OB STARS IN THE SOLAR NEIGHBORHOOD. I. ANALYSIS OF THEIR SPATIAL DISTRIBUTION

F. ELIAS
Facultad de Física, Departamento de Física Atómica, Molecular, y Nuclear,
Universidad de Sevilla, Apartado 1065, Seville, Spain

J. CABRERA-CAÑO
Facultad de Física, Departamento de Física Atómica, Molecular, y Nuclear, Universidad de Seville, Apartado 1065, Seville, Spain;
and Instituto de Astrofísica de Andalucía, CSIC, Apartado 3004, Granada, Spain

AND

E. J. ALFARO
Instituto de Astrofísica de Andalucía, CSIC, Apartado 3004, Granada, Spain

ABSTRACT

We present a newly developed, three-dimensional spatial classification method, designed to analyze the spatial distribution of early-type stars within the 1 kpc sphere around the Sun. We propose a distribution model formed by two intersecting disks, the Gould Belt (GB) and the local Galactic disk (LGD), defined by their fundamental geometric parameters. Then, using a sample of about 550 stars of spectral types earlier than B6 and luminosity classes between III and V, with precise photometric distances of less than 1 kpc, we estimate for some spectral groups the parameters of our model, as well as single-membership probabilities of GB and LGD stars, thus drawing a picture of the spatial distribution of young stars in the vicinity of the Sun.

Key words: solar neighborhood — stars: early-type

1. INTRODUCTION

With the naked eye it is possible to observe that the brightest stars in the sky are mainly distributed along two great circles forming an angle of about 20° between them: the Milky Way and a tilted strip known as the Gould Belt (GB; Herschel 1847; Struve 1847; Gould 1879, p. 354). Later studies found that the GB is better described as a planar distribution of O and B stars in the solar neighborhood, inclined with respect to the Galactic plane (Lesh 1968; Stothers & Frogel 1974).

Although this structure does not show a uniform stellar distribution (the bulk of the stars tend to form aggregates around the regions of Orion and Sco-Cen, among others), the fact that the kinematic behavior of the GB members is different from that of the local Galactic disk (LGD) stars of the same spectral types (Lesh 1968; Stothers & Frogel 1974), and that several features of the local interstellar medium such as dust (Gaustad & Van Buren 1993), neutral hydrogen (Lindblad 1967; Lindblad et al. 1973), or molecular clouds (Dame et al. 1987) seem to be associated with the system of OB stars, allow us to assume that we are witnessing a star-forming process with a spatial scale length of 1 kpc. Thus, the concept of a star formation complex proposed by Efremov (1978, 1995) and recently revised by Elmegreen et al. (2000) appears to have in the GB its closest example. Extensive reviews covering the history of research about the GB and describing the present state of our knowledge and understanding of this structure can be found in Pöppel (1997) and Grenier (2006).

1.1. *A Review of Classification Methods in the Literature*

When it comes to studying the stellar component of this complex, the first and most serious problem that arises is to isolate the GB members from the stars belonging to the LGD field. This problem has been addressed before by many authors, and the proposed methods are intimately related to their a priori hypotheses about the geometry of the GB (usually, either a real belt or toroid, or a disk).

The first and most intuitive way that we came across in the scientific literature of facing the problem is that once the positions of the stars in a three-dimensional frame (normally, Cartesian Galactic coordinates X, Y, Z) are known, then all or some of their three possible projections (i.e., Y vs. X, Z vs. X, and Z vs. Y) can be plotted in order to choose limit distance criteria for the GB according to the apparent positions of the stars in each of the coordinate planes. This commonly translates into defining a "box" for each projection inside which every star belongs to the GB. Examples of this procedure can be found in Pellegatti Franco (1983), Westin (1985), Lindblad et al. (1997), and Moreno et al. (1999), among others. The greatest advantage of this method is its simplicity, but undoubtedly there is an important contamination of LGD stars in the final selection of GB members, especially in the diffuse zone of intersection between the two systems.

A second method is based on maximum likelihood analysis of the star density projected over the celestial sphere. Assuming that the stars of both the GB and the LGD are confined around two great circles in the Galactic latitude versus longitude projection, the proposed stellar density of both structures decreases as an exponential function of the angular distance to those great circles. The analysis of those distributions provides information about the structure of the GB and the LGD without having to assign individual membership probabilities to the stars. Fine examples of this method are seen in Comerón et al. (1994), Torra et al. (1997), and Fernández (2005).

Those two lines of attack have in common that the classification is based on strictly spatial criteria obtained from two-dimensional projections. But while the first yields only a raw approximation of the GB members, the second allows us to estimate some structural parameters of the system, such as its inclination with respect to the Galactic plane, the longitude of its

ascending node, its angular thickness, and the fraction of stars belonging to each group. On the other hand, this model is based on two rather strict structural hypotheses:

1. The Sun is located in the center of the GB.
2. The GB stars are distributed along a toroidal geometry.

Such restrictions impose a limitation on other possible three-dimensional scenarios (such as a disklike structure), making them difficult to use within these premises.

Therefore, provided that we know the distances of the stars, the need for a three-dimensional analysis arises. Here we can take advantage of all the spatial information at our disposal to expand the number of possible structural scenarios. The first to propose such an approach to the problem were Stothers & Frogel (1974). They assumed that both the GB and the LGD could be represented by two crossed planes around which the stars are distributed by a law decreasing with the distance from each mid-plane, the parameters that define them being estimated by least-squares fitting. Then they assigned individual membership to the stars according to their vertical distance Z (in Cartesian Galactic coordinates) from each plane; i.e., a star belonged to the GB if its Z-distance to the GB midplane was smaller than that to the LGD midplane, and vice versa. This entire procedure is nested within an iterative algorithm that recalculates the equations of the planes and reassigns memberships until convergence is reached. The disadvantages of this method are that it produces "artificially sharp surfaces on the systems on the sides that face each other" (Stothers & Frogel 1974) and that, as seems unavoidable for any separation method based only on the spatial distribution of the stars, in the region of intersection between the planes the discrimination cannot be fully trusted.

1.2. *Objectives*

We have developed a new three-dimensional spatial classification method that allows us to estimate the mean planes that define the GB and the LGD, and the probabilities of a single star belonging to either of them. Essentially, as in Stothers & Frogel (1974), we obtain the mean planes by least-squares fitting, but instead of simply assigning membership by a Z-coordinate criterion, we define a parametric stellar density distribution for each plane that makes it possible to work with membership probabilities. An essential contribution is the introduction, for the first time in this kind of study, of the detection of outliers in the sample and the analysis of their effect on the estimation of the model parameters.

In this article we give a detailed description of our classification method (§§ 2.1 and 2.2), and then we test its validity and limitations using synthetic samples (§ 2.3). A real sample of OB stars from the *Hipparcos* catalog (Perryman et al. 1997), with precise photometric distances, is compiled in § 3, and then we apply our separation algorithm to this sample in order to obtain the structural parameters that best describe the GB and the LGD (§ 4). Finally, we develop a completeness correction to enhance our results (§ 5), and we summarize the most important conclusions of this article in § 6.

2. CLASSIFICATION METHOD

2.1. *The Spatial Model*

The most restrictive simplification that we must assume to build our model is that the stars belonging to the GB and the LGD are concentrated along two planes, with the distance d of their members to the mean planes being distributed according to a parametric probability density function (pdf). Two types of

functions have been tested in this study: exponential and Gaussian pdfs:

$$\phi(d) = \frac{1}{h} e^{-d/h}, \quad \phi(d) = \frac{1}{h} e^{-d^2/2h^2}, \qquad (1)$$

where h is, respectively, the scale height and the half-width. An exponential pdf, aside from being very practical in terms of calculus, is commonly used for describing the vertical distribution of the stars in models of galactic disks (Bahcall & Soneira 1984; Gilmore 1984; Chen et al. 2001). However, it has the annoying characteristic of having a noncontinuous derivative at its maximum. Thus, an apparently more realistic function such as a Gaussian pdf has been tested too.

Using heliocentric Galactic rectangular coordinates (X, Y, Z), where X is positive in the direction of the Galactic center, Y in the direction of Galactic rotation, and Z perpendicular to the Galactic plane so that they form an orthogonal, right-handed frame, the mean planes can be expressed by the standard Cartesian plane equation

$$a_1 X + a_2 Y + a_3 Z + a_4 = 0. \qquad (2)$$

Then, the distance of any point (x, y, z) to the plane is simply

$$d = \frac{|a_1 x + a_2 y + a_3 z + a_4|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}. \qquad (3)$$

Although this definition is useful in computational terms, a better geometric understanding is provided by the following parameters: the inclination with respect to the Galactic plane (i), the Galactic longitude of the ascending node (Ω), and the vertical distance to the Sun (Z_0), where

$$a_1 = \sin i \sin \Omega, \quad a_2 = -\sin i \cos \Omega,$$
$$a_3 = \cos i, \quad a_4 = -a_3 Z_0 \qquad (4)$$

if we normalize the a_i to $\left(a_1^2 + a_2^2 + a_3^2\right)^{1/2} = 1$ (see Fig. 1 for a visualization of the geometric scheme of the problem).

2.2. *Statistical Procedure*

The classification of the star sample is done according to the principles of Bayesian discriminant analysis, a description of which can be found in Cabrera-Caño & Alfaro (1990). If we know all the parameters that define the planes, we can construct the pdfs for the GB and the LGD, ϕ_{GB} and ϕ_{LGD}, from equation (1). Then, if we know the a priori probability of any star being a GB member, f_{GB}, the Bayes theorem can be written as

$$p_{GB} = \frac{f_{GB} \phi_{GB}}{f_{GB} \phi_{GB} + (1 - f_{GB}) \phi_{LGD}}, \qquad (5)$$

where p_{GB} is the a posteriori probability of belonging to the GB for a star with a known distance to both planes. Thus, following the Bayes minimum error rate decision rule, we can classify a star as a GB member if $p_{GB} > 0.5$, or as a LGD member if $p_{GB} \leq 0.5$.

Obviously, since we lack a preclassified sample from which to obtain the parameters of the planes, we must follow an iterative procedure to estimate them. Departing from some reasonable initial values of i, Ω, and Z_0 for the planes of the GB and the LGD, of f_{GB}, and of the scale heights h_{GB} and h_{LGD} that we use for the initial classification of the sample, we obtain a first result
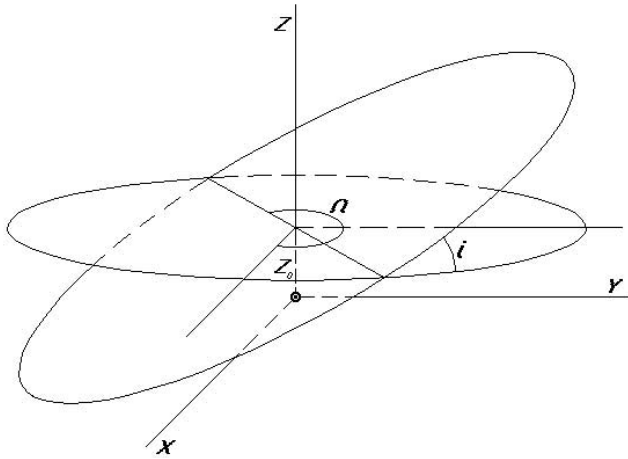
Fig. 1.—Geometric scheme of the two intersecting planes model. See the text for an explanation of the symbols.

that serves to calculate a second estimate of all the parameters, which we then use as the initial values for a third estimation, and so on, iteratively, until we reach a convergence.

2.2.1. *Main Algorithm*

The estimation algorithm is divided into the following steps:

1. The scale heights h_{GB} and h_{LGD} as parameters of the pdfs are simply estimated as the variance of the distances of the stars to the corresponding plane.

2. The probability f_{GB} is estimated as the ratio of the obtained number of GB members to the sample size.

3. The parameters of the planes are obtained by orthogonal least-squares fitting. If $r_i = (x_i, y_i, z_i)$ is the position vector of star i, we first calculate the mean position vector $\bar{r}$ and the matrix of moments $\mathbf{M}$:

$$\bar{r} = \frac{1}{N} \sum_{i=1}^{N} r_i, \tag{6}$$

$$\mathbf{M} = \frac{1}{N} \sum_{i=1}^{N} (r_i - \bar{r})^T (r_i - \bar{r}), \tag{7}$$

where N is the number of stars in the classified sample. We solve the eigenvalue equation

$$\det(\mathbf{M} - \lambda \mathbf{I}) = 0, \tag{8}$$

and then we solve for the eigenvector u

$$(\mathbf{M} - \lambda_0 \mathbf{I})u = 0, \tag{9}$$

where λ_0 is the smallest root of equation (8). Then the fitted orthogonal least-squares plane equation can be written

$$u \cdot (r - \bar{r}) = 0, \tag{10}$$

and from the values of the a_i coefficients, we can compute i, Ω, and Z_0 according to equation (4).

This iterative algorithm is fully nested within a bootstrap structure in order to simultaneously obtain error estimates of the model parameters. Thus, the procedure is performed once

with the true sample, and then it is repeated 99 more times with *pseudosamples* of the same size built by choosing stars at random—repetition being allowed—from the original batch.

2.2.2. *Detection of Outliers*

An important issue remains, however, which is that of the possible contamination of our sample by outliers that do not belong to either of our two distributions as we have defined them, i.e., that are too far away from the mean planes and cause deviations in the estimation of the parameters. This implies that an outlier will be found in zones of low density of probability in the sampling space (Cabrera-Caño & Alfaro 1985). Since we know the pdfs for the GB and the LGD, we can evaluate the total density D at the position of any star of the original sample:

$$D = f_{GB} \phi_{GB} + (1 - f_{GB}) \phi_{LGD}. \tag{11}$$

We consider as an outlier any star located at a point in space where the density D is lower than a certain threshold, which indicates regions so far from the midplanes that we can safely consider stars in these zones to not belong to our distributions. We have found after the examination of our system that a conservative value of $D < 10^{-3}$ fulfills our requirements to eliminate most of the possible outliers. All the outliers are removed from the sample before the iterative process begins and the new parameters are estimated. Once the new result has been obtained, we reintroduce the full sample and obtain the outliers as defined by the new planes, thus beginning a second iterative procedure that will lead to the final estimation of the parameters without contamination by outliers.

2.3. *Stability Analysis*

In order to study the behavior of our model, we have built a series of test samples with known parameters that we can compare with those estimated by the algorithm. These samples respond to the simplifications assumed in our model, i.e., that there are two clouds of points distributed along two crossed midplanes. In rectangular coordinates (X, Y, Z), the points are originally distributed at random in the (X, Y)-plane, while their height Z follows an exponential pdf (eq. [1]). Then each cloud of points is rotated as a whole by an angle i and then by an angle Ω, and displaced a vertical distance Z_0 (with their values obviously being different for each set of points). Finally, a distance limit r to the whole sample is chosen $[(a_1^2 + a_2^2 + a_3^2)^{1/2} \leq r]$. The number of points in each sample is about 700–800, which are the remaining "stars" after the distance cut (each sample originally has 1000 points that, to avoid undesired border effects, extend beyond the 1 kpc limit when we assign their random distances). A typical test sample after the separation procedure is shown in Figure 2, in which 5% of the LGD stars and 7.5% of the GB stars have been misclassified, mainly in the zone of intersection between the planes. Some isolated stars far away from the midplane of their corresponding distribution have also been misclassified.

In Figure 3 we plot the results of one series of tests of a simulated GB for each of the model parameters. Each point in Figure 3 (*top*) represents the mean result of the first iteration procedure for 100 samples with fixed true parameters ($i_{GB} = 17°$, *dashed line*). The initial values introduced at the beginning of the process are equal to the true parameters of the synthetic samples, except for the one whose behavior we want to analyze: in this case, i_{GB}. For each of these points we observe the following:

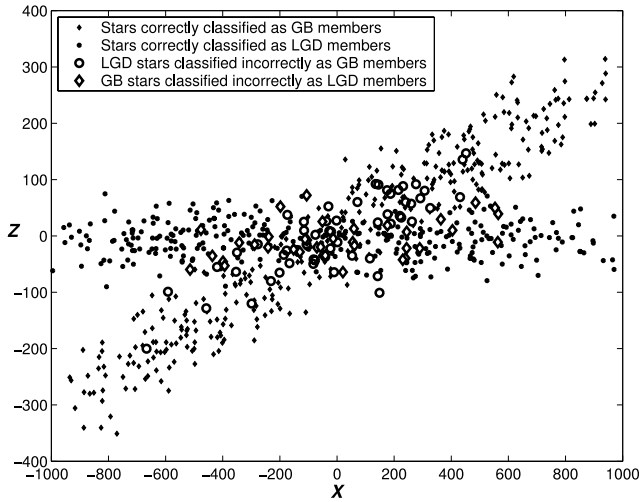1. Convergence is reached typically in less than 10 iterations, for a precision of $0°.01$.

Fig. 2.—Classification of a synthetic sample of stars with parameters $i_{GB} = 17°$, $\Omega_{GB} = 285°$, and $Z_0^{GB} = -10$.

2. The mean error of i_{GB}, estimated by the bootstrap procedure, is $0°.2$, so the number of iterations actually needed for a significant convergence is very low.

We also observe that the first iterative process leads to an underestimation of the inclination if the initial value is lower than the real one, or to an overestimation if the initial value is greater than the true inclination. A balance is achieved by the reintroduction of the full sample for a second iteration under the newly estimated parameters, which leads to a new, more re-

fined detection of the outliers. For instance, for an initial value of $i_{GB} = 14°$, an initial convergence is reached at $i_{GB} = 16°.4$; introducing $i_{GB} = 16°.4$ as the initial value leads to a final convergence of $i_{GB} = 16°.8$. Only two steps have been necessary to stabilize the result, this being the typical behavior in all the tests performed.

Similar tests for Ω and Z_0 are shown in the middle and bottom panels of Figure 3, respectively. While the results for the longitude of the ascending node are even more precise than those obtained for the inclination (convergence is reached at $\Omega_{GB} = 285° \pm 1°$ for a true value of $\Omega_{GB} = 285°$), an underestimation—in absolute value—of the vertical distance to the Sun is observed (convergence is now reached at $Z_0^{GB} = -6 \pm 2$ pc for a true value of $Z_0^{GB} = -10$ pc). Since this is the parameter of the plane most sensitive to the dispersion of the stars, it seems unavoidable that we must cope with a certain bias in the estimation of Z_0. In order to evaluate the magnitude of this systematic difference, we have performed some additional tests whose results are shown in Figure 4. One-hundred aleatory samples have been built for each value of Z_0^{GB} from 0 to 50, and then we have run the program to estimate Z_0^{GB} and its error for each case. We show their mean values in the figure, from which we can see how the estimation of Z_0^{GB} is affected by a systematic bias of around 30%.

We must note that similar tests performed systematically for Z_0^{LGD} show that there is no such bias in the case of the LGD (or, at least, it is smaller than the random errors in the estimation), probably due to its negligible inclination. This points out that it is the geometry of the problem that is responsible for such behavior of our algorithm, and thus, any pdf that we consider as a model for the vertical distribution of the planes does not alter this bias. We have corroborated this by simulating the LGD and GB
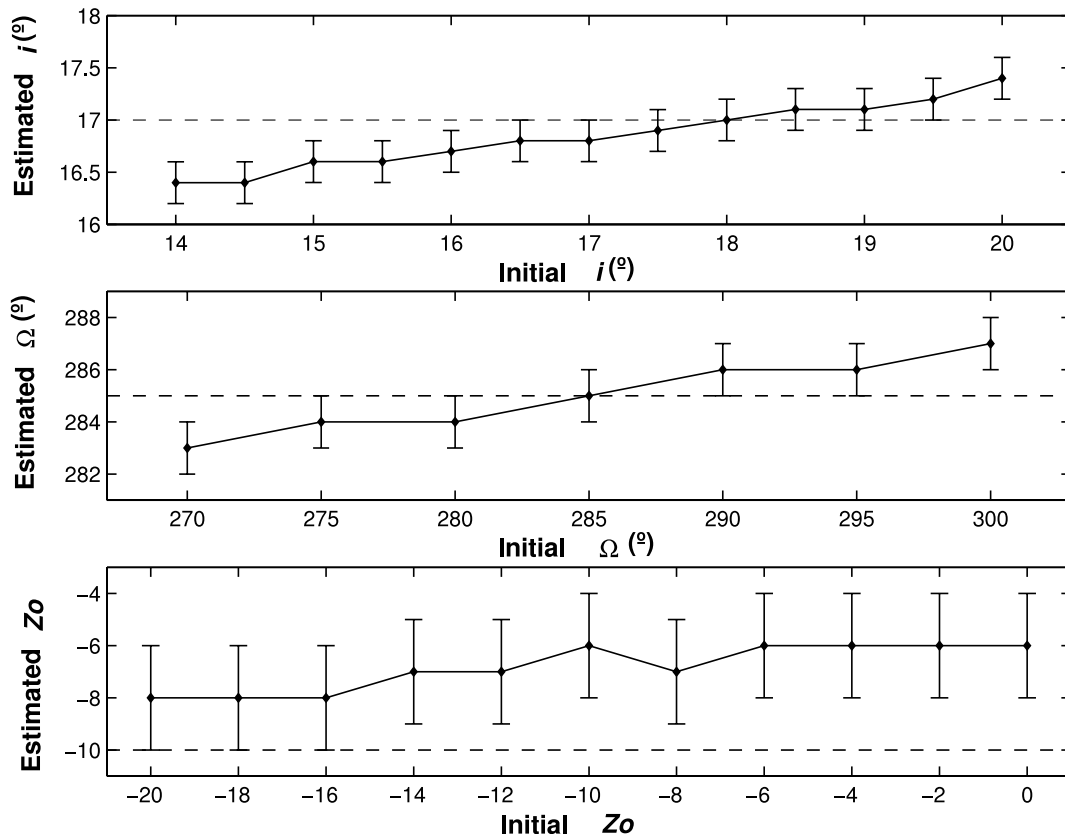

Fig. 3.—Estimation tests for a synthetic GB with the same parameter values as in Fig. 1 (marked here by dashed lines). Each data point represents the mean estimation of 100 tests with the corresponding initial value.
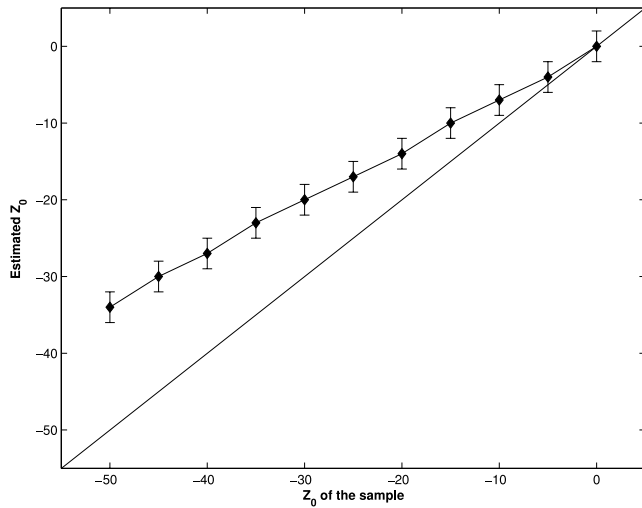
FIG. 4.—Estimation tests for a synthetic GB with different values of Z_0. Each data point represents the mean estimation of 100 tests with the corresponding true value of the sample.
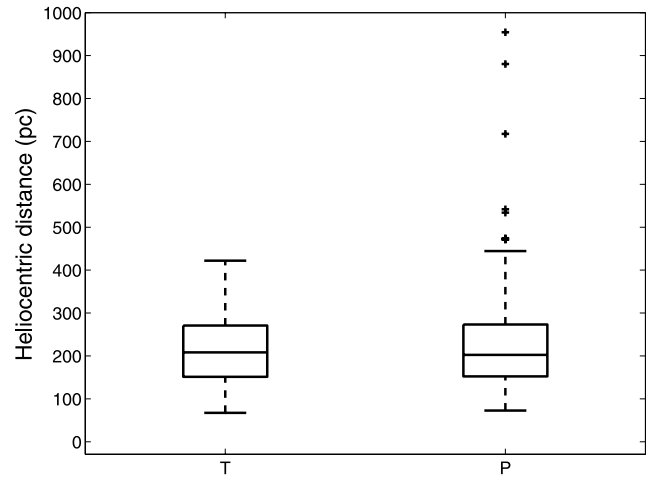

FIG. 5.—Box-plot comparison between *Hipparcos* trigonometric distances (T) and photometric distances in the Strömgren system (P) for stars with a relative error in the parallax lower than or equal to 20%.

systems with a Gaussian vertical distribution of the stars, and then performing the same tests with a Gaussian pdf model in our program. The results are consistent with the ones obtained working with exponential distributions; i.e., the LGD shows no significant bias in Z_0^{LGD}, and Z_0^{GB} is affected by a systematic bias of around 35%. When the simultaneous convergence of all the parameters (i, Ω, Z_0, and h for both planes, and f_{GB}) is sought, again only about 10 iterations for the first loop and two or three for the second loop are needed, thus confirming the stability of the model.

We want to remark on the importance and uniqueness of the detection of outliers in our procedure. Stars located in the extremes of the sample distribution may have a great weight in the estimation of the model parameters, but according to their low probability of belonging to the model distribution, they must be considered as probable outliers and thus eliminated from the process. This purge of outliers leads to quite robust results, which is essential when dealing with a system such as the GB, whose spatial distribution cannot be delimited without a certain degree of uncertainty.

3. STAR SAMPLE

We have selected a star sample from the *Hipparcos* catalog (Perryman et al. 1997) with spectral types from O to B6 and luminosity classes III, IV, and V. The spectral types listed in the *Hipparcos* catalog are a compilation from different sources; thus, a lack of homogeneity in the precision of the spectral type or the luminosity class may be present in the data. There are no systematic studies devoted to analyzing the reliability of the spectral information contained in the catalog. The closest work to such an analysis was performed by Abt (2004), who obtained and classified spectra for 584 stars belonging to *A Supplement to the Bright Star Catalog*. The comparison between this classification and that in the *Hipparcos* catalog yields that the estimated error in the spectral classification is ±1.2 subtypes and that 10% of the listed luminosity classes may be wrong. In this paper the spectral classification is only used as a rough estimate of the stellar ages. Thus, the uncertainties in the spectral classification do not have any influence on the main results of our study.

For each star, the following data were chosen:

1. HIP, *Hipparcos* identifier number.
2. Trigonometric parallax π (mas).

3. Standard error in trigonometric parallax, σ_π (mas).
4. Right ascension for the epoch J1991.5 in the International Celestial Reference System (ICRS) (degrees).
5. Declination for the epoch J1991.5 in the ICRS (degrees).

Also, $uvby$-β Strömgren photometry data from the catalog of Hauck & Mermilliod (1998) were used to calculate the photometric distance of every star in the sample through the Balona & Shobbrook (1984) calibration.

Since the relative error (σ_π/π) of the distances estimated by *Hipparcos* parallaxes grows with the trigonometric distance ($1/\pi$), we have decided to keep the trigonometric distance for a star only if the relative error in its parallax is lower than or equal to 10% (which corresponds approximately to distances closer than 100 pc from the Sun); otherwise, the photometric distance is chosen. Not only is the error in the Strömgren estimation independent of the distance, but we can also see (Fig. 5) that the distribution of photometric distances is very similar to the distribution of trigonometric ones, and that no systematic trends can be found. Comparison between the medians of both distance estimations for stars with a relative error in their parallaxes lower than or equal to 20% yields a difference between them of merely 1%. This is in good agreement with the studies performed by Kaltcheva & Knude (1998), who found no significant difference between *Hipparcos* trigonometric distances and those obtained from $uvby$ and Hβ photometry. Thus, we rest assured that the use of Strömgren photometric distances do not harm the precision of our three-dimensional picture of the solar neighborhood.

As we said, part of our sample has been selected with a relative error in its parallax lower than or equal to 10%. While the distance bias due to the nonlinear relationship between parallax and distance is negligible if σ_π/π is smaller than 10% (Arenou & Luri 1999), the Lutz-Kelker bias (Lutz & Kelker 1973) caused by truncating a sample based on the observed parallax relative error is more difficult to evaluate and depends on the parent population, as well as on the size of the selected subsample. In our case only 28 stars (about 5% of the total sample) have been selected by the parallax criterion. Thus, although this part of the sample may be affected by the Lutz-Kelker bias, the total sample shows little contamination by this effect, so we can assume that this truncation bias is negligible in our final selection.

Finally, we have eliminated the stars with an estimated distance greater than 1 kpc, because, according to results found in
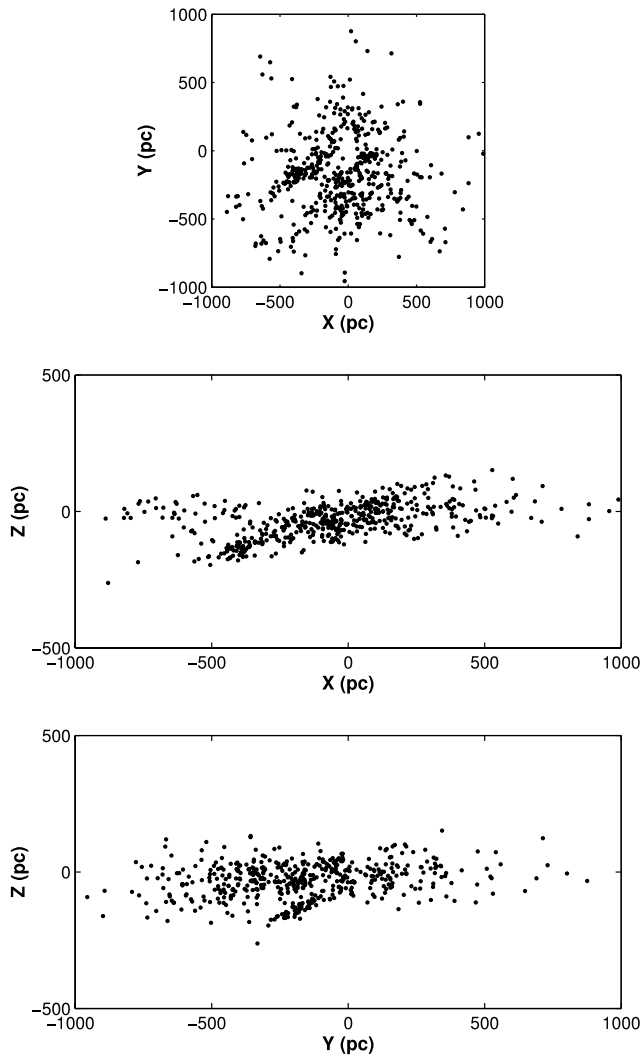
Fig. 6.—Spatial projections of the star sample.

the literature, the typical maximum radius of the GB is not greater than 700 pc (Stothers & Frogel 1974; Westin 1985; Comerón et al. 1994; Fernández 2005). The remaining sample is composed of 553 stars; Figure 6 represents their spatial distribution projections in the three Cartesian planes (X, Y) (top), (X, Z) (middle), and (Y, Z) (bottom).

4. STRUCTURAL RESULTS

We apply our model to the total star sample, as well as to four cumulative subsamples that comprise all the spectral types from O to B2, B3, B4, and B5, respectively. In all cases we find that the distribution of the stars can be satisfactorily described by our model of two intersecting planar systems, with the GB presenting a remarkable inclination with respect to the LGD. While the stars belonging to the latter tend to be homogeneously distributed on the (X, Y)-plane, the former shows a more clumpy, filamentary structure (yet the global inclination is maintained across the whole GB system). The numerical results that fit the model planes for the cumulative samples of increasing spectral types are displayed in Table 1. We have listed separately the solutions depending on whether an exponential or a Gaussian pdf has been used.

While some of the model parameters are found to be the same—within the estimated errors—regardless of the employed pdf, others show nonnegligible differences. The number of outliers and the fraction of stars belonging to the GB system (f_{GB}) especially display significant discrepancies. The number of outliers working with a Gaussian pdf is close to 20% of the total sample, meaning that about 100 out of 554 stars have a very low probability of belonging either to the GB or to the LGD. This fraction, however, is reduced to 10% when an exponential pdf is used. According to these results we consider that our sample is better described by an exponential law than by a Gaussian one. We must note that this cannot be extrapolated for the GB system as a whole, and thus must be seen only as the best fit for our observational sample.

We find an inclination of the GB between $14° \pm 1°$ and $17° \pm 0°.3$ for all samples, except for the youngest subsample of O–B2 stars, in which the range is reduced to $i_{GB} = 16° \pm 2°$–$17° \pm 1°$. It is a smaller value than those commonly found in the

TABLE 1
Estimated Parameters of the GB and LGD Planes

Sp	N	N'	n	f_{GB}	h_{GB} (pc)	h_{LGD} (pc)	i_{GB} (deg)	i_{LGD} (deg)	Ω_{GB} (deg)	Ω_{LGD} (deg)	Z_0^{GB} (pc)	Z_0^{LGD} (pc)
						Exponential pdf						
O–B2	181	162	3	0.53 ± 0.06	27 ± 4	34 ± 3	16 ± 2	2 ± 1	273 ± 7	354 ± 165	-14 ± 8	-10 ± 8
O–B3	301	267	2	0.54 ± 0.05	27 ± 3	34 ± 2	14 ± 1	2 ± 1	278 ± 5	352 ± 152	-19 ± 10	-10 ± 8
O–B4	341	303	3	0.54 ± 0.05	27 ± 3	35 ± 2	14 ± 1	2 ± 1	281 ± 5	354 ± 154	-17 ± 9	-17 ± 7
O–B5	484	433	1	0.54 ± 0.05	27 ± 3	35 ± 2	14 ± 1	1 ± 1	284 ± 3	356 ± 156	-13 ± 5	-17 ± 5
O–B6	553	498	1	0.54 ± 0.05	27 ± 3	34 ± 2	14 ± 1	1 ± 1	284 ± 3	355 ± 149	-13 ± 6	-16 ± 5
						Gaussian pdf						
O–B2	181	144	3	0.54 ± 0.06	18 ± 3	23 ± 3	17 ± 1	1 ± 1	277 ± 5	346 ± 147	-4 ± 5	-10 ± 5
O–B3	301	247	5	0.49 ± 0.05	18 ± 2	27 ± 2	16 ± 1	1 ± 1	280 ± 3	342 ± 126	-6 ± 5	-12 ± 4
O–B4	341	279	2	0.48 ± 0.06	18 ± 2	28 ± 2	16 ± 1	1 ± 1	280 ± 4	342 ± 116	-7 ± 3	-14 ± 4
O–B5	484	396	4	0.40 ± 0.04	15 ± 1	30 ± 2	17 ± 0.4	1 ± 1	282 ± 2	327 ± 78	-8 ± 3	-15 ± 3
O–B6	553	454	4	0.37 ± 0.03	13 ± 1	30 ± 2	17 ± 0.3	1 ± 1	284 ± 2	321 ± 58	-5 ± 3	-16 ± 3

Notes.—Sp stands for the spectral types in each cumulative sample, N gives the initial sample size, N' is the number of stars remaining in the sample after the elimination of outliers, and n is the number of iterations in the second loop needed to reach convergence (it must be noted that the initial values of the parameters given to the algorithm are the ones obtained as a result from the previous subsample, except in the O–B2 case). It must be also noted that h stands for both the scale height of the exponential model and the half-width of the Gaussian pdf solution. Errors are estimated by bootstrapping.

TABLE 2
GB PLANE PARAMETERS FOR SAMPLES OF DIFFERENT AGE

Sample Age (yr)	i_{GB} (deg)	Ω_{GB} (deg)
$<2 \times 10^7$...............................	18.9 ± 1.1	273.2 ± 3.1
$<3 \times 10^7$...............................	18.1 ± 0.9	270.9 ± 4.3
$(3\text{–}6) \times 10^7$...........................	14.3 ± 2.0	286.2 ± 5.7

NOTE.—From Westin (1985).

literature, yet we can compare it with those estimated by Westin (1985) in Table 2. In his study he also estimates a larger inclination for the youngest sample (albeit $2°$ larger than that found for our subsample with the earliest spectral types). Moreover, his result for the sample ranging from 30 to 60 Myr matches the inclination of $14°$ found for our later spectral types. Also, Torra et al. (2000a, 2000b), from a sample of *Hipparcos* OB stars with Strömgren photometric distances, estimate an inclination of $22°$ for stars younger than 30 Myr and an inclination of $16°$ for stars in the range of 30–60 Myr. It is worth noting that in the earlier work of Stothers & Frogel (1974), $i_{GB} = 18° \pm 0°.4$ for their O–B5 sample, and $i_{GB} = 19° \pm 1°$ for an O–B2.5 subsample, yet they find for stars in the range of spectral types from B3 to B5 that $i_{GB} = 16° \pm 1°$.

Comparing the values that we estimate for Ω_{GB}, we see that they are in very good agreement with those of Westin (1985). The earlier spectral types present a smaller value of Ω_{GB}, similar to that found by Westin (1985) for the youngest samples, while his result for the oldest age groups matches those we obtain when later spectral types are included. The values of $\Omega_{GB} = 275°–295°$ (depending on the sample age) found by Torra et al. (2000a, 2000b) are also in good agreement with our results. Other similar values were estimated by Torra et al. (1997), $\Omega_{GB} = 278°–290°$, and Comerón et al. (1994), $\Omega_{GB} = 284°.5$, from OB stars of the *Hipparcos* proposal.

Even though there is agreement in most of the estimations made by different authors of the GB geometric parameters around a certain range of values, it is true that this range is sometimes quite large (for instance, $i_{GB} \approx 14°–22°$). We venture not only that this is caused by the choice of different ages, spectral types, or heliocentric distance limits of the star samples but that the presence of outliers in those samples may be seriously affecting the results. It is perfectly possible that the weight of only a few stars with a low membership probability in the distribution sensibly alters the estimation, leading to unrealistic values of the parameters. Since the "limits" of the GB constitute a very diffuse zone whose boundaries cannot be easily defined, we consider that unless outliers are purged it is not possible to characterize its spatial distribution with enough confidence to define the series of parameters (such as the inclination or the longitude of the ascending node) that are usually employed to describe its geometry.

Our model also gives an estimation of the Sun's distance to the Galactic plane, which oscillates between 10 ± 8 pc (for the O–B2 and O–B3 subsamples) and 17 ± 7 and 17 ± 5 pc (for the O–B4 and O–B5 subsamples, respectively). Although this parameter is biased, as we had already foreseen in the simulations testing the model (Fig. 3 [*bottom*] and Fig. 4 perfectly illustrate that), the raw values we obtain are compatible with those found by Humphreys & Larsen (1995) from the Palomar Sky Survey star counts ($Z_\odot = 20.5 \pm 3.5$ pc), by Hammersley et al. (1995) from 2.2 and 3.5 μm maps with the DIRBE instrument of the *COBE* satellite and the Two-Micron Galactic Survey ($Z_\odot = 15.5 \pm 3$ pc), or by Cohen (1995) from the *IRAS* Point Source

Catalog in 12 and 25 μm ($Z_\odot = 15.5 \pm 0.7$ pc). However, the mean ages in these catalogs are much older than the young LGD we are dealing with in this work.

Comparison of the O–B5 subsample with the analysis of a star sample of the same spectral types by Stothers & Frogel (1974), who choose an exponential-type scale height model, shows an even better match with our results. They find that $Z_\odot = 24 \pm 3$ pc and that in the sphere of 200 pc around the Sun the scale height of the LGD is $h_{LGD} = 45 \pm 18$ pc. Their estimation of the scale height of the GB, $h_{GB} = 27 \pm 4$ for stars closer than 200 pc and $h_{GB} = 27 \pm 1$ for stars closer than 800 pc, is in perfect agreement with our results. A more recent study of the disk O–B5 stellar population by Maíz-Apellániz (2001), using a highly developed Gaussian model, concludes that $Z_\odot = 25.2 \pm 2.0$ pc and $h_{LGD} = 62.8 \pm 6.4$ pc. Also, working with a self-gravitating isothermal disk model, he finds that $Z_\odot = 24.2 \pm 2.1$ pc and $h_{LGD} = 34.2 \pm 3.3$ pc, the latter value matching almost exactly our estimation of $h_{LGD} = 35 \pm 2$ for the O–B5 subsample with an exponential model. It is worth noting too that our earlier O–B2 subsample yields similar results to those obtained by Reed (1997, 2000) for an O–B2 disk also using an exponential model ($h_{LGD} = 45 \pm 20$ pc, $Z_\odot = 9.5 \pm 3.5$ pc).

We also estimate a small inclination for the LGD between $1° \pm 1°$ and $2° \pm 1°$, which, although its high uncertainty points out that it is not a very significant value, is somewhat greater than that by Hammersley et al. (1995) of $0°.40 \pm 0°.03$. It is small enough, however, to make Ω_{LGD} remain not well defined, hence the large $(150°–160°)$ uncertainty that we should expect for pure geometric reasons when i_{LGD} is close to zero, as is the case.

Figure 7 displays the three possible projections for the classified O–B2 subsample (*left panels*) and the complete O–B6 sample (*right panels*) for the exponential pdf solution. The inclination of the GB is clearly seen in the (X, Z)-plane (*middle panels*); it is also interesting to observe the irregular spatial distribution of the GB in the (X, Y)-plane (*top panels*) presenting clustering and filamentary structures, in contrast with the more homogeneous LGD distribution. That is better seen in Figure 8, in which the spatial density distribution of both systems has been plotted. Note the clustering of the GB density distribution (*left*), the most prominent feature being located in the Orion region in the third Galactic quadrant.

5. COMPLETENESS CORRECTION

The spatial density distribution of the LGD shows that the star density decreases with increasing distance from the Sun, as we can see in Figure 8 (*right*). Certainly, such a distribution is not what we should expect in this region of the solar neighborhood. If we assume that on this scale the LGD should present a homogeneous star density, it is evident that what we observe in the figure is due to the incompleteness of our sample. Thus, under this hypothesis, we can use the two-dimensional density of the LGD as a measure of the star field incompleteness in the working region. This density map can be seen as a sort of "flat field" used for completeness correction that will allow us to improve the structural analysis of the GB.

We have to note, however, that the interstellar medium extinction pattern may also be playing an important role in the modulation of the observed density structure. While the effect of distance on the incompleteness of the sample is indistinctly reflected in the density correction for the GB and the LGD, the extinction pattern could show different features for both systems. Thus, this flat-field correction must be considered as a first approach to the problem of incompleteness, and not as a definitive solution. However, the main properties of the dust distribution in
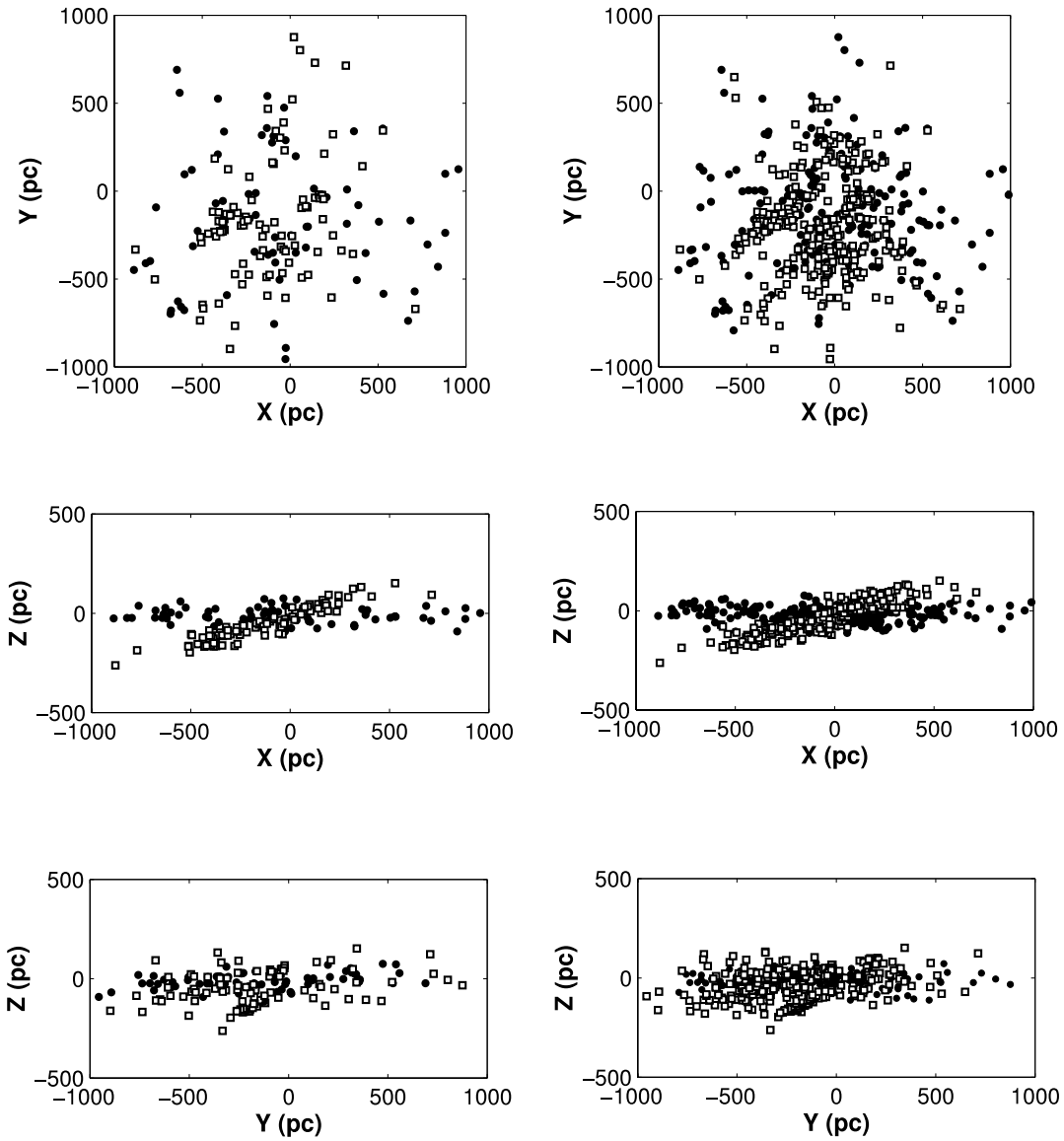
FIG. 7.—Classified O–B2 subsample (*left panels*) and full O–B6 sample (*right panels*). Circles represent the LGD, and squares the GB.
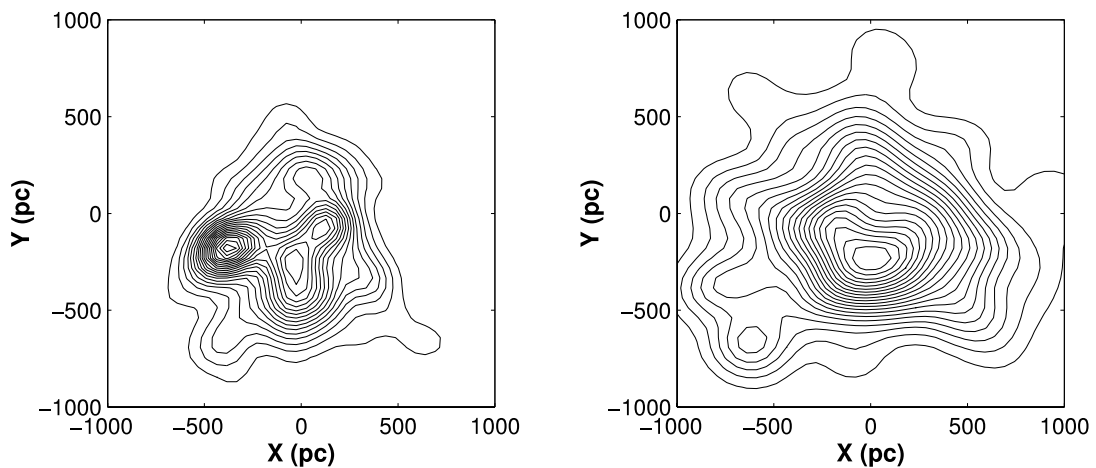

FIG. 8.—Spatial star density of the GB (*left*) and the LGD (*right*) for the classified full sample.
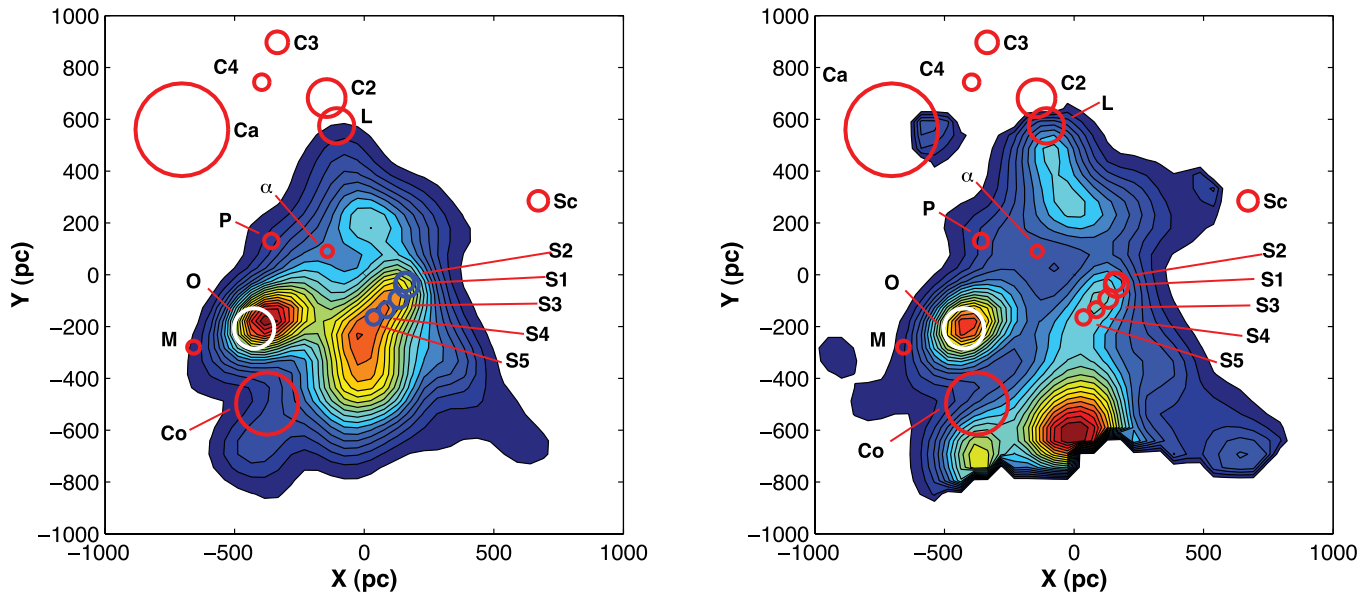
FIG. 9.—Spatial star density of the GB before (*left*) and after (*right*) the completeness correction. The circles represent the pre-*Hipparcos* OB associations, as listed by de Zeeuw et al. (1999). The letters stand for Sco OB2$_1$ (*S1*), Sco OB2$_2$ (*S2*), Sco OB2$_3$ (*S3*), Sco OB2$_4$ (*S4*), Sco OB2$_5$ (*S5*), Col 121 (*Co*), Ori OB1 (*O*), Mon OB1 (*M*), Per OB2 (*P*), α Per (α), Cam OB1 (*Ca*), Lac OB1 (*L*), Cep OB2 (*C2*), Cep OB3 (*C3*), Cep OB4 (*C4*), and Sct OB2 (*Sc*).

the solar neighborhood (see Pöppel 1997 for a detailed discussion) make us confident in the reliability of this approach.

5.1. *Model Parameters*

Dividing the GB star density by the LGD star density would be enough to provide a relative density that would enhance the structures that belong solely to the GB, but we can take a further step and introduce this correction in our model. We thus modulate the weight of any star (which originally was the probability of belonging to the correspondent plane) by a completeness function that simply is the inverse of the LGD star density at the (x, y)-position of the star, evaluated by a Gaussian kernel.

Following again the iterative process for the O–B6 full sample with an exponential pdf model, we reach convergence in the following values:

$$f_{GB} = 0.58 \pm 0.06, \quad h_{GB} = 31 \pm 4 \text{ pc}, \quad h_{LGD} = 34 \pm 5 \text{ pc},$$
$$i_{GB} = 14° \pm 1°, \quad i_{LGD} = 2° \pm 2°, \quad \Omega_{GB} = 287° \pm 6°$$
$$\Omega_{LGD} = 352° \pm 28°, \quad Z_0^{GB} = -15 \pm 12 \text{ pc},$$
$$Z_0^{LGD} = -12 \pm 12 \text{ pc}.$$

We observe that, within the error margin, the results are similar to those we had previously obtained (Table 1). This indicates that even with an incomplete sample, the structural results estimated by our model are still valid.

5.2. *Spatial Structure*

In Figure 9 we draw the nearest OB associations over the star density maps of the GB. As we can see, the completeness correction removes the contribution to the central peak by the decreasing star density with heliocentric distance, allowing the region around Scorpio to stand by itself. Also, the Orion peak moves farther away, toward the position of Ori OB1.

It is important to note that this is the first time that a completeness correction has been implemented in order to compensate for

the always incomplete samples constructed to study the spatial distribution of the GB. No significant difference can be found in the estimation of the GB structural parameters, even though the new density map of its spatial distribution (Fig. 9, *right*) shows important modifications of the original distribution (Fig. 9, *left*) in the (X, Y)-plane. This is probably due to the geometry of the system, dominated by the GB's inclination with respect to the Galactic plane (hence, the Z vertical distribution playing a fundamental role) and by the outer parts of the structure, better separated from the LGD. Nonetheless, the completeness correction is tremendously reassuring if we consider the difficulty of gathering a complete enough sample of GB stars, and it lets us look confidently at our results. It also provides a most interesting picture of the effects of incompleteness on the clumpy (X, Y)-distribution of the GB and its correspondence with OB associations.

5.3. *Associations*

Since the work by Blaauw (1965), it has been known that the young associations Sco-Cen, Per OB2, and Ori OB1 are part of the GB, while the field members could have originated from the disruption of older groups (Pöppel 1997). In fact, these three groups are well inside the typical radius of the GB, but most of the associations found within 1 kpc of the Sun seem to be outside the classical borders of the belt.

After the flat-fielding correction of the density map of the GB, we observe how new clumps appeared (Fig. 9, *right*). Within the uncertainty in the estimated distances of the associations, these clumps can be linked to well-known and cataloged associations (e.g., de Zeeuw et al. 1999). The GB membership of the associations is defined not only by the position of the centroid (a single point) but by a relative maximum in the stellar density pattern of the GB. Thus, young associations such as Cam OB1, Lac OB1, Col 121, and a clump (at $X = 0$, $Y \sim -600$ pc) that is connected to the Vela rift are likely GB members, extending the GB frontiers to larger distances.

Perhaps the most striking result is the inclusion of some Vela groups as probable GB members. We wonder, after a suggestion

from the referee, if this presence could be a spurious effect generated by the flat-field correction due to a different reddening pattern in the GB and the LGD. However, Vela is located close to the line of nodes in which the GB and the LGD coexist. Thus, the dust distribution is probably shared in this region by both systems, so it is not likely that it is responsible for the introduction of this maximum.

6. CONCLUSIONS

We have developed a new three-dimensional spatial classification method to estimate the structure of the GB and the LGD, working with single-star membership probabilities. This method, although based on a model of two intersecting planar systems, allows for a greater variety of spatial configurations than other classification techniques found in the literature, which is helpful in order to understand the complex structure of the GB. We have tested this method with artificial samples, and then we have applied it to a true sample of OB stars in the sphere of 1 kpc around the Sun. As a result, we have obtained that the distribution of young stars in the solar neighborhood is dominated by two different systems: the LGD and the inclined structure of the GB. Our hypothesis of two intersecting disks has been quite effective in characterizing both systems for their study, and although we have found that the GB is a clumpy, filamentary struc-ture, we have successfully estimated through our model the geometric parameters of the GB and the LGD, which are in good agreement with earlier estimations by other authors. The inclination of the GB is found to range over $i_{GB} = 14°–17°$.

Also, a completeness correction has demonstrated that the estimation of the global parameters that define our GB model are not affected by the decrease of the star density with heliocentric distance, providing robust parameter estimators. Yet this correction effectively refines the projected spatial density distribution of the GB on the (X, Y)-plane, removing false substructures and enhancing the true ones, some of which we can relate to the nearby OB associations.

REFERENCES

Abt, H. A. 2004, ApJS, 155, 175
Arenou, F., & Luri, X. 1999, in ASP Conf. Ser. 167, Harmonizing Cosmic Distance Scales in a Post-*Hipparcos* Era, ed. D. Egret & A. Heck (San Francisco: ASP), 13
Bahcall, J. N., & Soneira, R. M. 1984, ApJS, 55, 67
Balona, B. A., & Shobbrook, R. R. 1984, MNRAS, 211, 375
Blaauw, A. 1965, Koninkl. Ned. Akad. Wetenschap., 74, 4
Cabrera-Caño, J., & Alfaro, E. J. 1985, A&A, 150, 298
———. 1990, A&A, 235, 94
Chen, B., et al. 2001, ApJ, 553, 184
Cohen, M. 1995, ApJ, 444, 874
Comerón, F., Torra, J., & Gómez, A. E. 1994, A&A, 286, 789
Dame, T. M., et al. 1987, ApJ, 322, 706
de Zeeuw, P. T., Hoogerwerf, R., de Bruijne, J. H. J., Brown, A. G. A., & Blaauw, A. 1999, AJ, 117, 354
Efremov, Yu. N. 1978, Soviet Astron. Lett., 4, 66
———. 1995, AJ, 110, 2757
Elmegreen, B. G., Efremov, Y., Pudritz, R. E., & Zinnecker, H. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: Univ. Arizona Press), 179
Fernández, D. 2005, Ph.D. thesis, Univ. Barcelona
Gaustad, J. E., & Van Buren, D. 1993, PASP, 105, 1127
Gilmore, G. 1984, MNRAS, 207, 223
Gould, B. A. 1879, Uranometría Argentina (Buenos Aires: Coni)
Grenier, I. A. 2006, in The Young Local Universe, ed. A. Chalabaev, Y. Fukui, & T. Montmerle (Paris: Ed. Frontières), in press (astro-ph/0409096)
Hammersley, P. L., Garzon, F., Mahoney, T., & Calbet, X. 1995, MNRAS, 273, 206
Hauck, B., & Mermilliod, M. 1998, A&AS, 129, 431

Herschel, J. F. W. 1847, Results of Astronomical Observations Made During the Years 1834, 5, 6, 7, 8, at the Cape of Good Hope (London: Smith, Elder, & Co.)
Humphreys, R. M., & Larsen, J. A. 1995, AJ, 110, 2183
Kaltcheva, N., & Knude, J. 1998, A&A, 337, 178
Lesh, R. J. 1968, ApJS, 17, 371
Lindblad, P. O. 1967, Bull. Astron. Inst. Netherlands, 19, 34
Lindblad, P. O., Grape, K., Sandqvist, A., & Schober, J. 1973, A&A, 24, 309
Lindblad, P. O., Palouš, J., Lodén, K., & Lindegren, L. 1997, in *Hipparcos* Venice 1997, ed. B. Battrick (ESA SP-402; Noordwijk: ESA), 507
Lutz, T. E., & Kelker, D. H. 1973, PASP, 85, 573
Maíz-Apellániz, J. 2001, AJ, 121, 2737
Moreno, E., Alfaro, E. J., & Franco, J. 1999, ApJ, 522, 276
Pellegatti Franco, G. A. 1983, Ap&SS, 96, 195
Perryman, M. A. C., et al. 1997, The *Hipparcos* and Tycho Catalogs (ESA SP-1200; Noordwijk: ESA)
Pöppel, W. 1997, Fundam. Cosmic Phys., 18, 1
Reed, B. C. 1997, PASP, 109, 1145
———. 2000, AJ, 120, 314
Stothers, R., & Frogel, J. A. 1974, AJ, 79, 456
Struve, F. G. W. 1847, Études d'Astronomie Stellaire: Sur la Voie Lactée et Sur la Distance des Étoiles Fixes (St. Petersburg: Impr. l'Académie Impériale Sci.)
Torra, J., Fernández, D., & Figueras, F. 2000a, A&A, 359, 82
Torra, J., Fernández, D., Figueras, F., & Comerón, F. 2000b, Ap&SS, 272, 109
Torra, J., Gómez, A. E., Figueras, F., Comerón, F., Grenier, S., Mennessier, M. O., Mestres, M., & Fernández, D. 1997, in *Hipparcos* Venice 1997, ed. B. Battrick (ESA SP-402; Noordwijk: ESA), 513
Westin, T. N. G. 1985, A&AS, 60, 99