Structure of the velocity distribution of the Galactic disc. A maximum entropy statistical approach - Part I

Rafael Cubarsi

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

September 2009

Abstract

The maximum entropy approach is proposed to describe the local structures of the velocity distribution, which are collected through its sample moments. The method is used with several samples from the HIPPARCOS and Geneva-Copenhagen survey catalogues. For the large-scale distribution, the phase density function may be obtained by fitting moments up to sixth order as a product of two exponential functions, one giving a background ellipsoidal shape of the distribution and the other accounting for the skewness and for the slight shift in the ellipsoidal isocontours in terms of the rotation velocity. The small-scale distribution can be deduced from truncated distributions, such as velocity-bounded samples with $|\mathbf{V}| \leq 51$ km s^{-1}, which contain a complex mixture of early-type and young disc stars. By fitting up to ten-order moments, the maximum entropy approach gives a realistic portrait of actual asymmetries, showing a clear bimodal pattern: (i) around the Hyades-Pleiades stream, with negative radial mean velocity and (ii) around the Sirius-UMa stream, with slightly positive radial mean velocity. The "U-anomaly" along the radial direction is estimated straightforwardly $30 - 35$ km s^{-1} from the contour plots.

KEY WORDS: stars: kinematics – galaxies: kinematics and dynamics – galaxies: statistics – methods: statistical.
1991 MATHEMATICS SUBJECT CLASSIFICATION: 60, 62, 85.

1 Introduction

We present a further application of the maximum entropy approach (Cubarsi 2008) to describe the the still smooth structure appearing in the the velocity distribution of the local Galactic disc. Firstly, let us briefly explain how to interpret a maximum entropy density function, or better, what is the appropriate context for its use.
Up to a change of sign, Shannon's information entropy is defined as the Boltzmann H-functional, which first appeared in statistical mechanics in works by Boltzmann and Gibbs in the 19th century. However, it is not exactly the same concept.
Boltzmann's functional is used for non-equilibrium systems and is related to the irreversibility of dynamical processes in a uniform gas. For elastic collisions involving short-range forces and in the absence of boundaries, mass, momentum, and energy are conserved in binary encounters (e.g. Cercignani 1988). They are usually referred to as collisional invariants. There is only one distribution function, the Maxwellian distribution, fulfiling all of the following properties: it depends on a linear combination of the collisional invariants, the collision term of the Boltzmann equation is exactly zero, and it minimises Boltzmann's entropy. This solution represents a local equilibrium state, in the sense that other solutions to the Boltzmann equation will become closer to it as the time goes by. Depending on the potential, boundary conditions, and dissipative or collision effects (e.g. Villani 2002), maximum entropy solutions can be non-Maxwellian.
Shannon's information entropy[1] was introduced in communication theory to measure the redundancy of a language and the maximal compression rate, which is applicable to a message without any loss of information. It is defined for complex systems and is related to Boltzmann's entropy as a measure of the number of microstates associated with a given macroscopic configuration. On the other hand, the Fisher information was introduced as part of his theory of efficient statistics as a measure of the uncertainty. It is also related to Shannon's entropy, so that the entropy quantifies the variation of information. If we maximise the entropy subject to some constraints (e.g. statistics describing macroscopic properties) we get distributions containing maximum uncertainty that is compatible with these constraints.
For given mass and energy, the Fisher information takes its minimum value and Shannon's entropy its maximum value in the form of Maxwellian distributions. For a given covariance matrix, they take extreme values for Gaussian distributions. The number of constraints involved in the Lagrange multipliers may reach higher order moments, by reflecting more complex situations in which the stars interact with the potential and with themselves, as well as having different masses.

The maximum entropy approach will be used to describe the main kinematical features of solar neighbourhood stars by working from two large and kinematically representative local stellar samples. In the first case, the large-scale distribution of the local disc is inferred from Sample I (Cubarsi & Alcobé 2004), obtained by crossing the HIPPARCOS Catalogue with radial velocities from the HIPPARCOS Input Catalogue (ESA 1992). In the second case, the method is applied to Sample II from the Geneva-Copenhagen survey (GCS) catalogue (Nordström et al 2004). It has new and more accurate radial velocity data than the HIPPARCOS sample, and contains the total velocity space of F and G dwarf stars, which are considered the favourite tracer populations

[1]The quotation by Shannon, extracted from Martin & England (1981), is amusing: My greatest concern was how to call it. I thought of calling it 'information'. But the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John Von Neumann, he had a better idea. He told me: "You should call it entropy, for two reasons. In the first place your uncertainty has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage".

of the history of the disc. In both cases, the largest samples providing stable velocity moments are used. The preceding applications provide and confirm some general and well known trends in the background velocity distribution, such as the overall vertex deviation, the skewness, or the symmetry plane of the distribution. These stellar samples, which mainly contain thin and thick disc stars, can be sufficiently described from an exponential density function with a four-degree polynomial, although a six-degree polynomial provides a more accurate portrait of the local velocity distribution. According to the maximum entropy modelling, it is possible to interpret the velocity distribution as a product of two exponential functions, the one giving a background ellipsoidal shape of the distribution and the other, which is even and at least quadratic in the rotation velocity alone, acting as a perturbation factor that breaks the distribution symmetry.
On the other hand, the small-scale velocity distribution of the local disc can be deduced from truncated distributions. According to Alcobé & Cubarsi (2005), hereafter Paper I, a selection of stars with an absolute value of the total space motion $|\mathbf{V}| \leq 51$ km s^{-1} leaves the older disc stars aside. Such a selection is analysed in more depth and their properties described better. It contains a complex mixture of early-type and young disc stars for which a Gaussian mixture approach is not feasible. Thus, Sample III is built as a subsample of Sample I with $|\mathbf{V}| \leq 51$ km s^{-1}. Finally, Sample IV is drawn from Sample II under the same condition on the absolute velocity.
It is also possible to obtain a more detailed structure of the velocity distribution for specific subsamples, allowing the results of our approach to be compared with the small-scale structure sustained by moving groups. Among metallicity, colour, and other star properties, the eccentricity of the star's orbit is found to behave as a very good sampling parameter that allows distinguishing between different *eccentricity layers* within the thin disc, and allowing visualisation of the underlying structure of the distribution. In particular, for maximum eccentricity 0.3 and maximum distance to the Galactic plane 0.5 kpc, we get a representative thin disc sample.
For these truncated distributions, the density function needs a six-degree polynomial to describe their strong asymmetries and their main kinematic features. The improvement in the GCS catalogue over the HIPPARCOS catalogue provides a higher resolution contour plot for the inner thin disc, which in addition to describing a velocity distribution far from the ellipsoidal hypothesis, explains a clear bimodal structure. Therefore, the maximum entropy modelling can be presented as an alternative way instead of mixture models.

2 Improved parameter estimation

The maximum entropy method (Cubarsi 2008) led to solve the following Gramian system (Eq. 33, Cubarsi 2008)

$$\mathbf{Y} = \mathbf{G}_2\,\mathbf{X}; \quad \mathbf{Y} = [\mathbf{A}, \mathbf{B}, \mathbf{C}], \mathbf{X} = [\mathbf{a}, \mathbf{b}, \mathbf{c}] \tag{1}$$

Since the matrices $\mathbf{X}$ and $\mathbf{Y}$ consist of three column vectors, we dispose of a number of $3\binom{n+2}{3}$ equations. This number, for $n > 1$, is always greater than the number of independent unknowns (leaving out the normalisation factor). For example, in the case $n = 2$, we have 12 equations and 9 independent unknowns, because the symmetric coefficients λ_{12}, λ_{13}, and λ_{23} are equivalent to λ_{21}, λ_{31}, and λ_{32}, respectively, and similarly for higher values of n. In general, if the true distribution is indeed a maximum entropy distribution, the actual moments will be consistent with the symmetry of the coefficients, but a significant deviation from the maximum entropy property will produce some non-symmetric coefficients[2]. To avoid this situation, an equivalent overdeterminate system

[2]This is true for $n > 2$, but for $n = 2$ the coefficients λ_{ij} are related to the second central moments μ_{ij}^{-1}, which are necessarily symmetric.

Table 1: Coefficient submatrix relating the first column **A** of matrix **Y** and the first column **a** of matrix **X**.

$$
\begin{bmatrix}
1 & 0 & 0 & m_1 & m_2 & m_3 & 0 & 0 & 0 & m_{11} & 2\,m_{12} & 2\,m_{13} & m_{22} & 2\,m_{23} & m_{33} & 0 & 0 & 0 & 0 & \cdots \\
m_1 & 0 & 0 & m_{11} & m_{12} & m_{13} & 0 & 0 & 0 & m_{111} & 2\,m_{112} & 2\,m_{113} & m_{122} & 2\,m_{123} & m_{133} & 0 & 0 & 0 & 0 & \cdots \\
m_2 & 0 & 0 & m_{12} & m_{22} & m_{23} & 0 & 0 & 0 & m_{112} & 2\,m_{122} & 2\,m_{123} & m_{222} & 2\,m_{223} & m_{233} & 0 & 0 & 0 & 0 & \cdots \\
m_3 & 0 & 0 & m_{13} & m_{23} & m_{33} & 0 & 0 & 0 & m_{113} & 2\,m_{123} & 2\,m_{133} & m_{223} & 2\,m_{233} & m_{333} & 0 & 0 & 0 & 0 & \cdots \\
m_{11} & 0 & 0 & m_{111} & m_{112} & m_{113} & 0 & 0 & 0 & m_{1111} & 2\,m_{1112} & 2\,m_{1113} & m_{1122} & 2\,m_{1123} & m_{1133} & 0 & 0 & 0 & 0 & \cdots \\
m_{12} & 0 & 0 & m_{112} & m_{122} & m_{123} & 0 & 0 & 0 & m_{1112} & 2\,m_{1122} & 2\,m_{1123} & m_{1222} & 2\,m_{1223} & m_{1233} & 0 & 0 & 0 & 0 & \cdots \\
m_{13} & 0 & 0 & m_{113} & m_{123} & m_{133} & 0 & 0 & 0 & m_{1113} & 2\,m_{1123} & 2\,m_{1133} & m_{1223} & 2\,m_{1233} & m_{1333} & 0 & 0 & 0 & 0 & \cdots \\
m_{22} & 0 & 0 & m_{122} & m_{222} & m_{223} & 0 & 0 & 0 & m_{1122} & 2\,m_{1222} & 2\,m_{1223} & m_{2222} & 2\,m_{2223} & m_{2233} & 0 & 0 & 0 & 0 & \cdots \\
m_{23} & 0 & 0 & m_{123} & m_{223} & m_{233} & 0 & 0 & 0 & m_{1123} & 2\,m_{1223} & 2\,m_{1233} & m_{2223} & 2\,m_{2233} & m_{2333} & 0 & 0 & 0 & 0 & \cdots \\
m_{33} & 0 & 0 & m_{133} & m_{233} & m_{333} & 0 & 0 & 0 & m_{1133} & 2\,m_{1233} & 2\,m_{1333} & m_{2233} & 2\,m_{2333} & m_{3333} & 0 & 0 & 0 & 0 & \cdots \\
m_{111} & 0 & 0 & m_{1111} & m_{1112} & m_{1113} & 0 & 0 & 0 & m_{11111} & 2\,m_{11112} & 2\,m_{11113} & m_{11122} & 2\,m_{11123} & m_{11133} & 0 & 0 & 0 & 0 & \cdots \\
m_{112} & 0 & 0 & m_{1112} & m_{1122} & m_{1123} & 0 & 0 & 0 & m_{11112} & 2\,m_{11122} & 2\,m_{11123} & m_{11222} & 2\,m_{11223} & m_{11233} & 0 & 0 & 0 & 0 & \cdots \\
m_{113} & 0 & 0 & m_{1113} & m_{1123} & m_{1133} & 0 & 0 & 0 & m_{11113} & 2\,m_{11123} & 2\,m_{11133} & m_{11223} & 2\,m_{11233} & m_{11333} & 0 & 0 & 0 & 0 & \cdots \\
m_{122} & 0 & 0 & m_{1122} & m_{1222} & m_{1223} & 0 & 0 & 0 & m_{11122} & 2\,m_{11222} & 2\,m_{11223} & m_{12222} & 2\,m_{12223} & m_{12233} & 0 & 0 & 0 & 0 & \cdots \\
m_{123} & 0 & 0 & m_{1123} & m_{1223} & m_{1233} & 0 & 0 & 0 & m_{11123} & 2\,m_{11223} & 2\,m_{11233} & m_{12223} & 2\,m_{12233} & m_{12333} & 0 & 0 & 0 & 0 & \cdots \\
m_{133} & 0 & 0 & m_{1133} & m_{1233} & m_{1333} & 0 & 0 & 0 & m_{11133} & 2\,m_{11233} & 2\,m_{11333} & m_{12233} & 2\,m_{12333} & m_{13333} & 0 & 0 & 0 & 0 & \cdots \\
m_{222} & 0 & 0 & m_{1222} & m_{2222} & m_{2223} & 0 & 0 & 0 & m_{11222} & 2\,m_{12222} & 2\,m_{12223} & m_{22222} & 2\,m_{22223} & m_{22233} & 0 & 0 & 0 & 0 & \cdots \\
m_{223} & 0 & 0 & m_{1223} & m_{2223} & m_{2233} & 0 & 0 & 0 & m_{11223} & 2\,m_{12223} & 2\,m_{12233} & m_{22223} & 2\,m_{22233} & m_{22333} & 0 & 0 & 0 & 0 & \cdots \\
m_{233} & 0 & 0 & m_{1233} & m_{2233} & m_{2333} & 0 & 0 & 0 & m_{11233} & 2\,m_{12233} & 2\,m_{12333} & m_{22233} & 2\,m_{22333} & m_{23333} & 0 & 0 & 0 & 0 & \cdots \\
m_{333} & 0 & 0 & m_{1333} & m_{2333} & m_{3333} & 0 & 0 & 0 & m_{11333} & 2\,m_{12333} & 2\,m_{13333} & m_{22333} & 2\,m_{23333} & m_{33333} & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{bmatrix}
$$

of equations is built up, as explained in the next section, where the symmetric coefficients of tensors λ_n will not be repeated in the vector of unknowns. The system is solved by applying a least squares method, so that to get the minimum squared error of the fit, it is weighted in terms of the inverse sampling variances σ^2 of the moments up to order $n-2$, in the righthand side of Eq. 1. In addition, a predictor-corrector method is applied to evaluate the variance matrix of the unknowns, as detailed below.
The overdeterminate system of equations, which is equivalent to Eq. 1, is written as

$$\mathbf{y} = \mathbf{g}_2\mathbf{x}, \tag{2}$$

where the only unknowns are the non-identical elements of the symmetric tensors λ_n. It takes the following form:
The first column **A** (Table 2, Cubarsi 2008) of matrix **Y** and the first column **a** of matrix **X** are related by the coefficient submatrix of Table 1. The second column **B** of matrix **Y** and the first column **b** of matrix **X** are related by the coefficient submatrix of Table 2. The third column **C** of matrix **Y** and the third column **c** of matrix **X** are related by the coefficient submatrix of Table 3.

The resulting $\mathbf{g}_2$ matrix is obtained by stacking the three foregoing submatrices. Vector $\mathbf{x}$ now takes the form

$$\begin{aligned}\mathbf{x} = \; & (\lambda_1, \lambda_2, \lambda_3, 2\lambda_{11}, 2\lambda_{12}, 2\lambda_{13}, 2\lambda_{22}, 2\lambda_{23}, 2\lambda_{33}, 3\,\lambda_{111}, 3\,\lambda_{112}, 3\,\lambda_{113}, 3\,\lambda_{122}, 3\,\lambda_{123}, 3\,\lambda_{133}, \\ & 3\,\lambda_{222}, 3\,\lambda_{223}, 3\,\lambda_{233}, 3\,\lambda_{333}, ...)^{\mathrm{T}}\end{aligned} \tag{3}$$

The factors multiplying the elements of tensors λ_n in Table 2, other than those appearing in vector $\mathbf{x}$, have been carried over the elements of matrix $\mathbf{g}_2$.

Table 2: Coefficient submatrix relating the second column $\mathbf{B}$ of matrix $\mathbf{Y}$ and the first column $\mathbf{b}$ of matrix $\mathbf{X}$.

$$\begin{bmatrix}
0 & 1 & 0 & 0 & m_1 & 0 & m_2 & m_3 & 0 & 0 & m_{11} & 0 & 2\,m_{12} & 2\,m_{13} & 0 & m_{22} & 2\,m_{23} & m_{33} & 0 & \cdots \\
0 & m_1 & 0 & 0 & m_{11} & 0 & m_{12} & m_{13} & 0 & 0 & m_{111} & 0 & 2\,m_{112} & 2\,m_{113} & 0 & m_{122} & 2\,m_{123} & m_{133} & 0 & \cdots \\
0 & m_2 & 0 & 0 & m_{12} & 0 & m_{22} & m_{23} & 0 & 0 & m_{112} & 0 & 2\,m_{122} & 2\,m_{123} & 0 & m_{222} & 2\,m_{223} & m_{233} & 0 & \cdots \\
0 & m_3 & 0 & 0 & m_{13} & 0 & m_{23} & m_{33} & 0 & 0 & m_{113} & 0 & 2\,m_{123} & 2\,m_{133} & 0 & m_{223} & 2\,m_{233} & m_{333} & 0 & \cdots \\
0 & m_{11} & 0 & 0 & m_{111} & 0 & m_{112} & m_{113} & 0 & 0 & m_{1111} & 0 & 2\,m_{1112} & 2\,m_{1113} & 0 & m_{1122} & 2\,m_{1123} & m_{1133} & 0 & \cdots \\
0 & m_{12} & 0 & 0 & m_{112} & 0 & m_{122} & m_{123} & 0 & 0 & m_{1112} & 0 & 2\,m_{1122} & 2\,m_{1123} & 0 & m_{1222} & 2\,m_{1223} & m_{1233} & 0 & \cdots \\
0 & m_{13} & 0 & 0 & m_{113} & 0 & m_{123} & m_{133} & 0 & 0 & m_{1113} & 0 & 2\,m_{1123} & 2\,m_{1133} & 0 & m_{1223} & 2\,m_{1233} & m_{1333} & 0 & \cdots \\
0 & m_{22} & 0 & 0 & m_{122} & 0 & m_{222} & m_{223} & 0 & 0 & m_{1122} & 0 & 2\,m_{1222} & 2\,m_{1223} & 0 & m_{2222} & 2\,m_{2223} & m_{2233} & 0 & \cdots \\
0 & m_{23} & 0 & 0 & m_{123} & 0 & m_{223} & m_{233} & 0 & 0 & m_{1123} & 0 & 2\,m_{1223} & 2\,m_{1233} & 0 & m_{2223} & 2\,m_{2233} & m_{2333} & 0 & \cdots \\
0 & m_{33} & 0 & 0 & m_{133} & 0 & m_{233} & m_{333} & 0 & 0 & m_{1133} & 0 & 2\,m_{1233} & 2\,m_{1333} & 0 & m_{2233} & 2\,m_{2333} & m_{3333} & 0 & \cdots \\
0 & m_{111} & 0 & 0 & m_{1111} & 0 & m_{1112} & m_{1113} & 0 & 0 & m_{11111} & 0 & 2\,m_{11112} & 2\,m_{11113} & 0 & m_{11122} & 2\,m_{11123} & m_{11133} & 0 & \cdots \\
0 & m_{112} & 0 & 0 & m_{1112} & 0 & m_{1122} & m_{1123} & 0 & 0 & m_{11112} & 0 & 2\,m_{11122} & 2\,m_{11123} & 0 & m_{11222} & 2\,m_{11223} & m_{11233} & 0 & \cdots \\
0 & m_{113} & 0 & 0 & m_{1113} & 0 & m_{1123} & m_{1133} & 0 & 0 & m_{11113} & 0 & 2\,m_{11123} & 2\,m_{11133} & 0 & m_{11223} & 2\,m_{11233} & m_{11333} & 0 & \cdots \\
0 & m_{122} & 0 & 0 & m_{1122} & 0 & m_{1222} & m_{1223} & 0 & 0 & m_{11122} & 0 & 2\,m_{11222} & 2\,m_{11223} & 0 & m_{12222} & 2\,m_{12223} & m_{12233} & 0 & \cdots \\
0 & m_{123} & 0 & 0 & m_{1123} & 0 & m_{1223} & m_{1233} & 0 & 0 & m_{11123} & 0 & 2\,m_{11223} & 2\,m_{11233} & 0 & m_{12223} & 2\,m_{12233} & m_{12333} & 0 & \cdots \\
0 & m_{133} & 0 & 0 & m_{1133} & 0 & m_{1233} & m_{1333} & 0 & 0 & m_{11133} & 0 & 2\,m_{11233} & 2\,m_{11333} & 0 & m_{12233} & 2\,m_{12333} & m_{13333} & 0 & \cdots \\
0 & m_{222} & 0 & 0 & m_{1222} & 0 & m_{2222} & m_{2223} & 0 & 0 & m_{11222} & 0 & 2\,m_{12222} & 2\,m_{12223} & 0 & m_{22222} & 2\,m_{22223} & m_{22233} & 0 & \cdots \\
0 & m_{223} & 0 & 0 & m_{1223} & 0 & m_{2223} & m_{2233} & 0 & 0 & m_{11223} & 0 & 2\,m_{12223} & 2\,m_{12233} & 0 & m_{22223} & 2\,m_{22233} & m_{22333} & 0 & \cdots \\
0 & m_{233} & 0 & 0 & m_{1233} & 0 & m_{2233} & m_{2333} & 0 & 0 & m_{11233} & 0 & 2\,m_{12233} & 2\,m_{12333} & 0 & m_{22233} & 2\,m_{22333} & m_{23333} & 0 & \cdots \\
0 & m_{333} & 0 & 0 & m_{1333} & 0 & m_{2333} & m_{3333} & 0 & 0 & m_{11333} & 0 & 2\,m_{12333} & 2\,m_{13333} & 0 & m_{22333} & 2\,m_{23333} & m_{33333} & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{bmatrix}$$

Table 3: Coefficient submatrix relating the third column $\mathbf{C}$ of matrix $\mathbf{Y}$ and the third column $\mathbf{c}$ of matrix $\mathbf{X}$.

$$\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & m_1 & 0 & m_2 & m_3 & 0 & 0 & m_{11} & 0 & 2\,m_{12} & 2\,m_{13} & 0 & m_{22} & 2\,m_{23} & m_{33} & \cdots \\
0 & 0 & m_1 & 0 & 0 & m_{11} & 0 & m_{12} & m_{13} & 0 & 0 & m_{111} & 0 & 2\,m_{112} & 2\,m_{113} & 0 & m_{122} & 2\,m_{123} & m_{133} & \cdots \\
0 & 0 & m_2 & 0 & 0 & m_{12} & 0 & m_{22} & m_{23} & 0 & 0 & m_{112} & 0 & 2\,m_{122} & 2\,m_{123} & 0 & m_{222} & 2\,m_{223} & m_{233} & \cdots \\
0 & 0 & m_3 & 0 & 0 & m_{13} & 0 & m_{23} & m_{33} & 0 & 0 & m_{113} & 0 & 2\,m_{123} & 2\,m_{133} & 0 & m_{223} & 2\,m_{233} & m_{333} & \cdots \\
0 & 0 & m_{11} & 0 & 0 & m_{111} & 0 & m_{112} & m_{113} & 0 & 0 & m_{1111} & 0 & 2\,m_{1112} & 2\,m_{1113} & 0 & m_{1122} & 2\,m_{1123} & m_{1133} & \cdots \\
0 & 0 & m_{12} & 0 & 0 & m_{112} & 0 & m_{122} & m_{123} & 0 & 0 & m_{1112} & 0 & 2\,m_{1122} & 2\,m_{1123} & 0 & m_{1222} & 2\,m_{1223} & m_{1233} & \cdots \\
0 & 0 & m_{13} & 0 & 0 & m_{113} & 0 & m_{123} & m_{133} & 0 & 0 & m_{1113} & 0 & 2\,m_{1123} & 2\,m_{1133} & 0 & m_{1223} & 2\,m_{1233} & m_{1333} & \cdots \\
0 & 0 & m_{22} & 0 & 0 & m_{122} & 0 & m_{222} & m_{223} & 0 & 0 & m_{1122} & 0 & 2\,m_{1222} & 2\,m_{1223} & 0 & m_{2222} & 2\,m_{2223} & m_{2233} & \cdots \\
0 & 0 & m_{23} & 0 & 0 & m_{123} & 0 & m_{223} & m_{233} & 0 & 0 & m_{1123} & 0 & 2\,m_{1223} & 2\,m_{1233} & 0 & m_{2223} & 2\,m_{2233} & m_{2333} & \cdots \\
0 & 0 & m_{33} & 0 & 0 & m_{133} & 0 & m_{233} & m_{333} & 0 & 0 & m_{1133} & 0 & 2\,m_{1233} & 2\,m_{1333} & 0 & m_{2233} & 2\,m_{2333} & m_{3333} & \cdots \\
0 & 0 & m_{111} & 0 & 0 & m_{1111} & 0 & m_{1112} & m_{1113} & 0 & 0 & m_{11111} & 0 & 2\,m_{11112} & 2\,m_{11113} & 0 & m_{11122} & 2\,m_{11123} & m_{11133} & \cdots \\
0 & 0 & m_{112} & 0 & 0 & m_{1112} & 0 & m_{1122} & m_{1123} & 0 & 0 & m_{11112} & 0 & 2\,m_{11122} & 2\,m_{11123} & 0 & m_{11222} & 2\,m_{11223} & m_{11233} & \cdots \\
0 & 0 & m_{113} & 0 & 0 & m_{1113} & 0 & m_{1123} & m_{1133} & 0 & 0 & m_{11113} & 0 & 2\,m_{11123} & 2\,m_{11133} & 0 & m_{11223} & 2\,m_{11233} & m_{11333} & \cdots \\
0 & 0 & m_{122} & 0 & 0 & m_{1122} & 0 & m_{1222} & m_{1223} & 0 & 0 & m_{11122} & 0 & 2\,m_{11222} & 2\,m_{11223} & 0 & m_{12222} & 2\,m_{12223} & m_{12233} & \cdots \\
0 & 0 & m_{123} & 0 & 0 & m_{1123} & 0 & m_{1223} & m_{1233} & 0 & 0 & m_{11123} & 0 & 2\,m_{11223} & 2\,m_{11233} & 0 & m_{12223} & 2\,m_{12233} & m_{12333} & \cdots \\
0 & 0 & m_{133} & 0 & 0 & m_{1133} & 0 & m_{1233} & m_{1333} & 0 & 0 & m_{11133} & 0 & 2\,m_{11233} & 2\,m_{11333} & 0 & m_{12233} & 2\,m_{12333} & m_{13333} & \cdots \\
0 & 0 & m_{222} & 0 & 0 & m_{1222} & 0 & m_{2222} & m_{2223} & 0 & 0 & m_{11222} & 0 & 2\,m_{12222} & 2\,m_{12223} & 0 & m_{22222} & 2\,m_{22223} & m_{22233} & \cdots \\
0 & 0 & m_{223} & 0 & 0 & m_{1223} & 0 & m_{2223} & m_{2233} & 0 & 0 & m_{11223} & 0 & 2\,m_{12223} & 2\,m_{12233} & 0 & m_{22223} & 2\,m_{22233} & m_{22333} & \cdots \\
0 & 0 & m_{233} & 0 & 0 & m_{1233} & 0 & m_{2233} & m_{2333} & 0 & 0 & m_{11233} & 0 & 2\,m_{12233} & 2\,m_{12333} & 0 & m_{22233} & 2\,m_{22333} & m_{23333} & \cdots \\
0 & 0 & m_{333} & 0 & 0 & m_{1333} & 0 & m_{2333} & m_{3333} & 0 & 0 & m_{11333} & 0 & 2\,m_{12333} & 2\,m_{13333} & 0 & m_{22333} & 2\,m_{23333} & m_{33333} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{bmatrix}$$

It is well known that, if $\mathbf{V}_y$ is the variance matrix for vector $\mathbf{y}$, which is taken as the diagonal matrix of its sampling variances $\boldsymbol{\sigma_y^2}$, then the least squares system weighted by $\mathbf{V}_y^{-1}$ provides minimum variance estimates for $\mathbf{x}$ according to (e.g. Stuart & Ord 1987)

$$\mathbf{x} = (\mathbf{g}_2^{\mathrm{T}}\mathbf{V}_y^{-1}\mathbf{g}_2)^{-1}\mathbf{g}_2^{\mathrm{T}}\mathbf{V}_y^{-1}\mathbf{y}. \tag{4}$$

The minimum fitting error is then obtained from the weighted norm of the difference between the observed values and their theoretical predictions,

$$\chi^2 = (\mathbf{y}^{\mathrm{T}} - \mathbf{x}^{\mathrm{T}}\mathbf{g}_2^{\mathrm{T}})\mathbf{V}_y^{-1}(\mathbf{y} - \mathbf{g}_2\mathbf{x}). \tag{5}$$

The error on the results of the least squares fit, that is the variance $\boldsymbol{\sigma_x^2}$ of vector $\mathbf{x}$, is obtained from the diagonal matrix of

$$\mathbf{V}_x = (\mathbf{g}_2^{\mathrm{T}}\mathbf{V}_y^{-1}\mathbf{g}_2)^{-1}. \tag{6}$$

Some aspects of the fitting procedure must be pointed out:
(1) Some elements of vector $\mathbf{y}$ are exact values, -1 or 0, thus they have no associated error. However, an initial tolerance error may be assumed for vector $\mathbf{y}$, which can be associated with the finite domain beyond which the density function is negligible, as pointed out in Cubarsi (2008). This tolerance error is assumed to be constant and significantly small for all the components (e.g. 10^{-6}), and it is added to the sampling error σ_y^2 of the data. The final norm of the quadratic error is computed as $f_0 = \|\sigma_y^2\|$.
(2) In the least squares method, it is generally assumed that the elements of matrix $\mathbf{g}_2$ are evaluated from exact values, and does not contribute to the error of the estimates, although it is not true in the current case. To evaluate the part of the total fitting error due to the elements of matrix $\mathbf{g}_2$, an iterative procedure is started by assuming equal uncertainties for all equations, which are normalised to constant norm f_0. Let us call its predicted quadratic error ϵ_I^2.
(3) By starting from the predicted quadratic error, successive evaluations of the variance matrix $\mathbf{V}_y$ from the error propagation formula

$$\mathbf{V}_y = \mathbf{g}_2\mathbf{V}_x\,\mathbf{g}_2^{\mathrm{T}} \tag{7}$$

are carried out, so that we obtain a corrected quadratic error ϵ_F^2, which is used as a new predicted error, by normalising it to a constant norm f_0.
(4) The algorithm is stopped when a fixed point ϵ_0^2 is reached, that is, when ϵ_I^2 and ϵ_F^2 have the same direction. It is found that the final quadratic error ϵ_0^2 of the iterative process does not depend on the initial predicted error, so that the final variance matrix $\mathbf{V}_y$ is the result of a redistribution of weights provided by the matrix $\mathbf{g}_2$ of the least squares system.
(5) To compute the final fitting error χ^2, a total sampling variance $\varepsilon^2 = \epsilon_0^2 + \sigma_y^2$ is assumed as the sum of both independent quadratic errors. The fitting error χ^2 is only partially significant, since it is related to the initial value f_0, which depends on the initial errors of the data $\mathbf{y}$, although this way it is possible to compare the goodness of different fittings.

3 Large-scale structure

Several illustrations of the current functional approach are used to describe the main kinematical features of the solar neighbourhood. The first two cases give the whole velocity distribution of the local disc, which is usually fitted by a mixture of trivariate Gaussian distributions. In the first application, a nearly complete and kinematically representative local sample, Sample I with 13,678

stars, is used. It was obtained by crossing the HIPPARCOS Catalogue (ESA 1997) with radial velocities from the HIPPARCOS Input Catalogue (ESA 1992). To get a representative sample of the solar neighbourhood, it was limited to a trigonometric distance of 300 pc, where the only input data points were the velocity components (U, V, W) in a cartesian heliocentric coordinate system, with U toward the Galactic centre, V in the rotational direction, and W perpendicular to Galactic plane, positive in the direction of the North Galactic pole. In Paper I it was found that the optimal subsample containing thin and thick disc stars could be obtained by selecting stars with absolute heliocentric velocity $|\mathbf{V}| \leq 210$ km s^{-1}. The sample has undergone a deeper statistical analysis in Cubarsi et al. (2009), hereafter Paper II, where the subsamples selected from $|\mathbf{V}| > 400$ km s^{-1} contain, in addition to above disc populations, a fraction up to 1% of halo stars, with very stable computed moments. To compare results with the next sample, the velocity domain is limited to the absolute space velocity of 500 km s^{-1}, which only excludes five stars from the whole sample. The resulting Sample I is then composed of 13,673 stars. The distribution smoothly vanishes in reaching the velocity boundary, as shown in the last row of Fig. 1. For all practical purposes, the sample may be considered as unbounded. To compare between different fittings, Eq. 1 up to $n = 6$ is used, by taking up to tenth moments into account. According to §2, by normalising to the number N of equations, the squared error of the fit is then given by the expression

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\varepsilon_i^2} \, |y_i - g_{ij} x_j|^2 \, . \tag{8}$$

The maximum entropy procedure with $n = 2$ tries to represent the whole distribution from an unique ellipsoidal distribution. Thus, odd-order moments and even-order moments higher than four are not fitted. The resulting fitting error $\chi^2 = 83.93$ is not acceptable. Let us point out, however, that more than the value χ^2 itself, which is more significant is the increase or decrease in this quantity, since it may become distorted because over or underestimating observational errors, or undesired error distribution. The approach with $n = 4$ by using up to sixth moments gives a clearly improved result $\chi^2 = 1.05$. A symmetric distribution around the plane $W = 0$ is also quite admissible, with $\chi^2 = 2.85$, which is the same order as the previous error. The approach with $n = 6$ is also computed by fitting up to the tenth moments, which gives a very accurate fit, $\chi^2 = 0.02$, even with symmetry around the plane $W = 0$, $\chi^2 = 0.62$. The contour plots of the velocity distribution on each velocity plane are displayed in Fig. 1, as well as the bell-shaped sections of the density function in each velocity component. The coordinate system is centred in the mean velocity of the sample $(-10.83, -20.47, -7.32)$ km s^{-1}, referring to the Sun.

In a second example, the method is applied to Sample II, drawn from the GCS catalogue (Nordtröm et al. 2004, Holmberg et al. 2007). It has new and more accurate radial velocity data than the HIPPARCOS sample and contains the total velocity space of 13,240 F and G dwarf stars, which are considered the favourite tracer populations of the history of the disc. According to the authors, the main essential features of the sample are the lack of kinematic selection bias and the radial velocity data, which allowed to reject stars that have not taken part in the evolution of the local disc. The same cartesian heliocentric coordinate system is used. For this sample, according to the analysis in Paper II, the moments are computationally stable for all the velocity components in the range $400 \leq |\mathbf{V}| \leq 500$ km s^{-1}. The limitation up to an absolute velocity space of 500 km s^{-1} excludes five stars. The halo component is present in the total sample in a fraction less than 0.5 %. Therefore, for practical purposes, this sample may also be considered as unbounded. The results and the graphs are similar to those of Sample I.

In the next section we confirm that the GCS sample provides more accurate moments of the disc velocity distribution than HIPPARCOS' due to its more precise radial velocities. In Table

4, centred and non-centred velocity moments up to order four are listed for the GCS Sample II, along with their standard errors. These are moments for a mixture of thin disc (94%), tick disc (5.5%), and halo (0.5%), as discussed in Paper II. They allow some measures of spread and asymmetry of the distribution in the desired variables to be computed, as the non null skewness in the rotation velocity V, $\gamma_V = \mu_{030}\,{\mu_{020}}^{-\frac{3}{2}} = -3 \pm 0.5$ (in the Greek indices notation), which is zero in the other components. The moments also lead to a non-significant curtosis in the vertical velocity W, $c_W = \mu_{004}\,{\mu_{002}}^{-2} - 3 = 34 \pm 40$, or the non-vanishing vertex deviation on the UV plane $\delta = \frac{1}{2}\arctan[2\mu_{110}\,(\mu_{200} - \mu_{020})^{-1}] = 11.3^\circ \pm 2.3$.
Therefore, the main features of the maximum entropy distribution for Samples I and II, which show a reasonable deviation from an ellipsoidal distribution, may be easily deduced from Fig. 1, either for $n = 4$ or $n = 6$. The main features are:

(i) The velocity distribution is not symmetric around the mean, mainly in the rotation direction.

(ii) The whole distribution has a clear vertex deviation on the plane UV and no deviation on other planes.

(iii) There is some skewness in the variable V. As a consequence of both previous situations there is a wider distribution wing towards lower U and V velocities, which is likely caused by thick disc stars.

(iv) The curtosis in the W variable vanishes and is zero or very small in the U velocity (see Table 6).

(v) The plane $W = 0$ is basically a symmetry plane.

The resulting density function, according to the most significant polynomial coefficients, can be expressed as a product of two exponential functions in the form

$$f = \varphi_1(Q)\,\varphi_2(h) \tag{9}$$

where Q is a quadratic negative-definite form, which gives the background ellipsoidal shape of the distribution, with axis ratios 1:0.7:0.5, symmetry plane $W = 0$, as expected for disc stellar samples, and overall vertex deviation in the UV velocity components of about 12º. The function $\varphi_2(h)$ can be expressed in terms of the angular momentum integral h, and may be interpreted as a perturbation factor. It is even and is at least quadratic in the V velocity alone, which accounts for the skewness and the shift in the ellipsoidal isocontours in terms of the rotation velocity.
The bulk of the local velocity distribution does not show any substructure reflecting the existing moving groups, even by associating these moving groups with different proxy Gaussian components (Bovy et al. 2009). Then, it results in a smooth background distribution. However, by selecting specific subsamples by colour, or by using different analysis techniques where the resolution scale may vary, the substructures of the velocity distribution arise. We discuss it in the next section.

4 Truncated distributions

The next two examples are used for two new purposes: first, to test the ability of the maximum entropy method in reconstructing a truncated velocity distribution associated with a velocity bounded sample; second, to try a magnifying glass effect over the distribution and to focus on a specific velocity domain. According to Paper I, the selection of local stars with an absolute

Table 4: Centred moments $\mu_{\alpha\beta\gamma}$ and non-centred moments $m_{\alpha\beta\gamma}$ with their standard errors up to fourth order for the GCS Sample II.

$\alpha\beta\gamma$	$\mu_{\alpha\beta\gamma}$	$\Delta\mu_{\alpha\beta\gamma}$	$m_{\alpha\beta\gamma}$	$\Delta m_{\alpha\beta\gamma}$
100	0.00	± 0.30	-9.97	± 0.30
010	0.00	0.22	-18.59	0.22
001	0.00	0.16	-7.16	0.16
200	1205.49	± 30.08	1304.84	± 30.94
110	114.33	20.02	299.63	22.96
020	657.33	28.67	1002.95	34.28
101	-28.26	11.81	43.09	12.32
011	14.09	10.97	147.17	12.65
002	332.93	17.33	384.17	17.82
300	-5410.36	± 6929.41	-42446.74	± 7693.93
210	-29820.51	4514.01	-56357.61	5178.37
120	-1883.35	3918.62	-16130.85	4722.18
030	-50726.61	6922.64	-93813.02	8810.45
201	-1529.39	2007.04	-10306.27	2143.30
111	1567.29	1482.29	-192.55	1670.70
021	-1486.53	2027.51	-9189.67	2515.32
102	-2581.18	2450.85	-6005.75	2637.40
012	-9440.44	1724.77	-16784.25	2090.32
003	-5661.87	5733.57	-13178.21	6024.02
400	13430320.65	± 2125166.40	14374454.53	± 2291035.07
310	1460621.61	1270588.16	3175498.30	1469075.82
220	5317165.76	948783.69	7064506.56	1153512.90
130	322235.30	930673.62	1480869.49	1171917.48
040	11308141.98	1962964.12	16562919.85	2562819.68
301	-1192280.95	478081.09	-851132.23	485712.04
211	62741.79	322480.47	454274.09	354935.98
121	-315560.92	289850.40	-248097.44	344247.93
031	234834.16	517841.17	1003877.53	647019.02
202	1845208.39	421200.10	2010427.12	459870.47
112	315298.09	262885.04	506475.65	323652.27
022	1593831.34	251847.69	2140083.38	327148.87
103	139487.68	845660.69	321923.19	926678.48
013	492111.63	552721.46	941999.67	688288.39
004	4084082.40	2190306.74	4351176.92	2325987.63

value of the total space motion $|\mathbf{V}| \leq 51$ km s^{-1} had left the older disc stars aside, which are the originators of an important softening of the distribution. Such a selected group of stars contained a complex mixture of early-type and young disc stars for which a Gaussian mixture approach was unreliable because of the large fitting errors. This small-scale structure of the velocity distribution was strongly asymmetric in comparison to the background distribution. It was also observed in other analyses of the solar neighbourhood (e.g. Famaey et al. 2005, Soubiran & Girard 2005).

Sample III is then composed of 10,195 stars from the HIPPARCOS Sample I, with $|\mathbf{V}| \leq 51$ km s^{-1}. The maximum entropy approach for $n = 2$ gives a fitting error $\chi^2 = 0.66$, according to Eq. 8. Although it could seem a very low value compared to previous samples, we might bear in mind that Samples I and II contain stars with higher velocity than Sample III, which increases the uncertainty of the computed moments. Because of this, the fitting errors for Sample III are expected to be much smaller. Once again, we must pay attention to the variation in χ^2.

For $n = 4$, the approach is able to provide a more realistic, non-ellipsoidal map of the truncated distribution by fitting moments up to sixth order. In this case the fitting error is $\chi^2 = 0.12$. Nevertheless, for $n = 6$, the maximum entropy approach gives a much improved portrait by fitting up to tenth moments. The fitting error $\chi^2 = 0.008$ is about 10^2 times lower than the ellipsoidal approach. The contour plots of the velocity distribution on each velocity plane are displayed in Fig. 2. The coordinate system is centred in the heliocentric mean velocity $(-7.49, -11.25, -6.41)$ km s^{-1}.

Finally, Sample IV is drawn from the GCS catalogue by selecting 9,733 stars with absolute velocity lower than 51 km s^{-1}. As in the above example, the approaches with $n = 2$ ($\chi^2 = 1.09$) and $n = 4$ ($\chi^2 = 0.29$) are not able to provide a realistic map of the truncated distribution. However, for $n = 6$, with a fitting error $\chi^2 = 0.002$, more than 10^2 lower than the case $n = 4$, the maximum entropy approach gives a detailed portrait of actual asymmetries, in particular on the UV plane.

The contour plots of the velocity distribution on each velocity plane are displayed in Fig. 3. The coordinate system is centred in the mean heliocentric velocity $(-6.12, -11.23, -6.18)$ km s^{-1}. In Table 5, centred and non-centred velocity moments up to fourth order are listed as well with their standard errors. The skewness in the rotation velocity V is small, $\gamma_V = 0.23 \pm 0.03$, but non-zero, being similar in the U direction. The curtosis in the vertical velocity W, $c_W = 0.7 \pm 0.3$, is also very low. The vertex deviation on the UV plane is $\delta = 8.7^\circ \pm 0.6$. Although it is caused by both subjacent structures, it is nearly the same as the one obtained in Paper II for the thin disc component. The results are summarised in Table 6.

The improvement of the GCS catalogue over the HIPPARCOS catalogue, mainly for the bounded sample, provides a higher resolution contour plot of the velocity distribution, which in addition to describing a velocity distribution far from the ellipsoidal hypothesis, shows a clear bimodal structure, as displayed in Fig. 4.

The results are consistent with the contour plots obtained by Dehnen (1998) when inferring the velocity distribution of his total sample (AL), in particular for the innermost dark contour. Also, the shape of the velocity distribution for early-type stars (Skuljan et al. 1999) is similar to ours, which is now derived only from velocity moments. By using the GCS catalogue, Famaey et al. (2007) describe a similar small-scale structure of local stars; however, the entropy approach provides the smoothest density function that is also consistent with the data. In Figs. 3 and 4, two regions with higher probability densities are clearly identified, even using a large sample containing most of the thin disc with 9,733 stars. The highest peak is placed around the Hyades-Pleiades moving groups, and the lower peak around the Sirius-UMa stream. However, our method works in the opposite direction of methods based on an arbitrary number of mixture components, or on wavelet transforms on arbitrary smaller scales. As Bovy et al. (2009) point out, adding a new component could substantially increase the goodness of the fit over the model with less complexity, while still being far from the truth. Similarly, Dehnen (1998) points out that structures on scales of a few km s^{-1} are likely to be spurious. On the contrary, the maximum entropy approach is a technique for computing the simplest and smoothest approach to the distribution function that fulfils the provided set of moment constraints. For a good estimation, the only requirement is that the sample is bell-shaped enough and the moments have enough accuracy. The method tends to smoothing all the statistical fluctuations of the sample, since the moments are obviously means. However, as shown in the above examples and in the next sections, if more complexity or resolution is desirable, either a larger set of constraints must be taken into account or specific subsamples must be selected.

5 Analysis of samples

In the preceding sections, the method has been applied to four case examples to show how the fitting of the distribution function is getting more informative depending on the degree of the polynomial $\mathcal{P}_n$ and on the complexity of the sample. We now discuss some aspects of the results and samples. Samples I and II were chosen because they contain the maximum number of available stars with known velocity space, so that the velocity moments have minimum sampling variances. The main goal was to build the largest samples with stable velocity moments. However, these samples contain data with great uncertainty that could hamper the fitting of the distribution function.

The main source of data error, a matter of consequence for the HIPPARCOS sample, is the radial velocity, which is mostly measured from high proper motion stars. This may introduce a kinematical bias into the sample (Binney et al. 1997), although Skuljan et al. (1999) proved that

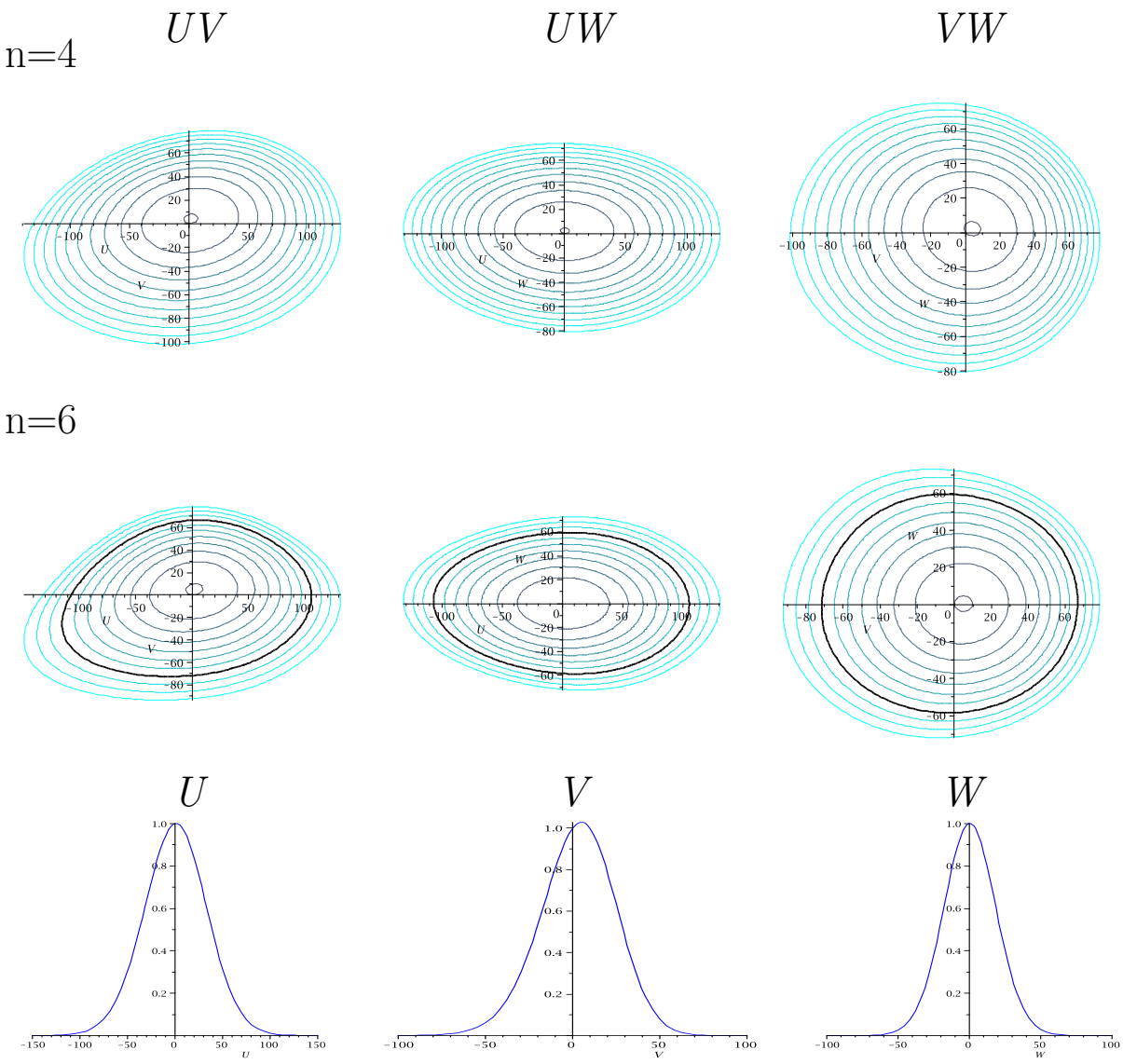

Figure 1: Contour plots of the local velocity distribution in terms of the peculiar velocities for HIPPARCOS' Sample I and Sample I'. The plots are centred on the mean heliocentric velocity $(-10.85, -19.93, -7.49)$ km s^{-1} of Sample I', with radial velocity errors up to 2.5 km s^{-1}. The case n=4, by fitting up to sixth moments, leads to more realistic contour plots than a pure ellipsoidal distribution (n=2), although n=6 provides a slightly improvement, by fitting up to tenth moments. The contours indicate levels $(\frac{1}{2})^k; k = 0, ..., 10$, and the black contour line corresponds to an approximate level 10^{-2} surrounding nearly the whole distribution, as confirmed in the last row, also for n=6, from the sections of the velocity density function (not normalised) in each velocity component.

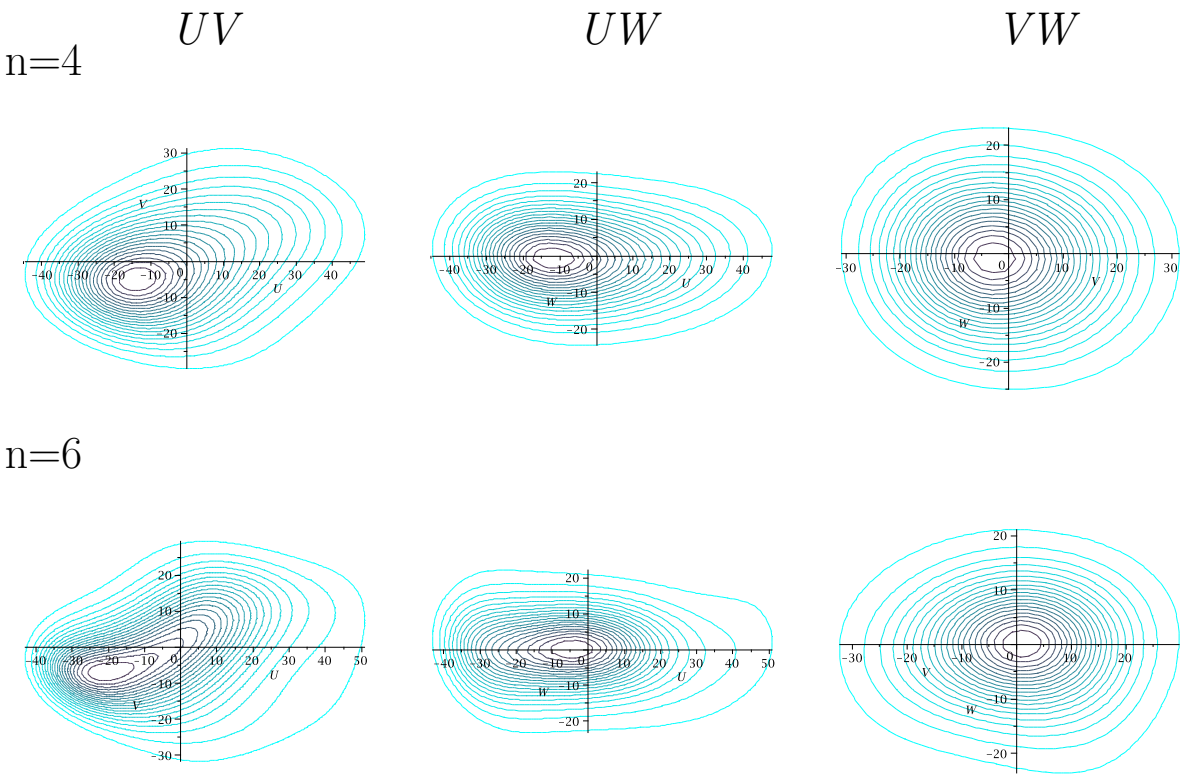

Figure 2: Contour plots of the velocity distribution for Sample III, from the HIPPARCOS catalogue, for stars with $|V| \leq 51$ km s^{-1}. The peculiar velocities are centred on the heliocentric mean velocity $(-7.49, -11.25, -6.41)$ km s^{-1}. The approach n=4 uses moments up to sixth order, and n=6 up to tenth order. The asymmetry of the velocity distribution, mainly on the UV plane, may be sufficiently described in the case n=6.

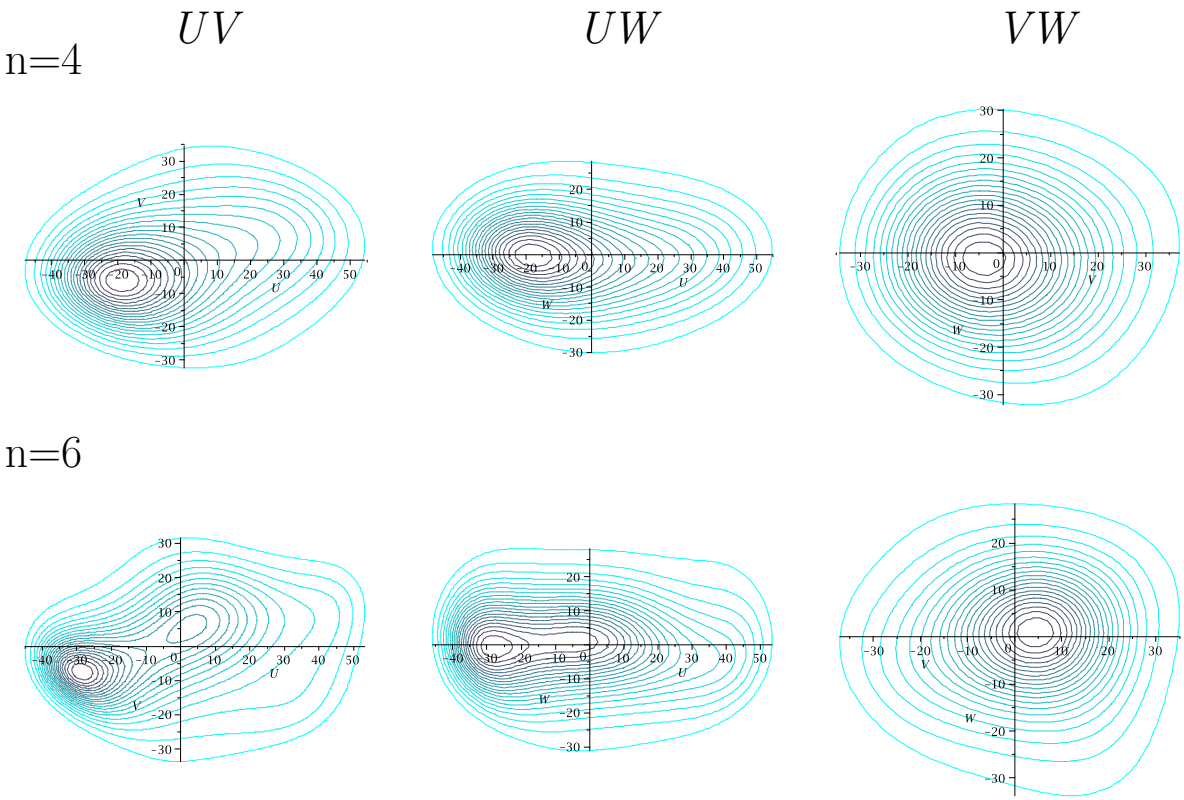

Figure 3: Contour plots of the velocity distribution for Sample IV, from the GCS catalogue, with $|V| \leq 51$ km s^{-1}. The peculiar velocities are centred on the heliocentric mean velocity $(-6.12, -11.23, -6.18)$ km s^{-1}. The strong asymmetry of the velocity distribution, mainly on the UV plane, may only be described in the case n=6, by fitting moments up to tenth order.

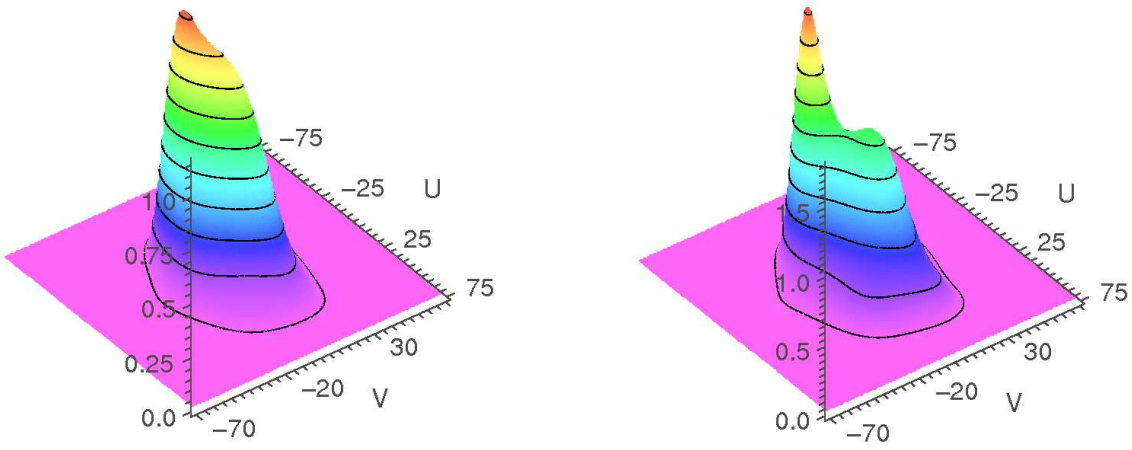

Figure 4: Density functions on the plane UV for the HIPPARCOS Sample III (left) and the GCS Sample IV (right). The plots show a bimodal structure around the Hyades stream (highest peak) and the Sirius-UMa stream (lowest peak) for a distribution far from the ellipsoidal hypothesis.

Table 5: Centred and non-centred moments with their standard errors up to fourth order for the GCS Sample IV.

$\alpha\beta\gamma$	$\mu_{\alpha\beta\gamma}$	$\Delta\mu_{\alpha\beta\gamma}$	$m_{\alpha\beta\gamma}$	$\Delta m_{\alpha\beta\gamma}$
100	0.00	± 0.22	-6.12	± 0.22
010	0.00	0.14	-11.23	0.14
001	0.00	0.13	-6.18	0.13
200	462.82	± 5.41	500.27	5.56
110	41.10	2.75	109.85	3.82
020	199.40	2.87	325.63	3.95
101	-5.00	2.53	32.82	2.81
011	-1.72	1.90	67.71	2.35
002	153.41	2.55	191.61	2.92
300	2173.59	± 157.78	-6551.81	± 276.03
210	-145.28	70.79	-6268.73	109.93
120	530.94	53.23	-2385.18	97.37
030	638.37	79.37	-7500.56	158.75
201	330.59	20.62	-2700.19	82.60
111	131.81	34.76	-480.42	47.48
021	96.46	40.55	-1877.37	64.93
102	337.15	46.93	-773.54	62.06
012	308.20	37.83	-1823.19	53.07
003	142.89	65.05	-2937.69	110.67
400	499489.15	± 10136.14	551669.32	± 9932.67
310	34551.28	3316.06	115444.19	4479.28
220	75113.83	1635.32	153794.57	2498.65
130	6595.67	1901.12	50161.86	3041.40
040	119660.84	4596.79	257920.40	6223.25
301	-1064.18	2965.11	32799.70	3070.77
211	-299.90	1022.52	32365.54	1356.19
121	-102.22	838.49	10219.59	1199.62
031	3803.90	1762.48	46257.99	2182.27
202	62272.91	1472.05	78158.49	1552.90
112	2098.65	788.56	8713.72	991.04
022	35099.21	1149.57	58305.32	1340.04
103	1025.32	1640.64	12177.32	1815.65
013	-1390.25	1394.34	25702.11	1649.39
004	86642.60	3447.49	119730.52	4100.98

the kinematic bias does not significantly affect the core of the disc distribution. Therefore, the description of disc kinematics from star velocities lower than $|\mathbf{V}| \approx 75$ km s^{-1} should not reflect such a bias. To see how the error in the radial velocity could change the shape of the distribution function, and in particular the computed velocity moments, we select some new samples (Sample I' and Sample II') with radial velocity errors up to 2.5 km s^{-1}, a similar value to the mean observational error (Figueras et al. 1997). Sample I' from the HIPPARCOS catalogue now contains 9,534 stars (70% of Sample I) and Sample II' from the GCS catalogue contains 11,514 stars (87% of Sample II). Clearly the GCS catalogue has stars with more accurate radial velocities. The velocity moments of Sample I' correspond now to a colder sample, with similar standard errors despite the small size of the sample. The diagonal second central moments are $(1310.09 \pm 45.20, 951.40 \pm 60.85, 345.39 \pm 16.85)$ instead of $(1431.46 \pm 45.23, 1073.89 \pm 54.95, 372.73 \pm 16.25)$ of Sample I. The velocity moments of Sample II' are not significantly changed and also have similar standard errors. Now, the diagonal second central moments are $(1236.14 \pm 31.60, 681.60 \pm 31.59, 344.70 \pm 19.60)$ compared to $(1205.49 \pm 30.08, 657.33 \pm 28.67, 332.93 \pm 17.33)$ of Sample II. The same criterion is applied to the bounded samples. Sample III' is 71% of Sample III, while Sample IV' is 86% of Sample IV. In these cases the velocity moments and their standard errors do not change at all. Although the respective fractions are similar to those in the previous samples, the moments remain stable, which confirms that the kinematic bias is associated with higher velocity stars. In all the cases a similar shape of the velocity distribution is obtained as well as a slightly improvement of the χ^2 fitting error (in particular for the complete Samples I' and II' with $n = 4$).

Another issue to clarify is the cut $|\mathbf{V}| \leq 51$ km s^{-1} for the bounded Samples III and IV. The main reason to choose them is the discontinuity noticed in Paper I, which has also been borne out in Paper II for the GCS sample. The recurrent segregation method used in these works (MEMPHIS algorithm) analysed the variations of two parameters accounting for a mixture approach: the

entropy of the partition and the fitting error. By increasing a sampling parameter, in that case the absolute star velocity, the first significative discontinuity of those parameters took place at 51 km s^{-1}. After this value the method was able to segregate thin and thick discs with a decreasing fitting error. Therefore, this is not an astronomical reason but a statistical fact. We can now investigate the astronomical facts. Since the GCS samples have more accurate velocities, the analysis is centred in this catalogue. In Fig. 5 the graphs show how the stars are distributed into populations in terms of absolute velocity, eccentricity, metallicity, and colour. According to Paper II, the three bands of the vertical axis represent the expected value of a star to belong to any Galactic component (thin disc in the bottom, thick disc in the middle and halo at the top). Except for the eccentricity plot, the blue dots correspond to stars with $|\mathbf{V}| \leq 51$ km s^{-1}, which clearly belong to the thin disc. This is also true for eccentricities, but now the blue dots correspond to stars with $|z_{max}| \leq 0.5$ kpc. In the $|\mathbf{V}|$ plot we see that a large fraction of thin disc stars are still beyond 51 km s^{-1}. They are mixed with thick disc stars, especially from 65 km s^{-1} onward. From the eccentricity plot we deduce that thin disc stars are below 0.3, as discussed in Paper II, but beyond 0.1 they are increasingly mixed with the thick disc. However, when the condition $|z_{max}| \leq 0.5$ is applied, no thick disc or halo stars are included for eccentricities below ≈ 0.3. On the other hand, it is well known that the metallicity is appropriate for distinguishing the halo from the disc, [Fe/H] ≤ -0.5, but not between thin and thick discs. For the Strömgren photometry, the $b-y$ colour is spread along the three main populations. Most of the thin disc stars of the sample may be found at any index between 0.2 and 0.6, with mode 0.3, while the thick disc has mode 0.4 and the halo 0.5, with slightly narrower distributions. Similarly, for the maximum height over the Galactic plane, $|z_{max}|$ (not shown), thin and thick disc stars are also mixed in the interval $|z_{max}| < 1$ kpc, but, like metallicity, the halo can be segregated.
In Fig. 6 the distributions obtained from the current method with $n = 6$ are plotted in terms of metallicity and colour on the three velocity planes for subsamples with $|\mathbf{V}| \leq 51$ km s^{-1}. All of them reproduce the bimodal structure of Fig. 3 with $n = 6$, except metalicities in the range $0 \leq$ [Fe/H] < 0.5 and colours with $0.205 \leq b-y < 0.300$, which correspond to the earliest F dwarfs of the thin disc, with negative radial mean velocity. However, such a bimodal shape comes from the velocity cut. For the whole CGS sample, the distributions for different metalicities and colours are similar to the deformed velocity ellipsoid of the whole sample, as shown in Fig. 7, though with a slightly different mean, depending on the colour and metallicity range.

6 Conclusions

The resulting features of the local velocity distribution are similar for the complete HIPPARCOS Sample I and GCS Sample II, although, as expected, the latter provides more accurate results due to the lower uncertainties of radial velocity measurements. The whole distribution shows a nearly constant vertex deviation in the pseudo ellipsoidal level curves, as well as a nearly constant axis ratio. According to the most significant polynomial coefficients, the resulting density function can be expressed as the product of two exponential functions in the form of Eq. 9. The background ellipsoidal shape has axis ratios 1:0.7:0.5 and a symmetry plane $W = 0$. The overall vertex deviation in the UV velocity components is approximately 11°. Some characteristic parameters of the distribution are summarised in Table 6. The function $\varphi_2(h)$ may be interpreted as a perturbation factor, and is even and at least quadratic in the V velocity alone. It accounts for the skewness and the shift in the velocity ellipsoids in terms of the rotation velocity. This is clearly visible on the UV and VW planes of Fig. 1, and may be interpreted with regard to the heating of disc stars, which is also correlated with a decreasing galactocentric rotation velocity, as expected

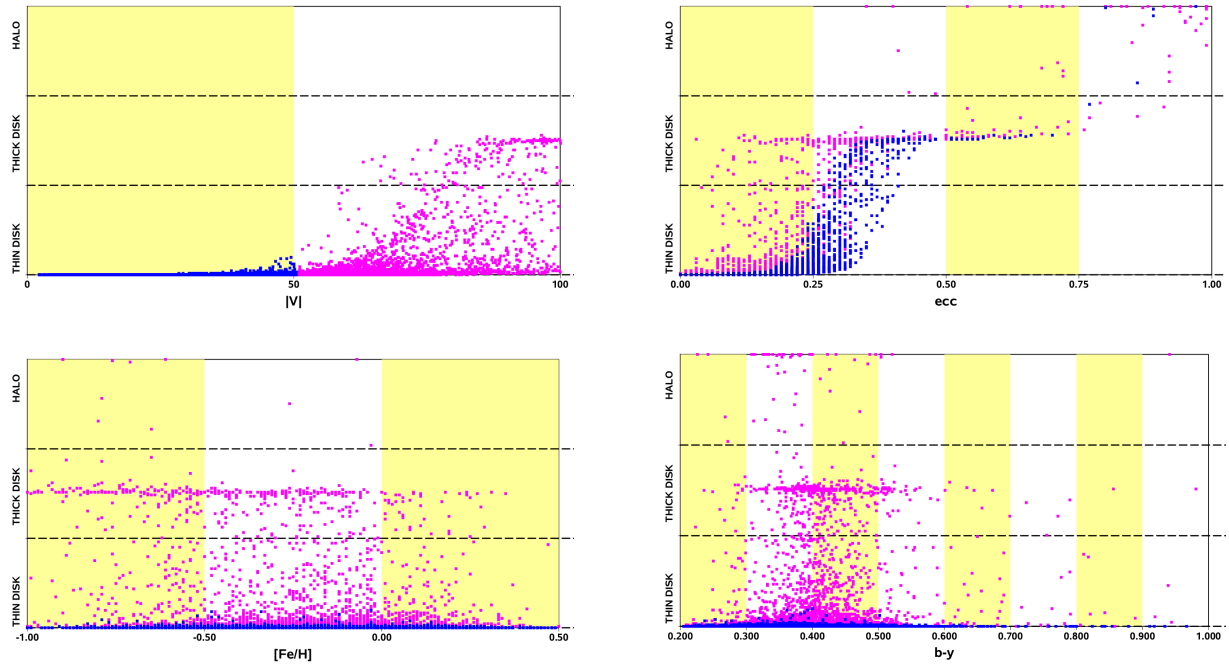

Figure 5: Distribution of GCS sample stars into populations in terms of absolute velocity, eccentricity, metallicity, and colour. The blue dots indicate stars with $|\mathbf{V}| \leq 51$ km s^{-1}, except for the eccentricity plot, which are for stars with $|z_{\max}| \leq 0.5$ kpc.

Table 6: Distribution parameters for HIPPARCOS and GCS samples with radial velocity errors up to 2.5 km s^{-1} (Samples I', II', III', and IV'). The displayed parameters are dispersions σ_U, σ_V, σ_W, vertex deviation δ on the UV-plane, curtosis c_U, c_V, c_W, and skewness γ_U, γ_V, γ_W.

Sample	σ_U	σ_V	σ_W	δ [°]	c_U	c_V	c_W	γ_U	γ_V	γ_W
HIP total	36.2 ±0.6	30.8±1.0	18.6±0.5	7.6±6.6	9.3 ±4.0	37.0 ±13.9	20.7 ±14.8	0.2 ±0.2	-4.5 ±0.8	-0.9 ±0.5
$\|\mathbf{V}\| \leq 51$	19.9 ±0.1	13.2 ±0.1	10.7 ±0.1	12.3 ±0.8	-0.4 ±0.1	0.1 ±0.3	1.5 ±0.5	0.3 ±0.0	0.2 ±0.0	0.2 ±0.1
GCS total	35.2 ±0.4	26.1 ±0.6	18.6 ±0.5	10.7 ±2.4	5.5 ±2.8	22.7 ±9.3	35.2 ±42.4	-0.1 ±0.2	-3.1 ±0.5	-0.9 ±1.0
$\|\mathbf{V}\| \leq 51$	21.5 ±0.1	14.1 ±0.1	12.4 ±0.1	8.7 ±0.6	-0.7 ±0.1	0.0 ±0.2	0.7 ±0.3	0.2 ±0.0	0.2 ±0.0	0.1 ±0.0

from Strömberg's law. The resulting overall distribution has zero curtosis in the W velocity, and, within 2σ level, in the U component for the more accurate GCS sample.

On the other hand, the entropy method offers an excellent estimation of the truncated velocity distributions of Samples III and IV, which only contain thin disc stars. For these subsamples, a Gaussian mixture approach was impossible in Paper I. This method can therefore be used as an alternative way to study multimodal distributions. For star velocities $|\mathbf{V}| \leq 51$ km s^{-1}, the tentative mixture (with a very large chi-squared fitting error) obtained in Paper I suggested a superposition of two enlarged pseudo-ellipsoidal distributions, mainly along the radial direction, with very overlapped wings, and a separation of 28 ± 9 km s^{-1} between means. For Sample IV, the separation of the two peaks (the "U-anomaly") along the radial direction may now be straightforwardly estimated in $30 - 35$ km s^{-1} from the contour plots obtained from the moment constraints. Figures 2 and 3 show a strongly asymmetric velocity dispersion on the plane UV, and nearly laminar isocontours of the W velocity component along the radial direction. There is a large radial velocity dispersion on the UV plane, in the direction of the gravitational gradient, and a very small dispersion in the direction perpendicular to the galactic plane. On the VW plane, the isotropy is slightly recovered. However, the higher resolution contour plots obtained from Sample IV allow more detailed analysis. A first look at the UW plane of Fig. 3 (n=6) shows

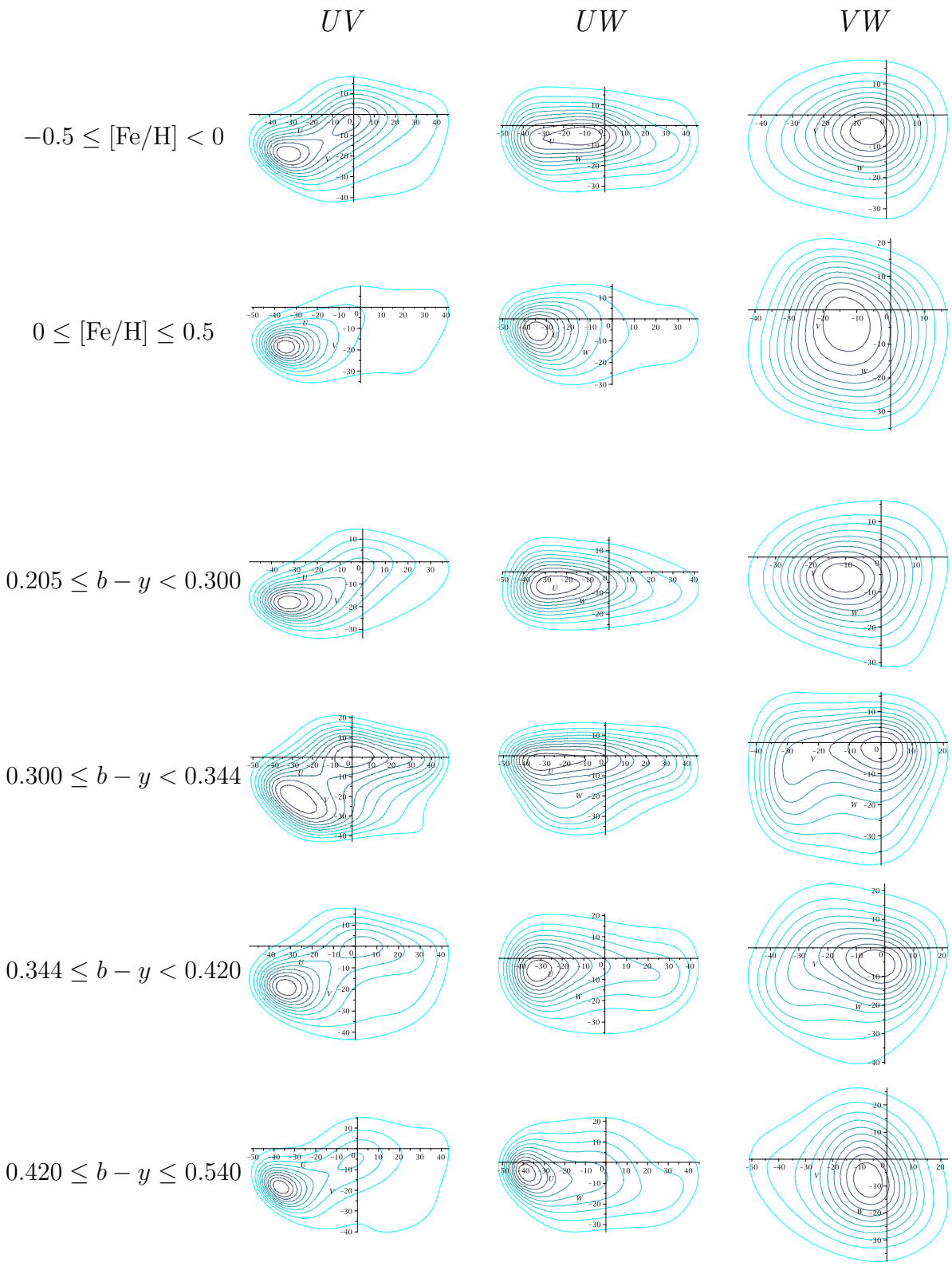

Figure 6: Contour plots of the velocity distribution for stars from the GCS catalogue with $|\mathbf{V}| \leq 51$ km s^{-1}, obtained for $n = 6$ in terms of metallicity and colour. The axes are labelled according to heliocentric velocities.

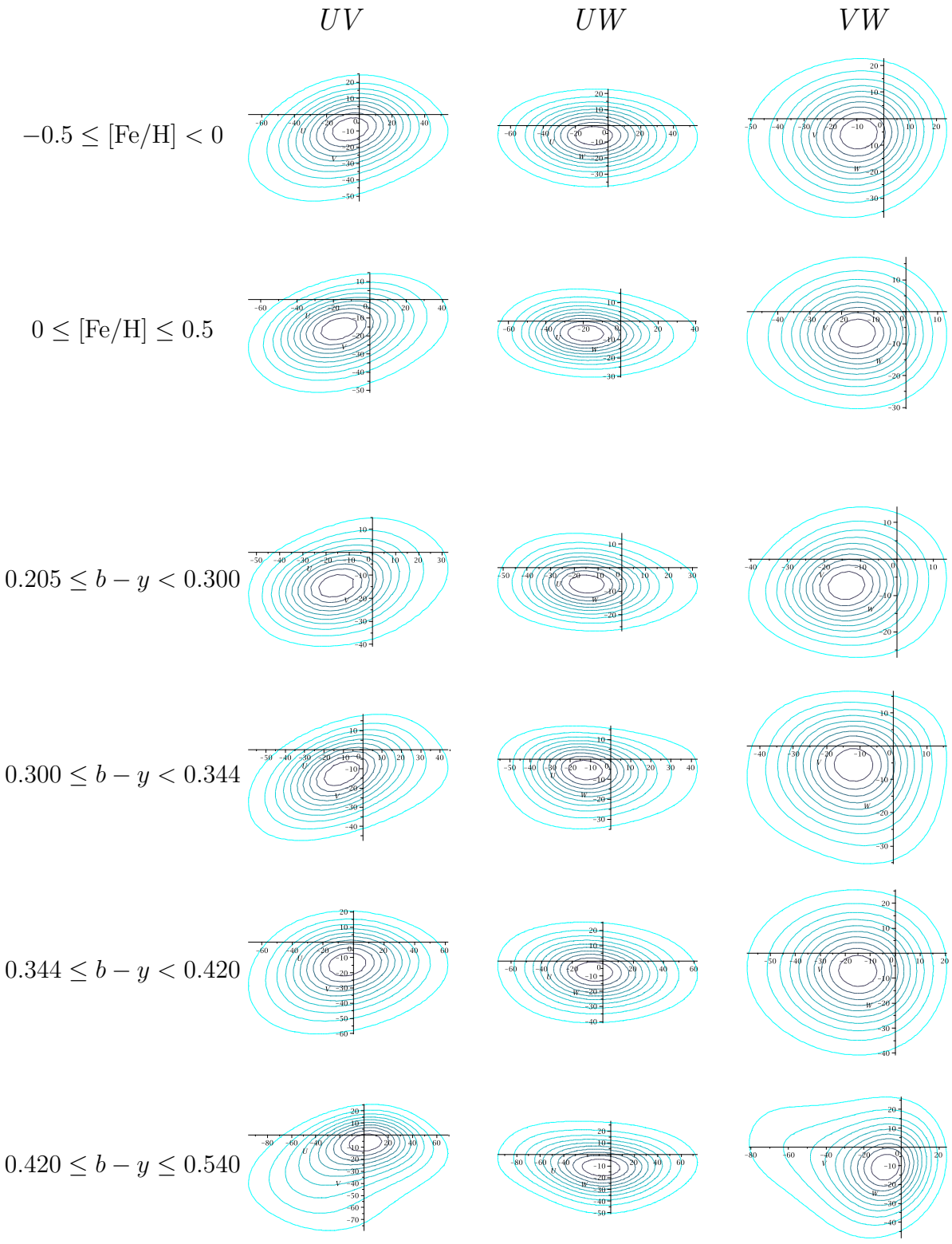

Figure 7: Contour plots of the velocity distribution for stars from the total GCS catalogue, obtained from the entropy approach with $n = 6$, in terms of metallicity and colour.

a core distribution with negative radial peculiar velocity, while there is a clear unimodal behaviour with positive peculiar rotation velocity on the VW plane. However, those high-density regions of the distribution are not simultaneous on the UV plane, but are associated with different large stellar groups.

References

Alcobé S., Cubarsi R. 2005, A&A 442, 929

Binney J., Dehnen W., Houk N., Murray C.A., Penston M.J. 1997, in Hipparcos- Venice'97, ESA SP-402, 473

Bovy J., Hogg D.W., Roweis S.T. 2009, ApJ, 700, 1794

Cercignani C. 1988, in The Boltzmann Equation and its Applications (Berlin: Springer-Verlag)

Cubarsi R., Alcobé S. 2004, A&A 427, 131

Cubarsi R. 2008, On the maximum entropy and the problem of moments: an application to stellar kinematics. Universitat Politecnica de Catalunya, Barcelona. E-prints http://hdl.handle.net/2117/3046

Cubarsi R., Alcobé S., Vidojević S., Ninković S. 2009, A&A http://dx.doi.org/10.1051/0004-6361/200912818

Dehnen W. 1998, AJ 115, 2384

ESA 1992, The Hipparcos Input Catalogue. ESA SP-1136

ESA 1997, The Hipparcos Catalogue. ESA SP-1200

Famaey B., Jorissen A., Luri X., Mayor M., Udry S., Dejonghe H., Turon C. 2005, A&A 430, 165

Famaey B., Pont F., Luri X., Udry S., Mayor M., Jorissen, A. 2007, A&A 461, 957

Figueras F. et al. 1997, in Hipparcos-Venice'97 (ESA SP-402), Battrick B. ed. , Noordwijk: ESA, 519

Holmberg J., Nordström B., Andersen J., 2007, A&A 475, 519

Martin N.F.G., England J.W. 1981, in Encyclopaedia of Mathematics and its applications, vol. 12, Mathematical theory of entropy. (Reading, Mass.: Addison-Wesley)

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

Skuljan J., Hearnshaw J.B., Cottrell P.L. 1999, MNRAS 308, 731

Soubiran C., Girard, P. 2005, A&A 438, 139

Strömberg G. 1925, ApJ 61, 363

Stuart A., Ord J.K. 1987, in Kendall's Advanced Theory of Statistics, vol. 1, Distribution Theory (London: Ch. Griffin & Co.)

Villani C. 2002, A review of mathematical topics in collisional kinetic theory, in Handbook of Mathematical Fluid Dynamics, vol. 1, pp. 71-305. Friedlander S. & Serre D. eds (Amsterdam: North-Holland)