

# Entropy of The Mixture Probability as Indicator of Population Discontinuities: MEMPHIS Algorithm

Rafael Cubarsi and Santiago Alcobé

Dept. Applied Mathematics IV, Technical University of Catalonia (UPC), Barcelona, Spain  
 email: rcubarsi@mat.upc.es, santiago@alcobe.net

**Abstract.** A statistical population is always associated with the distribution of a random variable. If the distribution is too complex, it is useful to think about a superposition of simpler components. For the stellar velocities, the distribution is generally assumed as a mixture of gaussian distributions, since these are particular solutions of Chandrasekhar equations for statistical equilibrium, easy to interpret dynamically. In any case, the nature of the phase space density function might be explained from the dynamics of large stellar groups, sharing a common potential, through integrals like the energy, the angular momentum, or more general quadratic integrals, rather than from the kinematics of particular groups of stars, such as those producing streaming motions. Another fact to point out is that a well defined population must be *adult*, that is, well described from some constant and consistent statistics. Otherwise, if a sample estimate varies depending on an external parameter, for example the mean age of a sample, then, either the type of distribution might be exchanged, or the population might be divided into an adequate number of subpopulations.

In this context, the MEMPHIS algorithm (Alcobé & Cubarsi 2005) has demonstrated to be useful in order to identify nearby stellar populations, by using the full space motions of a stellar sample obtained from the Hipparcos catalog (Cubarsi & Alcobé 2004). A sampling parameter, the modulus of the velocity  $|\mathbf{V}|_{\max}$ , produces a hierarchical set of nested samples allowing to detect significant population components of their velocity distribution. Although the method was designed to identify normal distributions, it also provides, as a less prior information pattern, a good approach to segregate non-gaussian populations. The entropy variations of the mixture probability allow us to estimate the number of populations without any prior assumption about such a number, according to the parameters of the Table below (sampling parameter, velocity dispersions, means –in  $\text{Km s}^{-1}$ – and vertex deviation are displayed in the  $UVW$ -cartesian heliocentric coordinates system). The main disk structure is supported by two gaussian populations, thin and thick disk. Two subcomponents were found within the thin disk, which have a high deviation from gaussianity in the radial direction, as a consequence of a non-random behaviour (Cubarsi & Alcobé 2006), but not in the other directions. Old disk stars were obtained as a broad wing of the young-thin disk. All together, with the early-thin population, they make up the thin disk component. In total 2+2 populations, since halo stars were not included in the sample.

$ \mathbf{V} _{\max}$	Population	%	$\sigma_U$	$\sigma_V$	$\sigma_W$	$U_0$	$V_0$	$W_0$	$\varepsilon$
51	Early-thin	37	$12 \pm 31$	$11 \pm 1$	$7 \pm 1$	$-21.3 \pm 4.3$	$-14.7 \pm 0.3$	$-5.9 \pm 0.2$	-
	Young-thin	38	$16 \pm 22$	$14 \pm 1$	$14 \pm 1$	$6.3 \pm 4.3$	$-7.7 \pm 0.3$	$-6.9 \pm 0.2$	-
209	Thin disk	91	$28 \pm 1$	$16 \pm 2$	$13 \pm 1$	$-10.6 \pm 0.4$	$-14.0 \pm 0.3$	$-7.2 \pm 0.3$	$10 \pm 2^\circ$
	Thick disk	09	$65 \pm 2$	$39 \pm 9$	$41 \pm 2$	$-14.7 \pm 2.9$	$-64.8 \pm 2.7$	$-8.3 \pm 2.7$	$7 \pm 3^\circ$

## References

- Alcobé, S., & Cubarsi, R. 2005, *A&A* 442, 929.  
 Cubarsi, R., & Alcobé, S. 2004, *A&A* 427, 131.  
 Cubarsi, R., & Alcobé, S. 2006, *A&A* preprint doi <http://dx.doi.org/10.1051/0004-6361:20065640>.