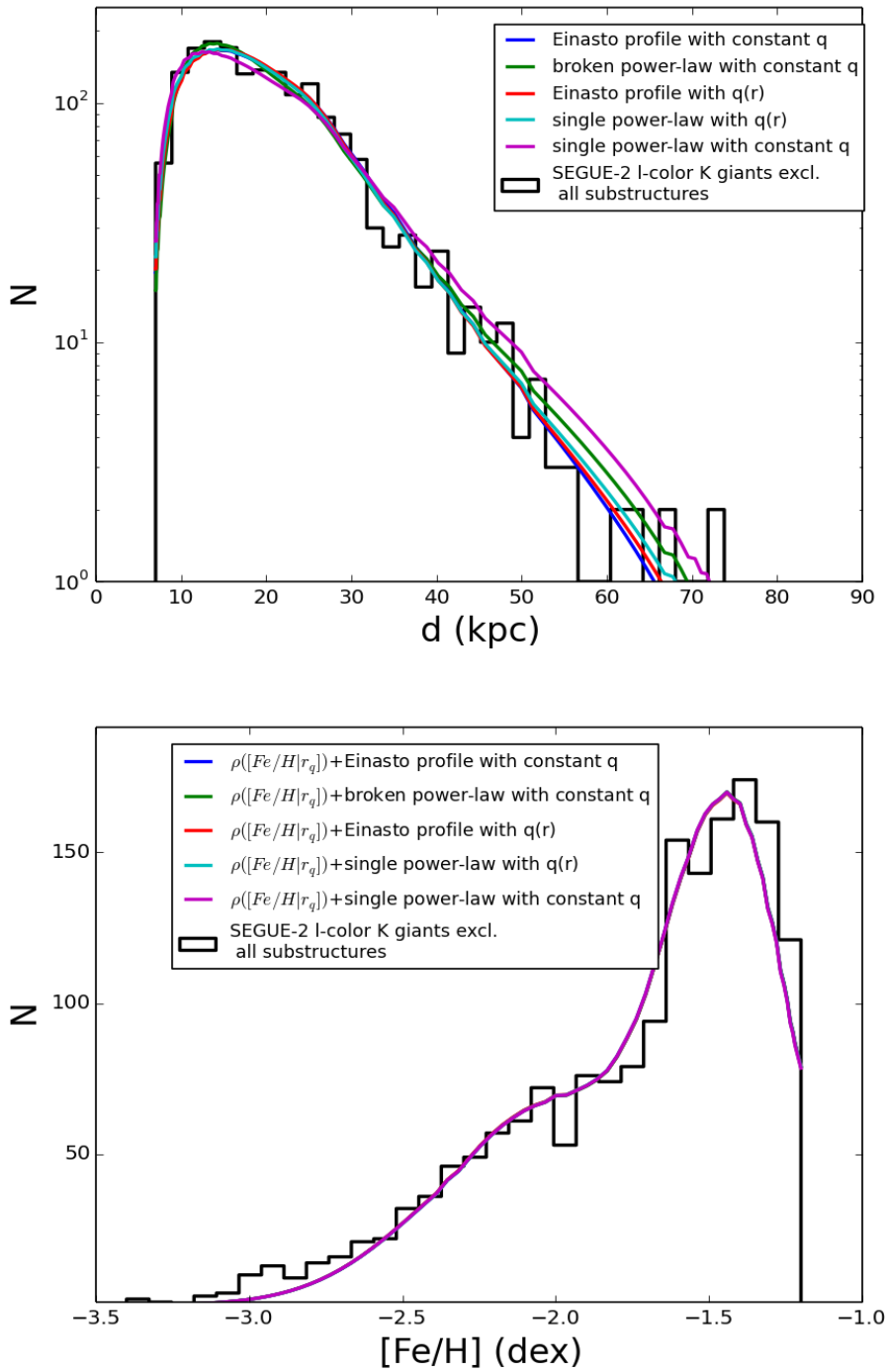

Fig. 7.—: Upper panel: Comparison between the observed distance distribution and the predicted distributions by the best-fitting models, marginalized over all other parameters (e.g. [Fe/H], l and b), and accounting for all selection effects. Lower panel: Analogous comparison between the observed metallicity distribution and the model predictions (now marginalized over $\mathcal{DM}$). All of the best-fitting models fit qualitatively well. Discrepancies in the [Fe/H] distribution reflect the fact that we only considered metallicity distribution that could be reflected in radially varying proportions of two Gaussian functions.

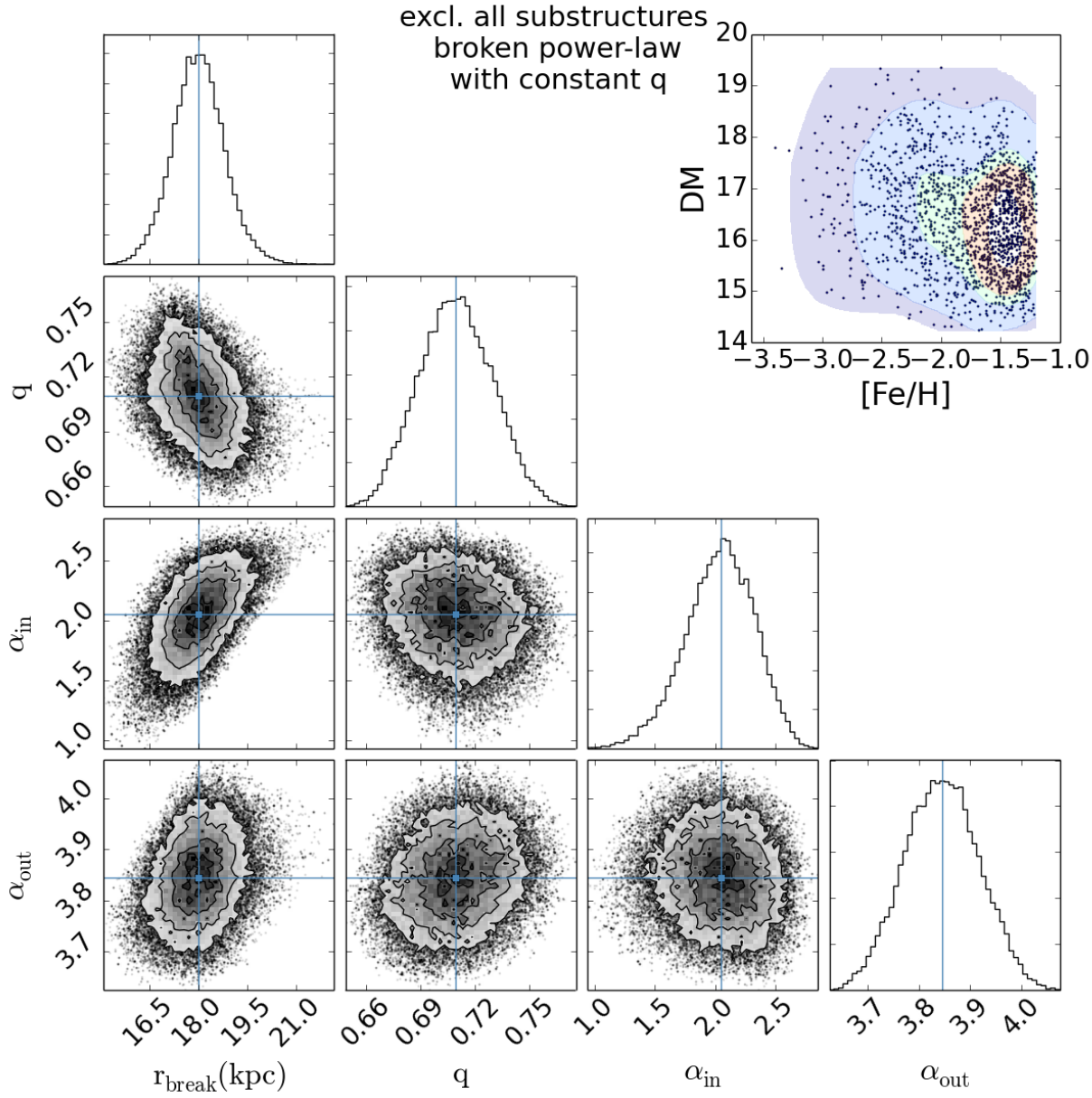

Fig. 8.—: Same as Figure 6, but for the broken power-law profile with parameters $(\mathrm{q}, \alpha_{\rm in}, \alpha_{\rm out}, \mathrm{r_{break}})$. The blue lines and squares mark the most likely value of each parameter. This broken power-law profile fits the data basically as well as the Einasto profile (see also Figure 7). Indeed, the actual prediction for the slope, $\partial \ln \nu_* / \partial \ln r$, in the radial regime constrained by the data is quite similar to that predicted by Einasto profile (Figure 13).

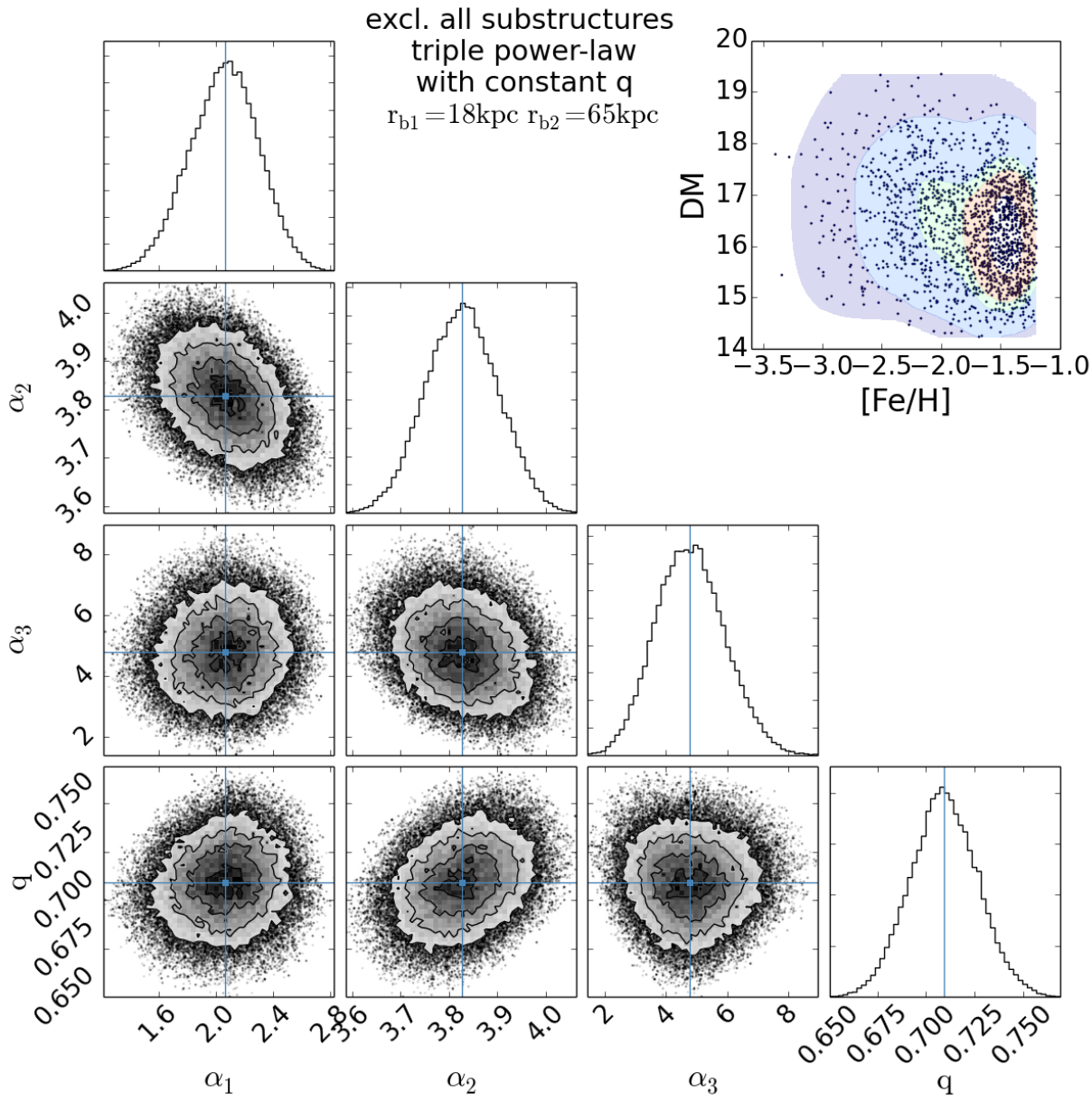

Fig. 9.—: Same as Figure 6, but for the triple (i.e. doubly-broken) power-law profile with parameters $(q, \alpha_1, \alpha_2, \alpha_3)$; the break radii are held fixed at 18 kpc and 65 kpc respectively (for comparison with Deason et al. (2014)). The blue lines and squares mark the most likely value of each parameter. The best-fitting TPL also fits the data well shown as top-right figure. The power-law index between 18 kpc and 65 kpc, α_2, is comparable to the index beyond 65 kpc, α_3. There is no strong drop beyond 65 kpc.

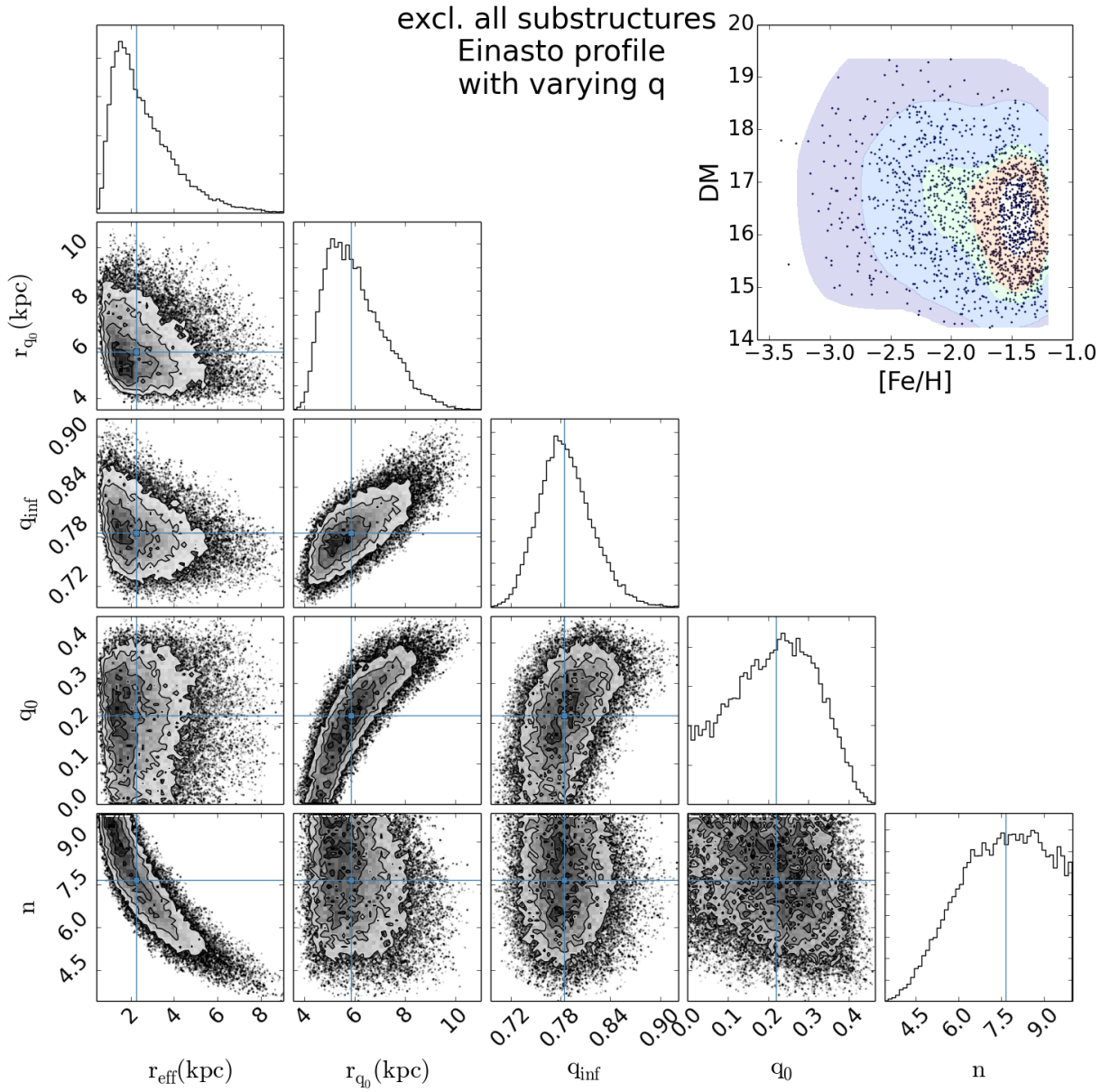

Fig. 10.—: Parameter *pdf*s for the Einasto profile with varying flattening (cf. §4.4). The fit implies a strong variation of flattening with radius, illustrated in Figure 12. Allowing for a varying flattening changes the parameters of Einasto profile quite strongly: The effective radius becomes formally very small and there is a strong covariance between $r_{\rm eff}$ and n. However, the actual prediction for the profile slope, $\partial \ln \nu_* / \partial \ln r$, is quite similar to that predicted by Einasto profile with a constant flattening shown as Figure 13.

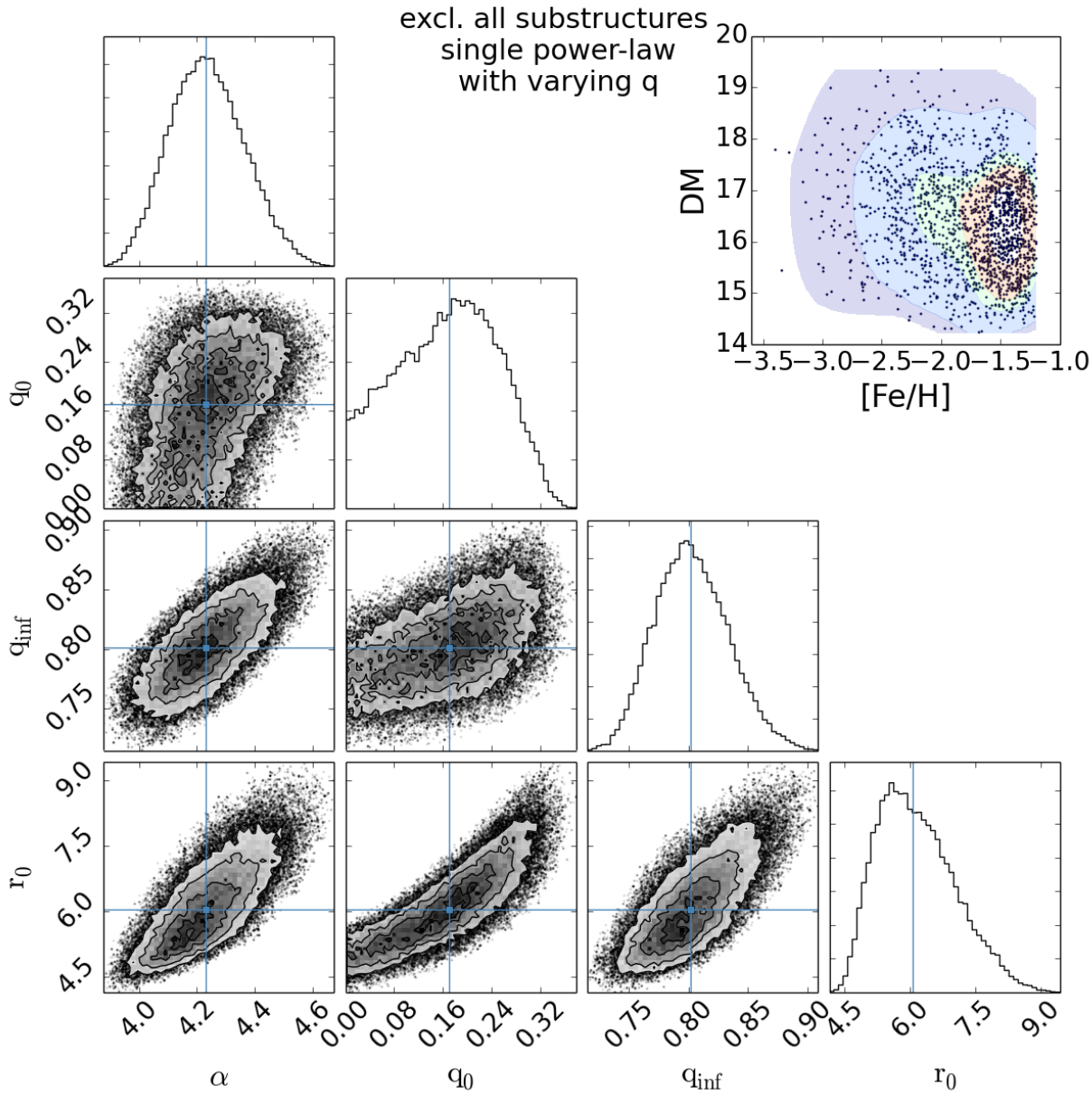

Fig. 11.—: Parameter *pdf*s for the single power-law profile with varying flattening (cf. §4.4). The blue lines and squares mark the median value of each parameter. The flattening profile is the same as for an Einasto fit (see also Figure 12). The actual predicted $\partial \ln \nu_* / \partial \ln r$ is similar to other models in the radial regime constrained by the data shown as Figure 13.

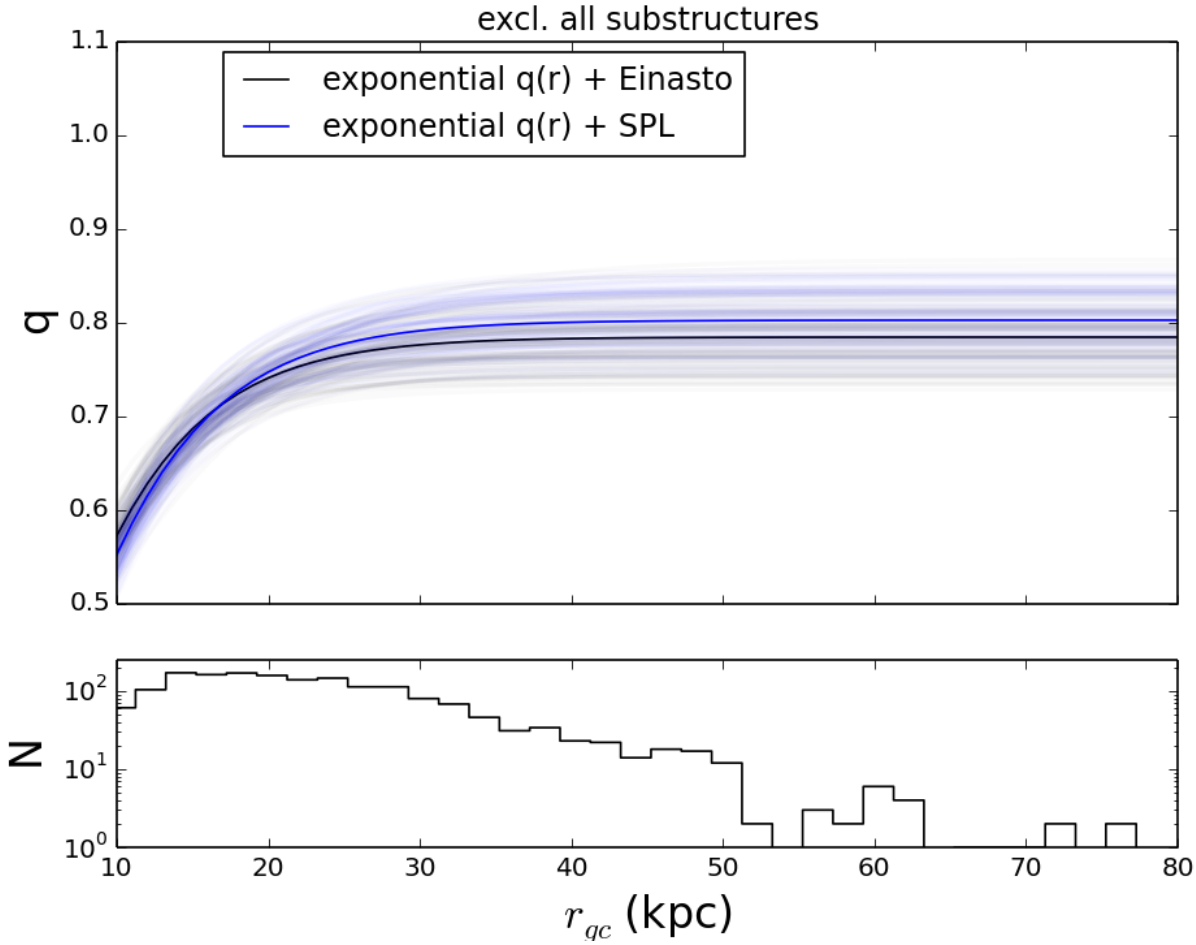

Fig. 12.—: The radial variation of the stellar halo's flattening (cf. §4.4). The coincidence of the curves for an Einasto and single power-law radial profile illustrates that the flattening profile is independent on the radial functional form. Over the observed range, the halo becomes much rounder at large radii, from q∼0.57 at 10 kpc to q∼0.78 at large radii. The bottom panel shows the actual radial distribution of the sample members used to constrain the fit.

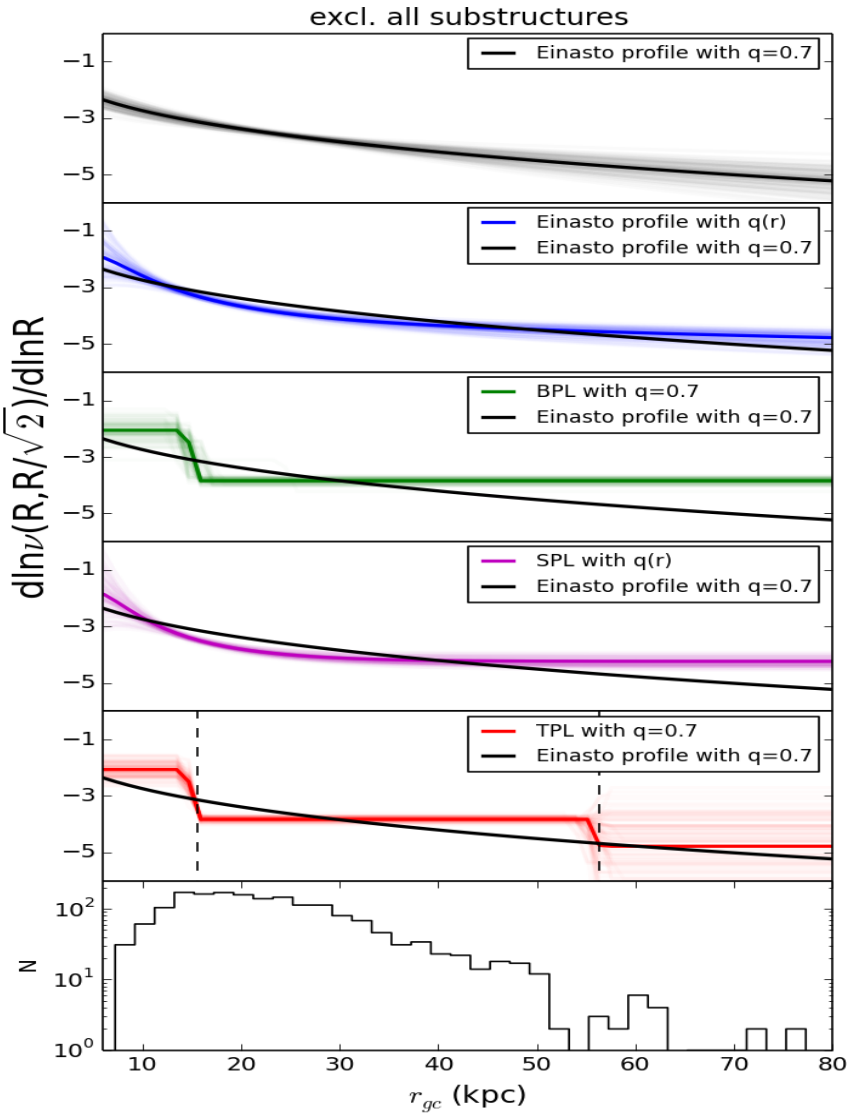

Fig. 13.—: (Modest) differences between the differing parameterizations of the radial profiles: the top five panels show the slope of the stellar density profile $\frac{\mathrm{dln}\nu(\mathrm{R},\mathrm{R}/\sqrt{2})}{\mathrm{dlnR}}$ (along an intermediate axis), for the Einasto profiles with constant (black) and variable (blue) flattening and the broken power-law (BPL) profiles with constant (green) flattening, and the single power-law with variable (magenta) flattening, and the twice broken power-law (TPL) (red) with constant flattening. The lines corresponding to radial profiles created from 100 samples of the parameter's *pdf*. The most likely profile from the top panel is repeated below for reference. The two vertical dashed lines indicate the two break radii (18 kpc and 65 kpc) of TPL in our fitting. Note that the break radius is in spheroidal coordinates r_q, while the X-axis here is spherical radius r_{GC}. Despite the fact that the Einasto profiles in the two cases (black and blue) have effective radii that differ by a factor of two, their radial profiles are very similar within the range constrained by the data. Especially, the Einasto profile and broken power-law with variable flattening (blue and magenta) show great consistency. Again, the bottom panel shows the distance distribution of the data. Most of the data are within 65 kpc, where all models have very similar predictions for the slope, $\partial \ln \nu(R, R/\sqrt{(2)})/\partial \ln R$.

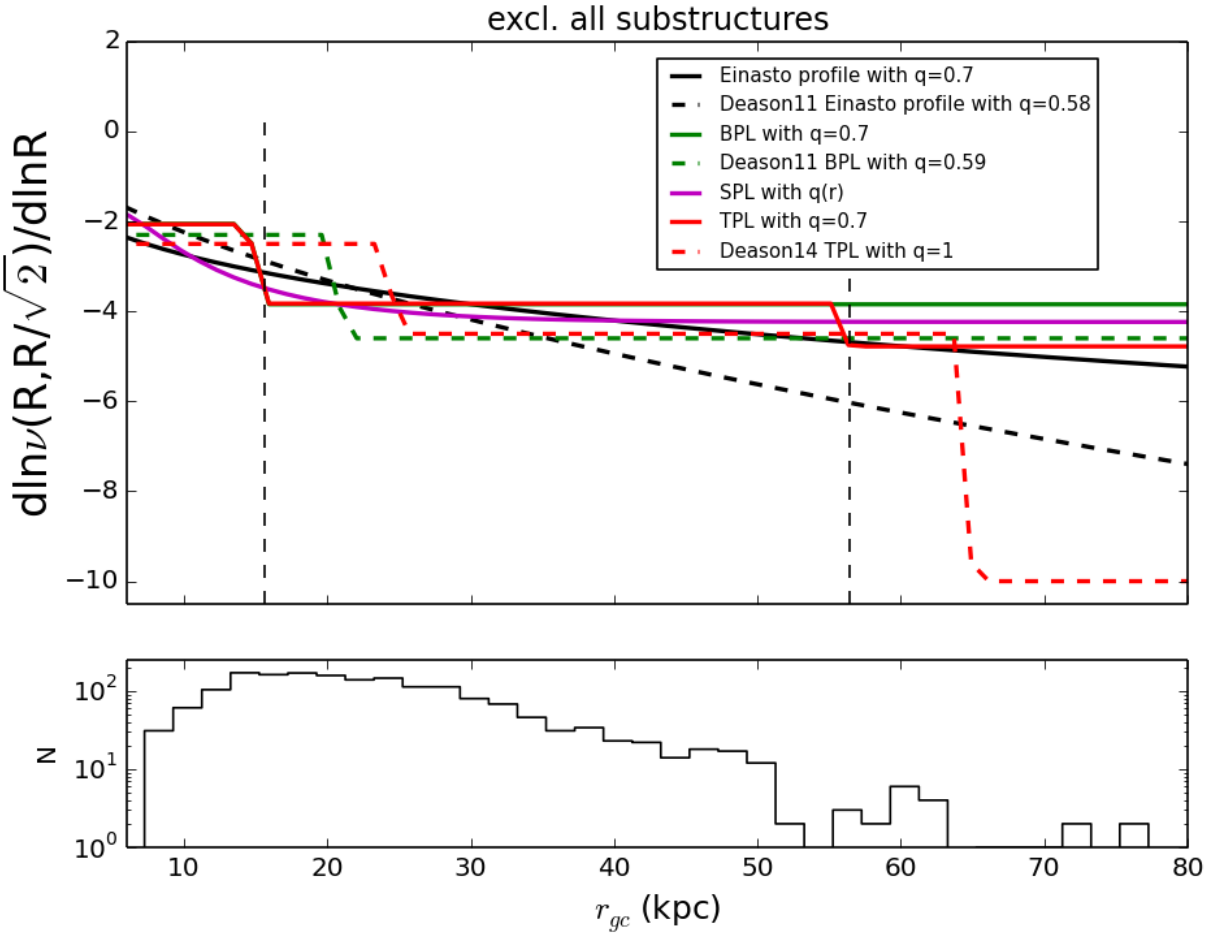

Fig. 14.—: Comparison of the radial density slope $\frac{\mathrm{dln}\nu(\mathrm{R},\mathrm{R}/\sqrt{2})}{\mathrm{dlnR}}$ found in the present analysis, based on K-giants, and the slopes found by Deason et al. (2011, 2014) from BHB stars; the differently colored lines show our best-fit broken power-law with constant flattening (green) and best-fit twice broken power-law with constant flattening (red), best-fit Einasto profiles with constant flattening (black) and the best-fit single power-law with variable flattening (magenta). The best-fit models of Deason et al. (2011, 2014) are shown on black, green and red but in dashed lines. The two vertical dashed lines indicate the two break radii (18 kpc and 65 kpc) of TPL in our fitting. Note that the break radius of broken power-law is in $\mathrm{r_q}$, but here the x-axis is in $\mathrm{r_{gc}}$. The lower panel shows the distance distribution of the data. Among the broken power-laws there are some differences between the present fit and the Deason et al. (2011, 2014) results, which used a different population as tracers. Notably, the K-giant sample does not point towards a steep halo drop beyond 65 kpc, although our substructure-cleaned sample contains only a few stars at such large radii.

Table 1:: A summary of our best-fitting models

Model	Parameters	N_p	$\Delta ln(\mathscr{L})$
Excl. all substructures			
[Fe/H]-model	$f_0 = 0.6 \pm 0.01$, $\gamma = -0.2 \pm 0.05$	2	
+			
SPL	$\alpha = 3.6 \pm 0.1$, q = 0.68 ± 0.02	2	34
Einasto	n = 3.1 ± 0.5, $r_{eff} = 15 \pm 2$ kpc, q = 0.7 ± 0.02	3	16
BPL	$\alpha_{in} = 2.1 \pm 0.3$, $\alpha_{out} = 3.8 \pm 0.1$, $r_{break} = 18 \pm 1$kpc,$q = 0.7 \pm 0.02$	4	14
TPL	$\alpha_1 = 2.1 \pm 0.2$, $\alpha_2 = 3.8 \pm 0.1$, $\alpha_3 = 4.8 \pm 0.8$, q = 0.7 ± 0.02, $\mathbf{r_{break1} = 18kpc}$, $\mathbf{r_{break2} = 65kpc}$	3	13
[Fe/H]-model	$f_0 = 0.6 \pm 0.01$, $\gamma = -0.26 \pm 0.06$	2	
+			
Einasto-q(r)	n = 7.7 ± 1.5, $r_{eff} = 2 \pm 1$ kpc, $q_0 = 0.2 \pm 0.1$, $q_{inf} = 0.78 \pm 0.05$, $r_0 = 6 \pm 1$ kpc	5	1
BPL-q(r)	$\alpha_{in} = 4.2 \pm 0.4$, $\alpha_{out} = 4.3 \pm 0.3$, $r_{break} = 22 \pm 23$ kpc ,$q_0 = 0.2 \pm 0.1$, $q_{inf} = 0.8 \pm 0.03$, $r_0 = 6 \pm 1$ kpc	6	0
SPL-q(r)	$\alpha = 4.2 \pm 0.1$,$q_0 = 0.2 \pm 0.1$, $q_{inf} = 0.8 \pm 0.03$, $r_0 = 6 \pm 1$ kpc	4	0 (–12760)
Incl. all substructures			
[Fe/H]-model	$f_0 = 0.6 \pm 0.01$, $\gamma = -0.15 \pm 0.04$	2	
+			
SPL	$\alpha = 3.4 \pm 0.1$, q = 0.74 ± 0.02	2	61
Einasto	n = 2.3 ± 0.2, $r_{eff} = 18 \pm 1$ kpc, q = 0.77 ± 0.02	3	24
BPL	$\alpha_{in} = 2.8 \pm 0.1$, $\alpha_{out} = 4.3 \pm 0.1$, $r_{break} = 29 \pm 2$kpc,$q = 0.77 \pm 0.02$	4	26
TPL	$\alpha_1 = 2.8 \pm 0.1$, $\alpha_2 = 4.3 \pm 0.2$, $\alpha_3 = 4.2 \pm 0.4$, q = 0.77 ± 0.02, $\mathbf{r_{break1} = 30kpc}$, $\mathbf{r_{break2} = 50kpc}$	3	26
[Fe/H]-model	$f_0 = 0.62 \pm 0.01$, $\gamma = -0.22 \pm 0.06$	2	
+			
Einasto-q(r)	n = 6.1 ± 1.7, $r_{eff} = 3 \pm 2$ kpc, $q_0 = 0.3 \pm 0.05$, $q_{inf} = 0.9 \pm 0.04$, $r_0 = 9 \pm 2$ kpc	5	0
BPL-q(r)	$\alpha_{in} = 4.2 \pm 0.3$, $\alpha_{out} = 4.5 \pm 0.4$, $r_{break} = 32 \pm 18$ kpc ,$q_0 = 0.3 \pm 0.1$, $q_{inf} = 0.9 \pm 0.04$, $r_0 = 9 \pm 1$ kpc	6	0
SPL-q(r)	$\alpha = 4.4 \pm 0.1$,$q_0 = 0.3 \pm 0.1$, $q_{inf} = 0.9 \pm 0.04$, $r_0 = 9 \pm 1$ kpc	4	0 (–17679)

We give the type of model, the best-fitting parameters of the model, the number of free parameters, the log-likelihood for the best-fitting model. Parameters which are kept fixed are highlighted in bold.