

# Classification of local stellar populations. The improved MEMPHIS algorithm - Part II

Rafael Cubarsi & Santiago Alcobé

Dept. Matemàtica Aplicada IV  
Universitat Politècnica de Catalunya  
08034 Barcelona, Catalonia, Spain  
E-mail:rcubarsi@ma4.upc.edu

July 2009

## Abstract

Discontinuities of the local velocity distribution which are associated with stellar populations are studied from the improved statistical method MEMPHIS (Maximum Entropy of the Mixture Probability from Hierarchical Segregation), by combining a sampling parameter, optimisation of the mixture approach, and maximum partition entropy of populations composing the stellar sample. The sampling parameter is associated with isolating integrals of the star motion and it is used to build a hierarchical family of subsamples. An accurate characterisation of the entropy graph is given where a local maximum of entropy takes place simultaneously with a local minimum  $\chi^2$  error. By working from different sampling parameters the method is applied to samples from HIPPARCOS and Geneva-Copenhagen survey (GCS) to obtain kinematic parameters and mixture proportions of thin disk, thick disk and halo. The sampling parameter  $P = |(U, V, W)|$ , absolute heliocentric velocity, allows to build an optimal subsample containing thin and thick disk stars, by leaving aside most of the halo population. The sampling parameter  $P = |W|$ , absolute perpendicular velocity, is able to build an optimal subsample containing a mixture of total disk and halo stars, although it does not allow an optimal segregation of thin and thick disks. Other sampling parameters like  $P = |(U, W)|$  or  $P = |V|$  are found to be less population informative. By comparing both samples, HIPPARCOS provides more accurate estimates for thick disk and halo, while GCS does for the total disk. In particular, the radial velocity dispersion of the halo fits perfectly into the empirical Titius-Bode like law  $\sigma_U = 6.6 \left(\frac{4}{3}\right)^{3n+2}$ , which was previously proposed for discrete kinematic components, where the values  $n = 0, 1, 2, 3$  stands for early-type stars, thin disk, thick disk, and halo populations. Population statistics are used to segregate thin disk, thick disk, and halo, and to obtain a more accurate bayesian estimation of the population fractions.

KEY WORDS: stars: kinematics – galaxies: kinematics and dynamics – galaxies: statistics – methods: statistical.

1991 MATHEMATICS SUBJECT CLASSIFICATION: 60, 62, 85.

# 1 Local populations

The application of the improved MEMPHIS algorithm (Cubarsi & Alcobé 2009) to actual samples leads to more accurate population estimates than in previous works. For the HIPPARCOS sample (ESA 1992, 1997), with the sampling parameter  $P = |(U, V, W)|$ , the absolute heliocentric velocity, the optimal sample is obtained for  $P = 230 \text{ km s}^{-1}$ , and contains nearly all the local disk, while beyond  $370 \text{ km s}^{-1}$  the samples include a significant part of the halo. Let us mention that, according to Paper I (Alcobé & Cubarsi 2005) and Paper II (Cubarsi & Alcobé 2004), the thick disk began to be included in the nested subsamples from an absolute heliocentric velocity of about  $85 \text{ km s}^{-1}$ , although the core of its distribution was entering around a velocity of  $125 \text{ km s}^{-1}$ . Although they may slightly vary depending on the sample, those values are consistent with those recently obtained by Vidojević & Ninković (2008, 2009) from a different stellar sample. According to MEMPHIS the optimal sample is composed of about 92% of thin disk stars, with velocity dispersions  $(28.4 \pm 0.6, 15.6 \pm 2.5, 12.8 \pm 0.7) \text{ km s}^{-1}$  and vertex deviation  $10^\circ \pm 2$ , and of 8% of thick disk stars, with velocity dispersions  $(66.8 \pm 2.3, 39.4 \pm 10.6, 41.8 \pm 1.9)$  and nearly null vertex deviation,  $5^\circ \pm 3$ . The differential rotation mean velocity is  $54.6 \pm 3.4 \text{ km s}^{-1}$ . On the other hand, for the sampling parameter  $P = |W|$ , the absolute perpendicular velocity, a subsample containing stars with  $P \leq 65 \text{ km s}^{-1}$  includes nearly all the local disk population, while beyond  $120 \text{ km s}^{-1}$  the samples contain a significant part of the halo. The optimal sample is obtained for  $P = 180 \text{ km s}^{-1}$ , by giving a mixture of total disk and halo. It is composed of about 98% of disk stars, with velocity dispersions  $(32.7 \pm 1.1, 15.2 \pm 10.7, 16.4 \pm 0.4)$  and  $8^\circ \pm 5$  of vertex deviation, and of 2% of halo stars, with velocity dispersions  $(156.4 \pm 10.4, 88.5 \pm 53.2, 70.5 \pm 4.3)$  and vanishing vertex deviation,  $-4^\circ \pm 4$ . Notice that the shape of the halo velocity ellipsoid is radial dominated, as expected near the Sun (Battaglia et al. 2005, Thom et al. 2005). The differential rotation mean velocity is  $216.5 \pm 20.8 \text{ km s}^{-1}$ . These population fractions have an uncertainty of about 1%, but they will be improved later.

For the GCS sample (Nordström et al. 2004), with the sampling parameter  $P = |(U, V, W)|$ , the optimal sample for a thin and thick disks mixture is obtained also for  $P = 230 \text{ km s}^{-1}$  and contains most of disk stars, while beyond  $375 \text{ km s}^{-1}$  the samples include a significant part of the halo. It is composed of about 96% of thin disk stars, with velocity dispersions  $(30.6 \pm 0.4, 17.5 \pm 1.4, 15.2 \pm 0.3)$  and  $9^\circ \pm 1$  of vertex deviation, and of 4% of thick disk stars, with velocity dispersions  $(68.8 \pm 2.7, 30.0 \pm 15.0, 40.0 \pm 2.0)$  and vertex deviation  $6^\circ \pm 2$ . The differential rotation mean velocity is  $66.6 \pm 4.4 \text{ km s}^{-1}$ . For the sampling parameter  $P = |W|$ , a subsample with  $P = 60 \text{ km s}^{-1}$  contains nearly all the disk population, while beyond  $120 \text{ km s}^{-1}$  the samples contain a significant part of the halo. Notice that this values are very similar to those from HIPPARCOS. The optimal sample is obtained for  $P = 170 \text{ km s}^{-1}$ , which contains a mixture of total disk and halo populations. It is composed of 99% of disk stars, with velocity dispersions  $(32.5 \pm 0.7, 17.8 \pm 5.7, 16.5 \pm 0.3)$  and vertex deviation  $8^\circ \pm 4$ , and of 1% of halo stars, with velocity dispersions  $(151.0 \pm 14.7, 107.9 \pm 54.0, 70.3 \pm 5.7)$  and vanishing vertex deviation,  $-3^\circ \pm 7$ . The differential rotation mean velocity is  $213.1 \pm 24.3 \text{ km s}^{-1}$ .

By comparing those results, it seems that HIPPARCOS sample leads to a better estimation of thick disk and halo kinematic populations, although GCS has more accurate estimates for the total disk, and for the thin disk. However, the greater thin disk velocity dispersions of the GCS sample suggest a kinematic bias of the sample towards older disk stars, as it could be expected from the star selection of F and K spectral types.

Indeed, above kinematical parameters are estimates of the distribution moments of each stellar population. They may then be used to classify the stars into disjoint samples, thin disk, thick

disk and halo, and to compute their actual sample moments. To this purpose, the foregoing mixture proportions are taken as initial values for an iterative method, which is based in the bayesian probability of a star to belong to a particular population, by leading to a more accurate estimation of the population fractions. Although initial and final fractions are quite similar for HIPPARCOS' sample (final thin disk 91.5%, thick disk 7.3%, halo 1.2%), for the GCS sample (final thin disk 93.8%, thick disk 5.7%, halo 0.5%) a small difference is very significant, since the halo fraction is less than a half of HIPPARCOS'. The resulting sample moments from both catalogues are very similar. Thus, for HIPPARCOS sample the velocity dispersions are: thin disk ( $28.2 \pm 0.2$ ,  $16.9 \pm 0.2$ ,  $12.5 \pm 0.1$ ), thick disk ( $69.3 \pm 1.3$ ,  $37.9 \pm 0.9$ ,  $42.9 \pm 0.9$ ), and halo ( $179.6 \pm 7.8$ ,  $89.0 \pm 6.3$ ,  $90.8 \pm 6.0$ ). For the GCS sample they are: thin disk ( $29.7 \pm 0.2$ ,  $17.8 \pm 0.2$ ,  $13.9 \pm 0.1$ ) thick disk ( $65.7 \pm 1.5$ ,  $36.7 \pm 1.1$ ,  $41.5 \pm 0.9$ ), and halo ( $178.6 \pm 12.2$ ,  $113.7 \pm 15.0$ ,  $110.5 \pm 14.2$ ). These sample moments will be verified by comparing with those obtained from an independent kinematical procedure, by using the orbital elements of the stars, as described in section §5. They are found to be mostly consistent, although, for the thin disk, the mixture model tends to give smaller values of  $\sigma_v$ , which are compensated by higher values of the rotation differential velocity between thin and thick disks. Our mixture proportions and kinematic parameters try to improve the currently accepted ones for the main Galactic components (e.g. Fuchs et al. 2009, Reddy et al. 2006, Freeman & Bland-Hawthorn, 2002, Helmi 2008), in particular, those of the halo.

## 2 Application

The described procedure in Cubarsi & Alcobé (2009) will be first applied to the HIPPARCOS sample used in Paper I to show the improvement of the results. Furthermore, it will be also applied to the disk representative GCS sample, with a similar number of stars than HIPPARCOS sample, but with better estimations of the radial velocities.

### 2.1 HIPPARCOS sample

This sample was obtained by crossing HIPPARCOS Catalogue (ESA 1997) with radial velocities from the HIPPARCOS Input Catalogue (ESA 1992) (see details in Paper II). In order to obtain a typical sample of the galactic disk the sample was limited to a trigonometric distance of 300 pc. The resulting total sample was composed by 13,678 stars, where the only input data were the velocity components ( $U, V, W$ ) in a cartesian heliocentric coordinate system, with  $U$  toward the Galactic centre,  $V$  in the Galactic rotational direction, and  $W$  perpendicular to Galactic plane, positive in the direction of the North Galactic pole.

The application of MEMPHIS algorithm to the HIPPARCOS sample in Paper I led to a precise kinematic description of thin and thick disk populations. The absolute value of the total space motion either refereed to the sun, or to the centroid of the major population, were used as sampling parameters, although the qualitative results were similar. Within the thin disk two non-gaussian subpopulations, early-type and young-disk stars, were found with a strong asymmetry in the radial velocity. Those populations, altogether with the background old disk stars, featured the whole thin disk, with a nearly gaussian distribution. Since no previous selection was done in order to exclude non-typical disk stars, the total thin disk population was the result of an addition of several possible combined phenomena over its stars, like moving groups coming from other Galactic regions, cluster evaporation, disk heating, spiral bar perturbation, etc., so that its final normal distribution was a good example of what the central limit theorem predicts.

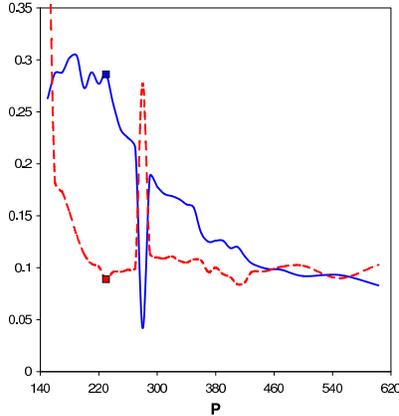


Figure 1: Entropy graph (continuous line) and  $\chi^2$  error of the mixture approach (dashed line, scaled to one tenth) for HIPPARCOS sample extended up to halo stars, for the sampling parameter  $P = |(U, V, W)|$ . The plateau region for optimal segregation of thin and thick disk stars, where both plots show opposite trends, takes place between  $P = 190$  and  $270 \text{ km s}^{-1}$  (absolute heliocentric velocity). There is a  $\chi^2$  local minimum at  $P = 230$ . For the disk and halo segregation, a slight opposite trend of the plots appears around  $P = 410$ , also with a  $\chi^2$  local minimum, although for  $P > 300$  there is no plateau region in the entropy graph.

### 2.1.1 Sampling parameters $P = |(U, V, W)|$ and $P = |W|$

By using the heliocentric velocity  $P = |(U, V, W)|$  as sampling parameter, from the value  $P = 145 \text{ km s}^{-1}$  onward both disk populations were well described according to a mixture of two gaussian populations, where the thin disk velocity dispersions became saturated around  $(28 \pm 1, 16 \pm 2, 13 \pm 1) \text{ km s}^{-1}$ . For those subsamples, as increasing the sampling parameter the new entering stars only contributed to increase the dispersion of the thick disk. An entropy maximum at  $P = 209$  gave an optimal mixture composed of about 9% of thick disk stars, with dispersions of  $(65 \pm 2, 39 \pm 9, 41 \pm 2)$ . However, we were unable to go farther to detect halo stars.

Now, with the improved algorithm, by combining both optimal criteria, the whole sample can be scanned up to  $P = 600 \text{ km s}^{-1}$ . We get simultaneous local extremes of  $\chi^2$  and entropy at  $P = 230 \text{ km}^{-1}$ , just before a deep entropy drop, and we also get a local minimum of  $\chi^2$  at  $P = 410$ , as it is shown in Fig. 3. Notice the opposite trend of  $\chi^2$  and entropy around those values, mainly around  $P = 230$ , in what can be considered a plateau region of the entropy graph. Nevertheless, after  $P = 300$  there is no plateau region, since the entropy is mostly decreasing. For  $P > 230$ , stars coming from the new halo population are merged to the increasing subsamples, where an entropy drop takes place around  $P = 280$ , simultaneously with a peak of  $\chi^2$ . Afterwards, the component Pop-I wraps the whole disk, while halo stars are confined in Pop-II. Therefore, the value  $P = 230$  now provides an improved optimal mixture estimation for thin and thick disks. The mixture proportion, dispersions, differential mean velocity, and vertex deviation are shown in Table 1. The value  $P = 410$  might give a first estimation for a total disk and halo mixture. The trend of the diagonal central velocity moments in terms of the sampling parameter is shown in Fig. 4 (in a logarithmic plot) where we see that the disk heating process remains nearly constant after  $P = 410$  (Pop-I, dashed line). The halo moments do not reach steady values.

If the alternative sampling parameter  $P = |W|$ , absolute value of the star velocity perpendicular to the Galactic plane, is used, we obtain the entropy and  $\chi^2$  plots displayed in Fig. 5. There is a small plateau region of the entropy graph, with a short local minimum of  $\chi^2$ , around  $P = 65 \text{ km s}^{-1}$ , although there is not a significative opposite trend of both graphs. Indeed, the  $\chi^2$  graph is mostly

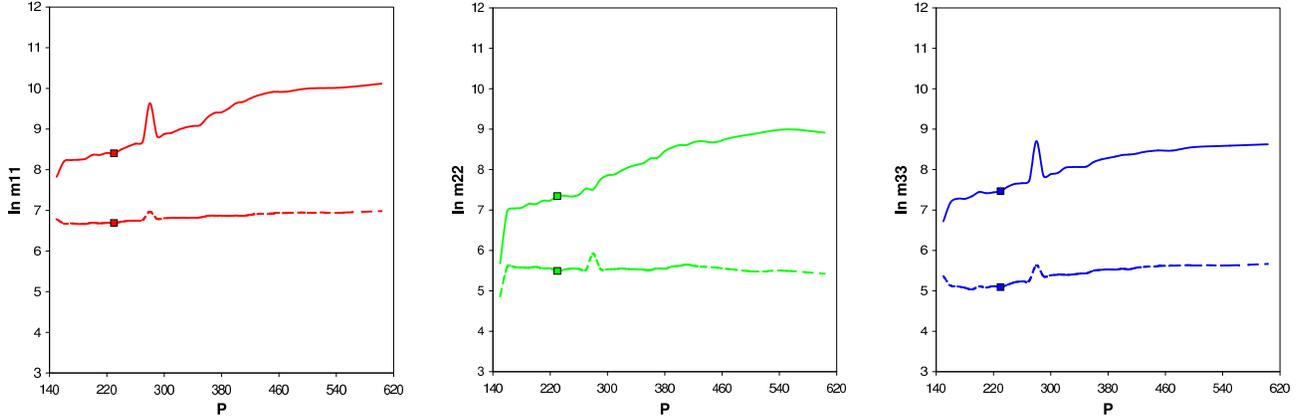


Figure 2: Diagonal central velocity moments (in logarithmic plots) in terms of the sampling parameter  $P = |(U, V, W)|$  for populations segregated from several nested HIPPARCOS subsamples (dashed line for Pop-I and continuous line for Pop-II). The diagonal moments correspond to the velocity components  $U$  (left),  $V$  (centre) and  $W$  (right). The value  $P = 230 \text{ km s}^{-1}$  provided an optimal mixture estimation for thin and thick disks.

decreasing up to  $P = 90$ . For  $P = 65$  the velocity moments (Fig. 6) give typical perpendicular velocity dispersion of thin and thick disk stars, however, the other velocity dispersions have values corresponding to a mixture of disk stars, Pop-I, and halo stars, Pop-II. On the other hand, an opposite behaviour of both plots takes place around  $P = 170$ , where a neat plateau region of the entropy graph begins. In this region both graphs become stable, and around  $P = 180$  and  $P = 250$  small simultaneous local extremes take place. The diagonal central moments, Fig. 6, plotted in terms of the sampling parameter show that halo and disk velocity dispersions clearly reach asymptotic values. Therefore we can take  $P = 180$ , with the minimum  $\chi^2$  error, as the optimal sampling parameter to characterise the mixture of disk and halo stars (Table 1), although there is no significant difference from other neighbour values.

Thus,  $P = |W|$  is not appropriate to distinguish between thin and thick disks, but it is very good to discriminate between total disk and halo stars. Such a good behaviour is also seen when the differential mean velocities are plotted. They can be compared for both sampling parameters in Fig. 7, where the estimation provided by the parameter  $P = |W|$  is totally stable from  $P = 80$  onward.

### 2.1.2 Titius-Bode-like law for the halo

It is interesting to point out that the estimation of the radial velocity dispersion of the halo obtained from the HIPPARCOS sample confirms the Titius-Bode-like law (TBLL) that was suggested in Paper I (Table 6) for the same sample. This empiric law was there written as  $\sigma_U(x) = 6.6 \left(\frac{4}{3}\right)^x$ . The value  $x = 2$  gave the radial velocity dispersion of the early-type stars,  $12 \text{ km s}^{-1}$ , the value  $x = 5$  gave the radial velocity dispersion of the thin disk,  $28 \text{ km s}^{-1}$ , which now has been borne out, the value  $x = 8$  gave the radial velocity dispersion of the thick disk,  $66 \text{ km s}^{-1}$ , which has been also confirmed, and, now, the value  $x = 11$  gives the radial velocity dispersion of the halo,  $156 \text{ km s}^{-1}$ , which is exactly the one obtained in Table 1. We are still not able to add any comment to those of Paper I for a physical explanation of this empirical law, but we may certainly certify this fact, and associate the discrete local stellar populations with the radial dispersions expressed

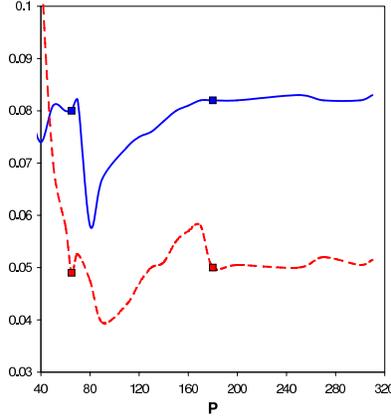


Figure 3: Entropy (continuous line) and  $\chi^2$  error (dashed line, scaled to one twentieth) for HIPPARCOS sample extended up to halo stars, for the sampling parameter  $P = |W|$  (absolute value of the star velocity perpendicular to the Galactic plane). The short plateau region around  $P = 65$  km s $^{-1}$  provides the segregation of the thin disk from thick disk plus halo stars. The  $\chi^2$  behaviour is mainly decreasing up to  $P = 90$ . After  $P = 160$  both plots show an opposite trend with optimal values between  $P = 180$  and  $P = 250$ . The sampling parameter  $P = |W|$  is better than the absolute heliocentric velocity to segregate the whole disk from the halo, but it is unable to segregate properly thin and thick disks.

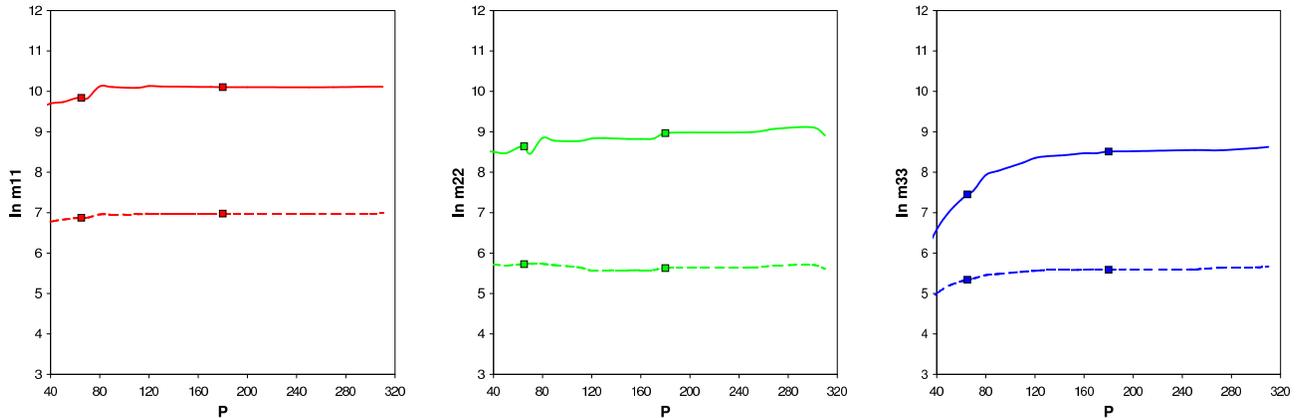


Figure 4: Diagonal central moments (logarithmic plot) in terms of the sampling parameter  $P = |W|$  for both populations segregated from the nested HIPPARCOS subsamples (dashed line for Pop-I and continuous line for Pop-II). The value  $P = 180$  km s $^{-1}$  provides the optimal estimations for disk and halo stars, although they are nearly constant for  $P > 120$ . Notice that the value  $P = 65$  provides the perpendicular velocity dispersions of thin and thick disks, but other velocity components are accounting for a mixture of thin disk in Pop-I and thick disk plus halo stars in Pop-II.

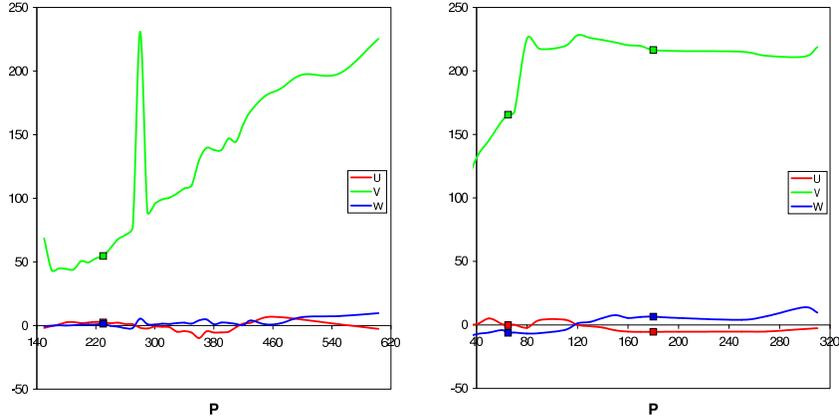


Figure 5: Differential mean velocities between Pop-I and Pop-II for HIPPARCOS subsamples. The velocity components  $U$  (red),  $V$  (green), and  $W$  (blue) are depicted in terms of the absolute heliocentric velocity on the left, and of the absolute perpendicular velocity on the right. On the left plot, for  $400 < P < 600 \text{ km s}^{-1}$  where the total disk and a partial halo are segregated, the differential rotation velocity always increases without reaching an asymptotic value. On the right plot, from  $P = 120$  onward we get constant estimates for the differential rotation mean velocity of the halo, which is approximately at rest (if we assume the usual local rotation velocity of  $V = 220 \text{ km s}^{-1}$ ).

by the equation

$$\sigma_U(n) = 6.6 \left(\frac{4}{3}\right)^{3n+2} ; \quad n = 0, 1, 2, 3 \quad (1)$$

## 2.2 GCS sample

We compare the previous results with those of the Geneva-Copenhagen survey catalogue (Nordström et al. 2004). The same Cartesian heliocentric coordinates system ( $U, V, W$ ) is used. The sample contains the total velocity space of 13,240 stars of F and G dwarf stars, which are considered the favourite tracer populations of the history of the disk. According to the authors, the GCS catalogue might be a kinematically unbiased sample, containing about 97% of thin disk and 3% of thick disk. Nevertheless, some contamination by the halo may exist according to Famaey et al. (2007), which could be removed by considering only stars with  $[\text{Fe}/\text{H}] > -0.5$ . To our purpose, and similarly as in the HIPPARCOS sample, such a contamination could be useful to contrast the kinematic parameters of the three populations.

In regard to the composition of the sample, some remarks can be done. At first, the selection of only specific spectral type stars could imply some kinematical bias of the sample. As we can see in Fig. 8, the velocity distribution in the  $V$  component is not as clock shaped, and is not as bimodal shaped, as it is in the HIPPARCOS sample. In contrast, the  $U$  and  $W$  velocity distributions are quite clock shaped, although HIPPARCOS distributions have higher peaks. In any case MEMPHIS method may still be a good approximation for slightly non-gaussian partial distributions, and can provide good estimations of mixture proportions and velocity moments of populations. For the GCS samples it is also possible to recognise the substructure of early-type and young disk stars within the thin disk, as described in Paper I for the HIPPARCOS sample, although the subpopulations are strongly non-gaussian. However, once the number of thin disk stars added to the nested subsamples is enough large to be approached by a gaussian partial distribution, and once the number of thick disk stars is also enough large to conform an

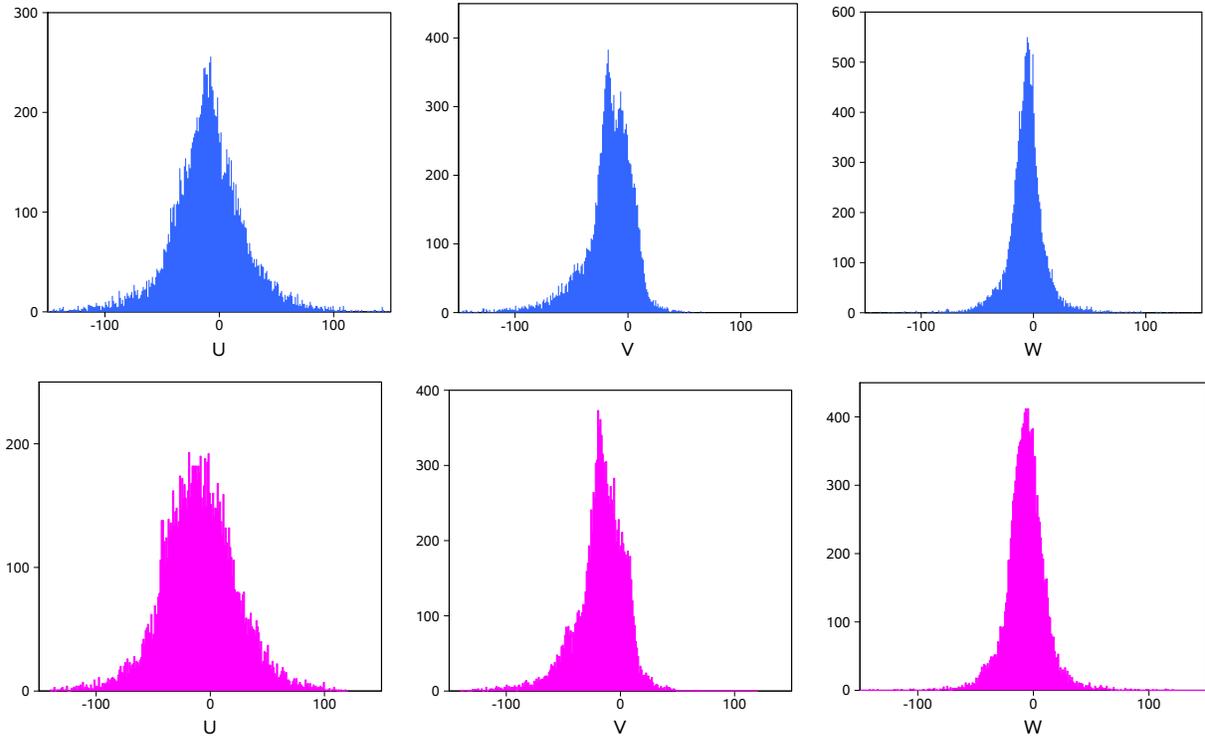


Figure 6: Distribution of star velocity components  $U$ ,  $V$ , and  $W$  for HIPPARCOS sample, 13,678 stars (upper row), and GCS sample, 13,240 stars (lower row). The GCS sample has slightly lower size than the HIPPARCOS sample since it only includes F and G stars. Most of GCS stars come from the HIPPARCOS sample, generally with improved velocity estimations. Such a slight different composition can be seen from the histograms. For example, the central peaks of the  $U$  and  $V$  velocity components are higher in HIPPARCOS'. Also, the  $V$  velocity component of the HIPPARCOS sample has a slightly clearer bimodal shape, and a more rounded right wing.

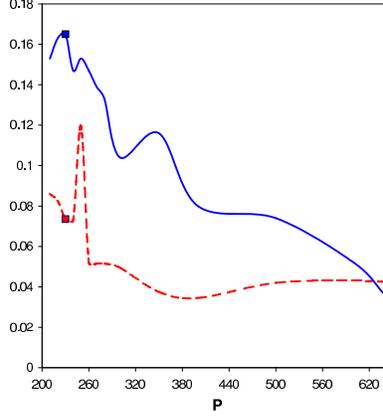


Figure 7: Entropy (continuous line) and  $\chi^2$  error (dashed line, scaled to one twentieth) for GCS sample with the sampling parameter  $P = |(U, V, W)|$  (absolute heliocentric velocity). The lower  $\chi^2$  value in the interval with higher entropy levels, around  $P = 230 \text{ km s}^{-1}$ , can be used for a estimation of thin and thick disks. There is not a clear opposite trend of both graphs until  $P = 300$ , although there is no plateau region for these values.

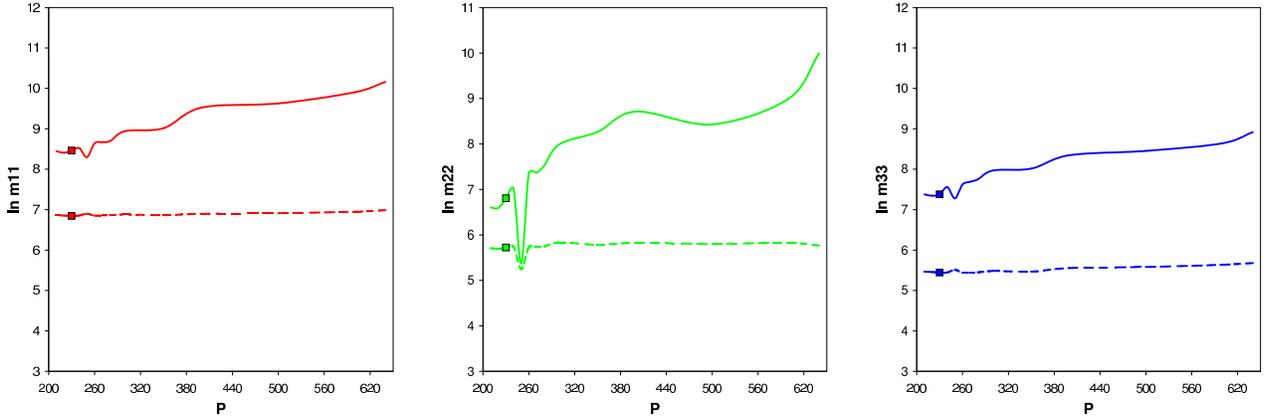


Figure 8: In logarithmic plots, diagonal central moments in terms of the sampling parameter  $P = |(U, V, W)|$  (absolute heliocentric velocity), for both populations segregated from nested GCS subsamples (dashed line for Pop-I and continuous line for Pop-II). For  $P > 400 \text{ km s}^{-1}$  the halo is continuously merged to Pop-II, but not in an asymptotic way.

independent population, the segregation can be well performed.

### 2.2.1 Sampling parameters $P = |(U, V, W)|$ and $P = |W|$

If we use the sampling parameter  $P = |(U, V, W)|$ , we get the entropy and  $\chi^2$  plots of Fig. 9. There is a local minimum of  $\chi^2$  at  $P = 230 \text{ km s}^{-1}$ , simultaneously with a local maximum of entropy, although a general opposite trend of both graphs does not appear until  $P = 300$ . After the initial high entropy levels there is not any plateau region in the entropy graph. If we take a look at the plots of the diagonal moments, Fig. 10, by comparing with HIPPARCOS results we observe that in  $P = 230 \text{ km s}^{-1}$  the algorithm is really detecting, but not in the same stable way, a mixture of thin (96%) and thick (4%) disks (Table 1), while for  $P > 350$  the algorithm is indeed continuously merging halo stars in the Pop-II component, up to the end of the whole sample (ca. 1% of halo stars).

If the absolute velocity perpendicular to the Galactic plane is taken as sampling parameter,  $P =$

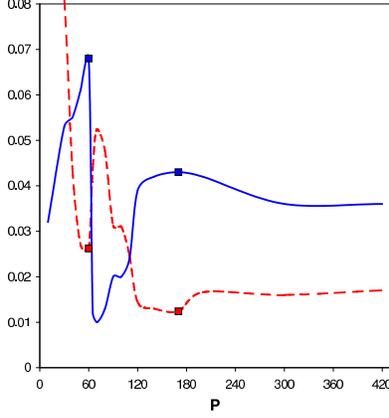


Figure 9: Entropy (continuous line) and  $\chi^2$  error (dashed line, scaled to one fiftieth) graphs for GCS sample, with sampling parameter  $P = |W|$ . At  $P = 60 \text{ km s}^{-1}$  and 170 simultaneous extremes of  $\chi^2$  and entropy take place. The plateau region for  $P > 140$  provides a perfect segregation of disk and halo populations, while  $P = 60$  only detects a mixture of thin and thick disks in the  $W$  velocity.

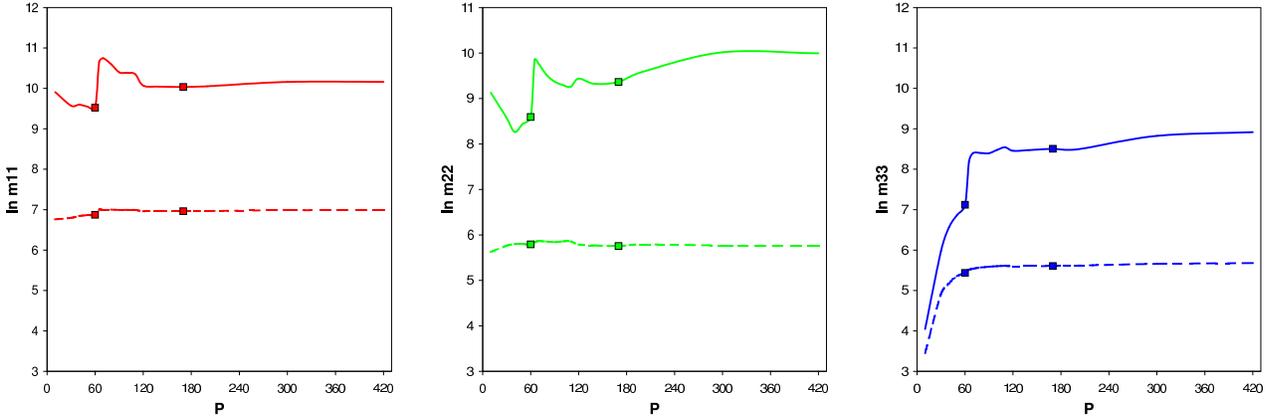


Figure 10: Diagonal central velocity moments in terms of the sampling parameter  $P = |W|$  for both populations segregated from the GCS subsamples (dashed line for Pop-I and continuous line for Pop-II). The value  $P = 170$  provides the optimal mixture estimation for disk and halo stars. Notice the saturation of the disk, Pop-I, for  $P > 60 \text{ km s}^{-1}$ .

$|W|$ , things become much more improved. In Fig. 11 we observe that the entropy graph shows a peak at  $P = 60 \text{ km s}^{-1}$ , and a clear plateau region for  $P > 140$ , with simultaneous extremes of  $\chi^2$  and entropy at  $P = 170$ , where both plots show opposite trends. Looking at the graph of moments, Fig. 12, we see that, similarly as for HIPPARCOS sample, the value  $P = 60$  only provides a neat segregation of thin and thick disks for the  $W$ -velocity component, while in other velocity components it mixes up thick disk and halo in the second population Pop-II. On the contrary, for  $P = 170$ , and also for greater values, the algorithm provides a perfect segregation of disk (99 %) and halo (1%). The optimal estimations are also listed in Table 1. Such a segregation is borne out by the plots of the differential mean velocities, Fig. 13, where the stability of estimates for  $P = |W|$  (right plot) contrasts with those of  $P = |(U, V, W)|$  (left plot).

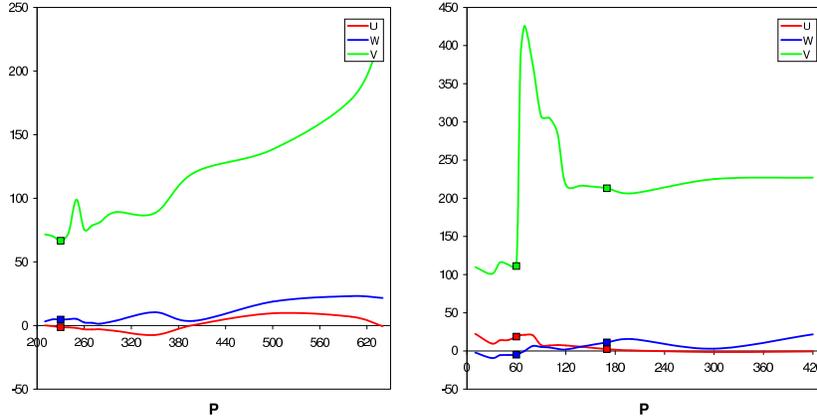


Figure 11: Differential mean velocities between Pop-I and Pop-II. The velocity components  $U$  (red),  $V$  (green), and  $W$  (blue) are depicted in terms of the sampling parameter  $P$ , absolute heliocentric velocity  $|(U, V, W)|$  on the left, and absolute perpendicular velocity  $|W|$  on the right, for GCS subsamples. For the absolute heliocentric velocity, in the region  $P > 260$  km s<sup>-1</sup>, the differential rotation velocity always increases without reaching an asymptotic value, but for the absolute perpendicular velocity, from  $P > 120$  we get some constant estimates with a differential rotation mean velocity corresponding to an halo which is approximately at rest, like in the HIPPARCOS sample.

### 2.2.2 Sampling parameters $P = |(U, W)|$ and $P = |V|$

We also present the results obtained by working with other sampling parameters. The sampling parameter  $P = |(U, W)|$ , which is related to the integral  $I_4$ , produces the entropy and  $\chi^2$  plots of Fig. 14 (left). Around  $P = 200$  km s<sup>-1</sup> and for  $P > 250$  km s<sup>-1</sup> both plots are really behaving with opposite trends, although there is not a clear plateau region in the entropy graph. The two local extremes at  $P = 200$  and  $280$  give account for a mixture of thin disk in Pop-I, and thick disk plus halo in Pop-II. This is clearly seen in the graphs for the diagonal central moments, Fig. 15. On the other hand, in Fig. 14 (right) we also display the differential mean velocities between Pop-I and Pop-II. In nearly all the plotted interval the rotation velocity is typical from halo stars, and around  $P = 320$  km s<sup>-1</sup> there are some moving groups with antigalactic rotation (the null rotation would correspond to  $V \sim 200$  km s<sup>-1</sup>). Indeed this is a similar behaviour we found in Fig. 13 (right) by using the perpendicular velocity as sampling parameter, between  $P = 80$  and  $100$ . Once more, for the greatest values of the sampling parameter, the halo is approximately at rest.

Alternatively, the sampling parameter  $P = |V|$ , which is related to the angular momentum integral  $I_2$ , provides us with the entropy and  $\chi^2$  plots of Fig. 16 (left). The  $\chi^2$  error trend is basically decreasing, while the entropy shows a neat local maximum in  $P = 100$  km s<sup>-1</sup>. Hereafter both plots run with the same decreasing trend. There is no clear plateau region in the entropy graph. The segregation obtained from this maximum entropy value corresponds, according to the velocity moments of Fig. 17, to a thin and thick disk mixture, very similar to the one obtained when working with the sampling parameter  $P = |(U, V, W)|$ . At the end of the graph, for the whole sample, we obtain typical moments of a disk and halo mixture. A similar interpretation applies to the differential mean velocities, Fig. 16 (right). In all the plotted interval the differential rotation velocity is basically increasing up to reach the typical value around  $220$  km s<sup>-1</sup>, or even closer to more recent values like  $235$  km s<sup>-1</sup> (Orlov et al. 2006).

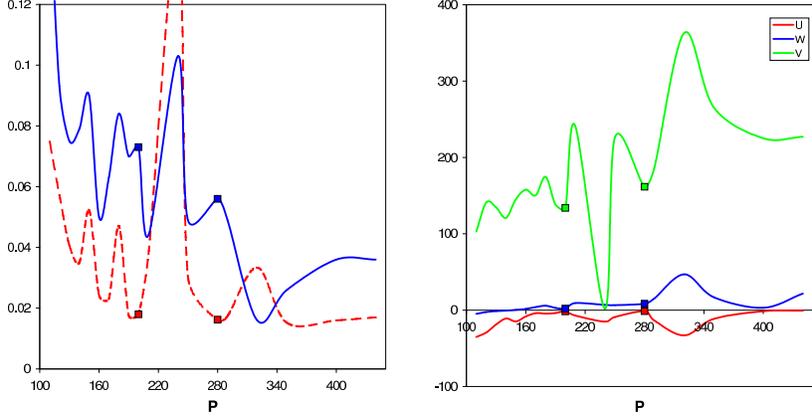


Figure 12: Left: Entropy graph (continuous line) and  $\chi^2$  error (dashed line, scaled to one fiftieth) in terms of the sampling parameter  $P = |(U, W)|$  for GCS subsamples. For  $P > 250 \text{ km s}^{-1}$ , as well as around  $P = 200$ , both plots show opposite trends. Some segregations could be tried for  $P = 200$  and  $P = 280$ . Right: Differential mean velocities in terms of  $P = |(U, W)|$ . Around  $P = 320 \text{ km s}^{-1}$  there are stars with antigalactic rotation. Such a behaviour also appeared in Fig. 13 (right), when working with the sampling parameter  $P = |W|$ . For the highest values of the sampling parameter the halo is approximately at rest.

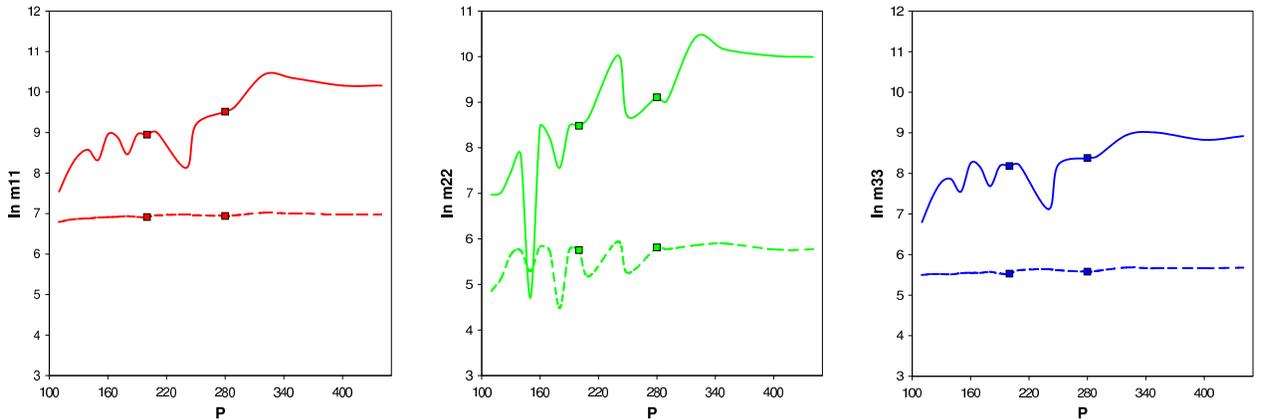


Figure 13: Diagonal central velocity moments (logarithmic plot) in terms of the sampling parameter  $P = |(U, W)|$  (dashed line for Pop-I and continuous line for Pop-II) for GCS subsamples. The values  $P = 200 \text{ km s}^{-1}$  and  $280$ , where the entropy and the fitting error take local extremes, Pop-II velocity moments give account for mixed thick disk and halo values.

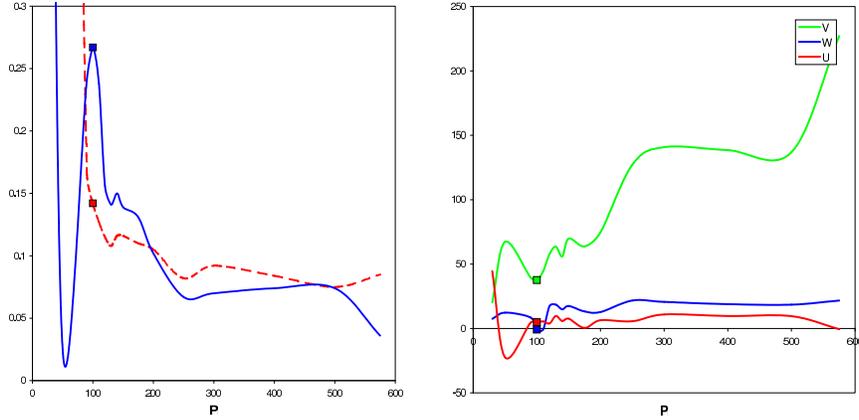


Figure 14: Left: Entropy graph (continuous line) and  $\chi^2$  error (dashed line, scaled to one tenth) in terms of the sampling parameter  $P = |V|$  for GCS subsamples. The entropy shows an only local maximum at  $P = 100 \text{ km s}^{-1}$ . Hereafter both plots mostly run with the same decreasing trend. Right: Differential mean velocities in terms of the sampling parameter. In all the plotted interval the rotation velocity is basically increasing up to reach the typical value for a halo at rest.

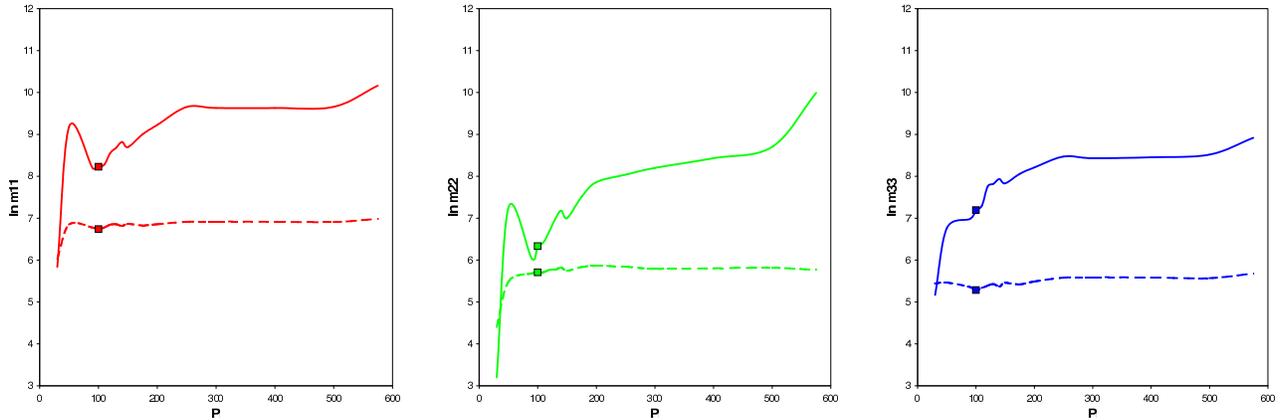


Figure 15: Diagonal central velocity moments (logarithmic plot) in terms of the sampling parameter  $P = |V|$  (dashed line for Pop-I and continuous line for Pop-II) for GCS subsamples. The value  $P = 100 \text{ km s}^{-1}$  gives an approximate thin and thick disk segregation, although with a smooth transition up to the greatest values of the sampling parameter.

Sample	$\#S(P)$	$P$	Pop.	$\sigma_U$	$\sigma_V$	$\sigma_W$	$U$	$V$	$W$	$\varepsilon [^\circ]$
HIP	13,551	$ (U, V, W)  = 230$	t 92%	$28.4 \pm 0.6$	$15.6 \pm 2.5$	$12.8 \pm 0.7$	$-10.7 \pm 0.3$	$-13.9 \pm 0.4$	$-7.2 \pm 0.2$	$10 \pm 2$
			T 8%	$66.8 \pm 2.3$	$39.4 \pm 10.6$	$41.8 \pm 1.9$	$-13.0 \pm 0.6$	$-68.6 \pm 3.1$	$-8.5 \pm 0.2$	$5 \pm 3$
	13,671	$ W  = 180$	D 98%	$32.7 \pm 1.1$	$15.2 \pm 10.7$	$16.4 \pm 0.4$	$-10.8 \pm 0.3$	$-16.6 \pm 0.4$	$-7.2 \pm 0.2$	$8 \pm 5$
			H 2%	$156.4 \pm 10.4$	$88.5 \pm 53.2$	$70.5 \pm 4.3$	$-5.3 \pm 2.5$	$-233.1 \pm 20.4$	$-13.7 \pm 3.6$	$-4 \pm 5$
GCS	13,185	$ (U, V, W)  = 230$	t 96%	$30.6 \pm 0.4$	$17.5 \pm 1.4$	$15.2 \pm 0.3$	$-10.0 \pm 0.3$	$-15.2 \pm 0.3$	$-6.9 \pm 0.2$	$9 \pm 1$
			T 4%	$68.8 \pm 2.7$	$30.0 \pm 15.0$	$40.0 \pm 2.0$	$-8.6 \pm 0.7$	$-81.8 \pm 4.3$	$-11.6 \pm 0.5$	$6 \pm 2$
	13,235	$ W  = 170$	D 99%	$32.5 \pm 0.7$	$17.8 \pm 5.7$	$16.5 \pm 0.3$	$-9.9 \pm 0.3$	$-16.9 \pm 0.3$	$-7.1 \pm 0.2$	$8 \pm 4$
			H 1%	$151.0 \pm 14.7$	$107.9 \pm 54.0$	$70.3 \pm 5.7$	$-12.2 \pm 2.1$	$-230.0 \pm 24.1$	$-18.3 \pm 3.3$	$-3 \pm 7$

Table 1: Optimal segregations for HIPPARCOS (HIP) and GCS samples. The displayed quantities are: size of the optimal sample, sampling parameter, segregated population (t=thin disk, T=thick disk, D=total disk, H=halo) and mixture proportion, velocity dispersions, mean velocities (both in  $\text{km s}^{-1}$ ), and vertex deviation. These are the population estimates which are consistent with the mixture model. The greater uncertainties of those related to rotation velocity are due to the error propagation from the mixture model, since partial moments and differential velocity are in gear.

## 2.3 Precise determination of mixture proportions

The estimation of the mixture proportions has an intrinsic uncertainty due to the segregation method itself. For mixtures of pure normal populations the segregation is performed with a high degree of accuracy, of about 0.5–1%, as it was tested by working from simulated samples in Paper II. In general, the error depends on the gaussianity of the components. Otherwise, if a non-pure normal population is approximated by a gaussian distribution alone, as in the case of a cumulative population Pop-I containing the thin and thick disks, the uncertainty of the mixture proportion may increase. To give some more accurate mixture proportions, once the kinematic parameters of the three segregated populations have been determined, we apply an iterative method based in the bayesian probability of a star to belong to any particular population.

With the mixture proportions given in Table 1 as initial values, we evaluate the total velocity distribution function as the mixture

$$f(v) = \sum_{i=1}^3 p(A_i) f(v|A_i) \quad (2)$$

where the populations  $\{A_1, A_2, A_3\}$  refers either to the halo, thick disk and thin disk, respectively. Notice the notation  $f(v|A_i) = f^{(i)}(v|X)$ . Each gaussian component  $f(v|A_i)$  is evaluated according to its six central second moments  $\Sigma_i$  (the covariance matrix), which determinant is noted as  $|\Sigma_i|$ , and to its mean velocity  $m_i$ . Thus, we can write

$$f(v|A_i) = \frac{1}{(2\pi)^{\frac{3}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2} (v - m_i)^t \cdot \Sigma_i^{-1} \cdot (v - m_i)} \quad (3)$$

Then, for each star of the sample, by applying Bayes' theorem, we compute the probability of belonging to any of the populations, in basis to its actual velocity,

$$p(A_i|v) = \frac{p(A_i) f(v|A_i)}{f(v)} \quad (4)$$

and we then evaluate the expected value  $y = E(i|v)$  of the index  $i$  representing the populations,

$$y = \sum_{i=1}^3 i p(A_i|v) \quad (5)$$

which takes continuous values in the interval  $[1, 3]$ .

We may assume that the value  $y$  is uniformly distributed in the interval  $[1, 3]$ . Then, if  $1 \leq y \leq 1.67$  the star has greater probability of belonging to the halo population  $A_1$ , if  $1.67 < y \leq 2.33$  the star would belong to the thick disk  $A_2$ , and if  $2.33 < y \leq 3$  it is assigned to the thin disk  $A_3$ . Thus, the stars are now labelled according to the most likely population. We can then evaluate the mixture proportions. These values are substituted into Eq. 2 and Eq. 4, and the process may be iterated up to convergence.

For both samples the mixture proportions become stable after 4 iterations, by leading to values of Table 2. Fig. 18 displays the expected value  $y$  in terms of the number of stars, in logarithmic scale, where it is possible to see three well defined levels corresponding to the classification of most of the stars into three discrete populations.

The resulting thin disk, thick disk, and halo components are shown in Fig. 19 for the HIPPARCOS sample, and in Fig. 20 for GCS sample. This classification of stars into population samples allows

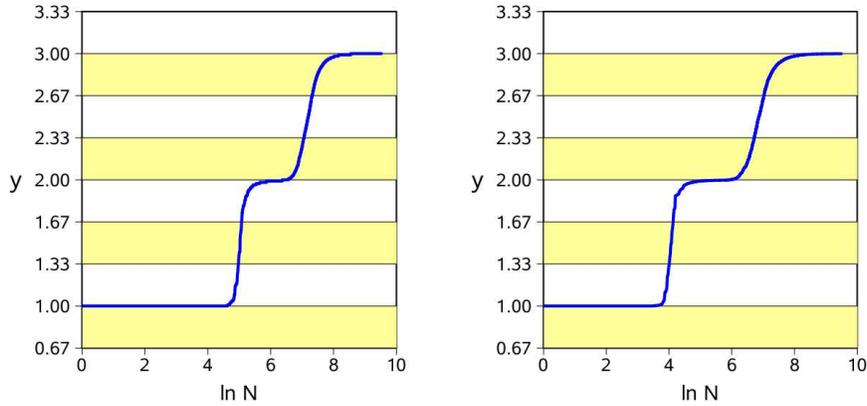


Figure 16: Expected value  $y$  in terms of the number of stars (in logarithmic scale, HIPPARCOS on the left, and GCS on the right) where it is possible to see three well defined levels of classification for most of the stars. However there is still a significant number of stars between thick and thin disks, in the range  $2 < y < 3$ . If  $1 \leq y \leq 1.67$  the star has greater probability of belonging to the halo population, if  $1.67 < y \leq 2.33$  the star would belong to the thin disk, and if  $2.33 < y \leq 3$  it is assigned to the thin disk.

to compute the sample moments of thin disk, thick disk and halo according to Table 2, which provide very similar kinematic parameters for the population samples drawn from both catalogues. Let us note that the resulting population samples depend on the way of to assign the population representative range to the expected value  $y$ . Thus, if instead of assuming the boundaries between populations as 1-1.66-2.33-3, which correspond to an uniform distribution of  $y$ , we had assumed boundaries corresponding to equidistant points from 1, 2 and 3, that is ranges 1-1.5-2.5-3, we had obtained population fractions t 89.9%, T 9.0%, H 1.1% for HIPPARCOS, and t 95.5%, T 4.1%, H 0.4% for GCS. However, we find more plausible the former choice.

It is interesting to remark the difference between Tables 1 and 2. Table 1 describes the population distribution parameters, which are consistent with the mixture model. In particular, the rotation velocity dispersions have greater uncertainties since they geared to the differential rotation velocity between populations. Therefore the uncertainties are the consequence of applying error propagation techniques to the mixture model. On the other hand Table 2 lists the population sample moments, which are obtained by segregating the whole samples according to the population distribution moments of Table 1. Thus, the standard errors are sampling variances which do not reflect the mixture model.

### 3 Discussion

In Cubarsi & Alcobé (2009) we had described the properties of a sampling parameter used to build a set of hierarchical subsamples. It was related to isolating integrals of the star motion, which may take similar values among stars of a specific population, but very different from other populations. However, no isolating integral alone is capable of isolate single populations, since actual population distributions always overlap. In the best of situations it is possible to use a sampling parameter allowing to build a family of nested subsamples each one containing, at most, one more population than others. Thus, by continuously increasing the sampling parameter the new stars entering the subsample belong either to a single population, or are a mixture of two nearby populations. Hence, in addition to the sampling parameter, a segregation algorithm is

Sample	# <i>S</i>	Pop.	$\sigma_U$	$\sigma_V$	$\sigma_W$	<i>U</i>	<i>V</i>	<i>W</i>	$\varepsilon$ [°]
HIP	12,516	t 91.5%	28.2 ± 0.2	16.9 ± 0.2	12.5 ± 0.1	-11.2 ± 0.3	-14.7 ± 0.2	-7.1 ± 0.1	13 ± 1
	1,003	T 7.3%	69.3 ± 1.3	37.9 ± 0.9	42.9 ± 0.9	-7.1 ± 2.2	-60.8 ± 1.2	-9.7 ± 1.4	4 ± 2
	159	H 1.2%	179.6 ± 7.8	89.0 ± 6.3	90.8 ± 6.0	-0.61 ± 14.2	-234.9 ± 7.0	-10.8 ± 7.2	-6 ± 3
GCS	12,415	t 93.8%	29.7 ± 0.2	17.8 ± 0.2	13.9 ± 0.1	-10.4 ± 0.3	-15.2 ± 0.2	-7.1 ± 0.1	11 ± 1
	763	T 5.7%	65.7 ± 1.5	36.7 ± 1.1	41.5 ± 0.9	-3.4 ± 2.4	-59.3 ± 1.3	-7.1 ± 1.5	5 ± 2
	62	H 0.5%	178.6 ± 12.2	113.7 ± 15.0	110.5 ± 14.2	1.6 ± 22.7	-230.8 ± 14.4	-16.4 ± 14.0	-5 ± 10

Table 2: Segregation of populations for HIPPARCOS and GCS samples. The displayed quantities are: size of the sample, segregated population (t=thin disk, T=thick disk, H=halo), mixture proportion, velocity dispersions, mean velocities (both in  $\text{km s}^{-1}$ ), and vertex deviation. The more accurate mixture proportions have been determined by applying an iterative method based in the bayesian probability of a star to belong to any particular population. The velocity dispersions, which are computed from population sample moments, are obtained by segregating the whole samples according to the population distribution moments of Table 1. Thus, the standard errors are sampling variances which do not reflect the mixture model.

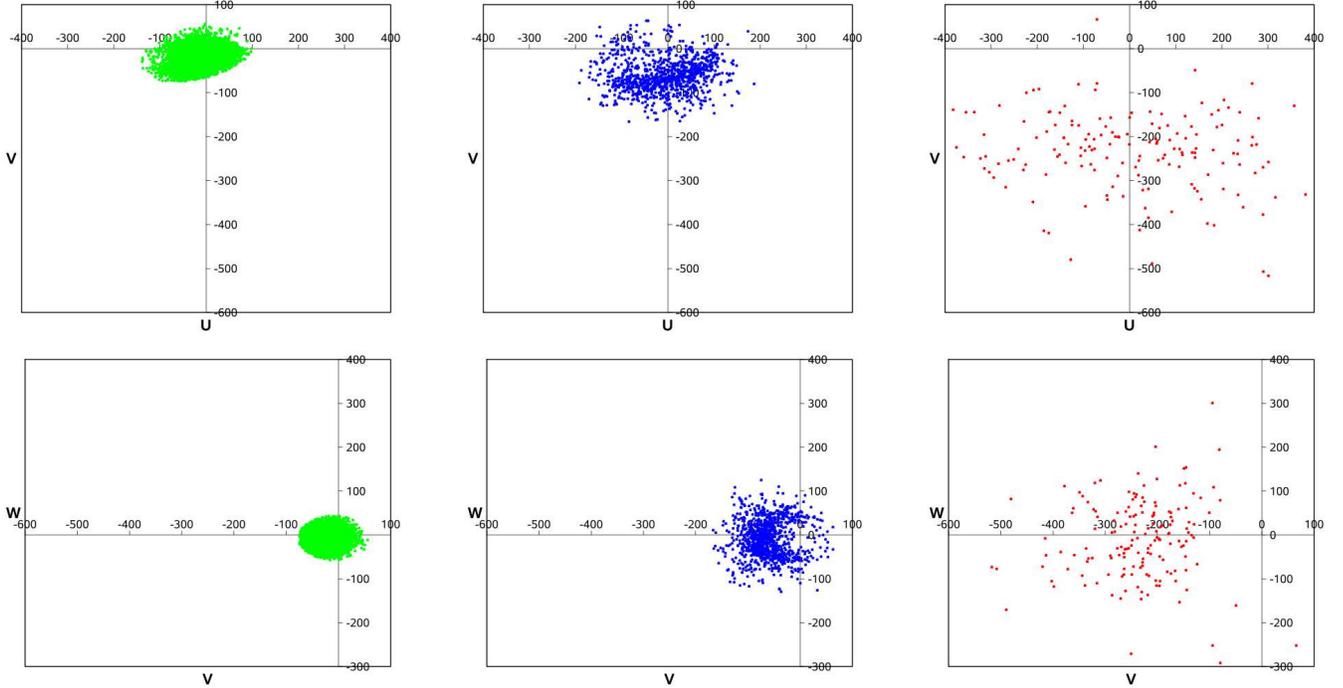


Figure 17: Distribution of thin disk (left), thick disk (centre), and halo (right) for HIPPARCOS catalogue stars. The distribution on the plane  $UV$  of heliocentric velocities is shown in the first row, and on the plane  $VW$  in the second row.

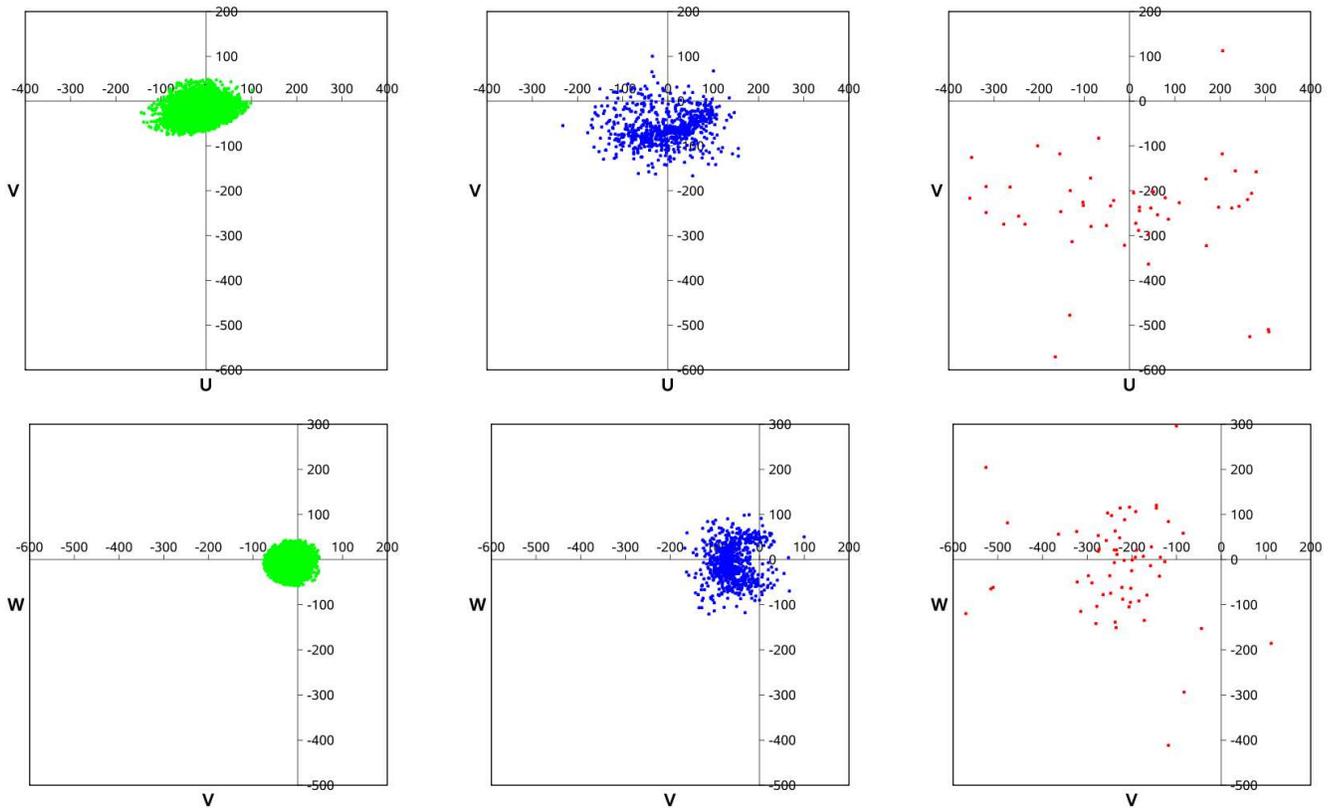


Figure 18: Distribution of thin disk (left), thick disk (centre), and halo (right) for GCS catalogue stars. The distribution on the plane  $UV$  of heliocentric velocities is shown in the first row, and on the plane  $VW$  in the second row.

needed to identify the mixture. Now, several sampling parameters have been tested for stellar populations of the solar neighbourhood. The sampling parameter  $P = |(U, V, W)|$ , absolute value of the heliocentric velocity, which is a combination of the three basic integrals of motion, allows to build an optimal subsample containing thin and thick disk stars, by leaving aside, for  $P = 230 \text{ km s}^{-1}$ , most of the halo population. The sampling parameter  $P = |W|$ , absolute value of the perpendicular velocity, depending on the so called Oort's third integral, which is valid near the Galactic plane, is able to build, for  $P = 180 \text{ km s}^{-1}$ , an optimal subsample containing a well defined mixture of total disk and halo stars, although it does not allow an optimal segregation of thin and thick disks. The sampling parameter  $P = |(U, W)|$ , which is related to a combination of the energy and angular momentum integrals, is much less stable than the others in estimating kinematic populations, and allow to isolate the thin disk component from a component containing thick disk and halo stars altogether. Finally,  $P = |V|$ , which is related to the angular momentum integral, works in a similar way than the absolute value of the heliocentric velocity, but with less accuracy.

The application to the HIPPARCOS sample with the sampling parameter  $P = |(U, V, W)|$  has been carried out to compare with previous results of Paper I. We have obtained a better estimation of thin and thick disk kinematic parameters, which are listed in Table 1. Moreover, the kinematic parameters and the population fraction for the halo have now been determined by using the sampling parameter  $P = |W|$ .

The use of the GCS sample had two motivations. First, the sample has new and more accurate radial velocity data than HIPPARCOS sample. Second, while HIPPARCOS sample is unconstrained, in the sense that all the available neighbour stars are included, GCS sample should only contain disk representative stars. It is therefore interesting to compare both results. The thin and thick disk kinematic parameters are also obtained working from  $P = |(U, V, W)|$ , although they are not obtained as stable as for HIPPARCOS, and the thick disk kinematic estimates have greater uncertainty. Similarly,  $P = |W|$  provides a good characterisation of the total disk and halo kinematics, although halo estimates are worse than for HIPPARCOS. Thus, HIPPARCOS sample leads to better estimations of thick disk and halo populations. Total disk characterisations are similar for both samples, although GCS has more accurate estimates. The greater thin disk velocity dispersions of the GCS sample suggest a kinematic bias of the sample towards older disk stars. Therefore, by working from the GCS sample we do not get a dramatic improvement of the local kinematic portrait, and we have preferred to use the moment estimates obtained from HIPPARCOS in order to characterise the local populations. We shall try some explanations to justify such a decision.

First, let us note that the errors in the estimates corresponding to the rotation velocity are greater than those of other components. For the GCS samples they are also greater than for the HIPPARCOS samples. The reasons are basically two. On one hand, when trying a disk-halo segregation, thin and thin disk stars are joined in the population Pop-I, while halo stars remain in Pop-II. Since thin and thick disk stars have a significantly different rotation mean velocity, the gaussian approach given by Pop-I leads to an artifactual error in such a velocity component. This is not so evident in the other velocity components, since they have nearly the same means. In addition, the GCS sample lags a part of the wing of the thin disk distribution, mainly in the rotation direction (Fig. 8), which leads to a greater error of the gaussian mixture approach. Indeed, Holmberg et al. (2007) remark that the velocity distribution is not like a two normal mixture, although we are actually using a three normal mixture. Also notice that estimates of minor populations have always greater uncertainties, by leading, in this case, to worse thick disk and halo kinematic estimates.

For the same reason, the partial lag of early-type or younger stars within the thin disk component of the GCS sample leads to slightly greater thin disk velocity dispersions than for the HIPPARCOS sample. In the HIPPARCOS sample, those stars had been mostly included (Paper I) in the nested subsamples up to an absolute heliocentric velocity  $P < 50 \text{ km s}^{-1}$ . In total 10,128 stars, from which 10,124 stars are contained within the thin disk component (12,516 stars, Table 2). Therefore, they set up 81% of the total thin disk. Instead, for the GCS sample, there are 9,568 stars with  $P < 50$ , from which 9,564 stars are also contained within the thin disk component (12,415 stars, Table 2). This only is 77% of the thin disk. Hence the younger stars, which contribute to a lower value of the thin disk velocity dispersions, are underweighted in the GCS catalogue, which provides some characteristic dispersions biased towards older disk stars.

If we centre our attention to the total disk and halo segregation obtained either for HIPPARCOS or GCS samples by selecting the subsamples from the absolute perpendicular velocity  $P = |W|$ , we see that both populations become saturated as enlarging the velocity samples, by increasing the sampling parameter, as it is shown in Figs. 6 and 12. This happens for Pop-I, containing thin and thick disk stars, as well as for Pop-II, containing halo stars. Notice that this is also quite evident for disk stars, when working from the absolute heliocentric velocity  $P = |(U, V, W)|$ , as it is shown in Figs. 4 and 10, although in this case the halo shows a continuous increasing estimation of velocity dispersions. Therefore, it seems that the question about the likely disk unsaturation stated by Holmberg et al. (2007) may now be answered: The apparent unsaturation was likely due to some contamination with halo stars. Now, by segregating the halo from the disk, we find a stable estimation of the kinematical parameters for both local Galactic components.

On the other hand, the mixture proportions listed in Table 1 have been taken as initial values for an iterative method leading to a more accurate estimation of the population fractions. For each star, the bayesian probability of belonging to a particular population is evaluated, and the star is assigned to the most likely population. The process is iterated up to reach a constant population assignation. The resulting mixture proportions and sample moments are listed in Table 2. Although initial and final fractions are quite similar, for the GCS sample a small difference is very significant, since it only contains 0.5% of halo stars, less than a half of HIPPARCOS fraction. Working from the segregated population samples, the sample moments have been calculated, which are totally equivalent either they have been obtained from HIPPARCOS or from GCS samples. HIPPARCOS sample leads once more to more accurate sample estimates for thick disk and halo populations, which is clearly due to the size of the population samples.

The results on population proportions for the GCS sample are also consistent with those of Nordström et al. (2004), Holmberg et al. (2007) and (Famaey et al. 2007). We find that the halo represents about 0.4 – 0.5%, and thick disk would have a contribution of about 4.1 – 5.7% of the whole sample (these authors are suggesting about 1% and 3% respectively). Then, according to our kinematic estimates, some supposed halo stars are in fact thick disk stars. This may be related to the exclusion of halo stars from  $[Fe/H] < -0.5$  (Famaey et al. 2007). Indeed, our accurate estimation of the population fraction has labelled 62 halo stars, from which 60 stars are  $[Fe/H] \leq -0.5$ , 1 star (Hipp. number 103,311) is  $[Fe/H] = -0.07$ , and 1 star has no metallicity value (Hipp. number 24,186). However, we must warn that within the thick disk component there is a fraction of 33% with  $[Fe/H] \leq -0.5$ , and within the thin disk there is still a fraction of 7% stars with this metallicity.

Finally, in regard to the multivariate mixture model, it seems that the results from the HIPPARCOS sample have given a more complete kinematical description of the solar neighbourhood than those of the GCS sample, even despite to the improved accuracy of radial velocities in the later sample. It seems that the preselection work of the GCS sample in order to get a representative disk

sample has partially masked the actual features of the thin disk velocity distribution, or, at least, has weakened the MEMPHIS working hypothesis of less informative partial distributions. On the contrary, the unconstrained HIPPARCOS sample has left our algorithm to do the statistical work for a better estimation of population kinematical parameters.

When applying Conan Doyle's suggestion "Some facts should be suppressed, or, at least, a just sense of proportion should be observed in treating them", the first arising question might be "which ones of the facts?". In the current work, our response may well have been: let the statistics do the work.

## References

- Alcobé S., Cubarsi R. 2005, *A&A*, 442, 929
- Battaglia G., Helmi A., Morrison H., Harding P., Olszewski E.W., Mateo M., Freeman K.C., Norris J., Shectman S.A., 2005, *MNRAS*, 364, 433B
- Cubarsi R. 1992, *AJ*, 103, 1608
- Cubarsi R., Alcobé S. 2004, *A&A*, 427, 131
- Cubarsi R., Alcobé S. 2009, in *The improved MEMPHIS algorithm - Part I. Maximum partition entropy and minimum  $\chi^2$  error*. Dept. Matemàtica Aplicada IV (Barcelona: UPC e-prints)
- ESA, 1992, *The Hipparcos Input Catalogue*, ESA SP-1136
- ESA, 1997, *The HIPPARCOS Catalogue*, ESA SP-1200
- Famaey B., Pont F., Luri X., Udry S., Mayor M., Jorissen, A. 2007, *A&A*, 461, 957
- Freeman K., Bland-Hawthorn J. 2002, *ARA&A*, 40, 487F
- Fuchs B., Dettbarn C., Rix H.W., Beers T.C., Bizyaev D., Brewington H., Jahrei H., Klement R., Malanushenko E., Malanushenko V., Oravetz D., Pan K., Simmons A., Snedden S. 2009, *AJ*, 137, 4149F
- Helmi A. 2008, *A&ARv*, 15, 145H
- Holmberg J., Nordström B., Andersen J. 2007, *A&A*, 475, 519
- Nordström B., Mayor M., Andersen J., Holmberg J., Pont F., Jørgensen B.R., Olsen E.H., Udry S., Mowlavi N. 2004, *A&A*, 418, 989
- Orlov V.V., Mylläri A., Stepanischev S., Ossipkov L.P. 2006, *A&AT*, 25,161O
- Reddy B.E., Lambert D.L., Allende-Prieto C. 2006, *MNRAS*, 367, 1329R
- Thom C., Flynn C., Bessell M.S., Hninen J., Beers T.C., Christlieb N., Holmberg J.D., Gibson B.K. 2005, *MNRAS*, 360, 354T
- Vidojević S., Ninković S. 2008, *SerAJ*, 177, 39
- Vidojević S., Ninković S. 2009, *AN*, 330, 46