A RIGOROUS REANALYSIS OF THE IRAS VARIABLE POPULATION: SCALE LENGTHS, ASYMMETRIES, AND MICROLENSING

SERGEI NIKOLAEV AND MARTIN D. WEINBERG

Department of Physics and Astronomy, University of Massachusetts, Amherst, MA 01003-4535

Received 1997 January 8; accepted 1997 May 5

ABSTRACT

Previous work reported a bar signature in color-selected IRAS variable stars. Here, we estimate the source density of these variables while consistently accounting for spatial incompleteness in the data using a likelihood approach. The existence of the bar is confirmed with a shoulder at \( a \approx 3 \) kpc, an axis ratio of \( a:b = 2.2 - 2.7 \), and a position angle of \( 19^\circ \pm 1^\circ \). The ratio of nonaxisymmetric to axisymmetric components gives a similar estimate for the bar size of \( a = 3.3 \pm 0.1 \) kpc and a position angle of \( \phi_0 = 24^\circ \pm 2^\circ \). We estimate a scale length of \( 4.00 \pm 0.55 \) kpc for the IRAS variable population, suggesting that these stars represent the old disk population.

We use this density reconstruction to estimate the optical depth to microlensing for the large-scale bar in the Galactic disk. We find an enhancement over an equivalent axisymmetric disk by up to 40%, but this still too small to account for the MACHO result. In addition, we note a possibility for a significant asymmetry at positive and negative longitudes with optical depths comparable to those in Baade's window. Independent of our reconstruction, an infrared microlensing survey may be a sensitive tool for detecting or constraining structural asymmetries.

More generally, this is a pilot study for Bayesian star count analyses. The Bayesian approach allows the assessment of prior probabilities to the unknown parameters of the model; the resulting likelihood function is straightforwardly modified to incorporate all available data.

Subject headings: Galaxy: structure — infrared: stars — methods: statistical — stars: statistics

1. INTRODUCTION

Weinberg (1992, hereafter Paper I) identified color-selected variables in the IRAS Point Source Catalog (PSC) with asymptotic giant branch (AGB) stars based on color consistency and the circumstantial sensitivity of the IRAS survey to long-period variables (cf. Harmon & Gilmore 1988). These were then used as rough standard candles to infer a large-scale asymmetry in the stellar distribution. The identification of IRAS variables with AGB stars was strengthened by an in-depth study of a bright subset (Allen, Kleinmann, & Weinberg 1993). Carbon-selected AGB stars (carbon stars) have also proven to be effective tracers (see e.g. Metzger & Schechter 1994). Advantages of AGB tracers are reviewed in Weinberg (1994). In general, standard candle analyses have the advantage over \( \mu \) or star count analyses in that they provide direct information about the three-dimensional structure of the Galaxy. However, uncertainties in their selection and intrinsic properties may bias any inference, and especially for the IRAS-selected sample, the census is incomplete. Nonetheless, these samples are an important check on the more widely used star count and \( \mu \)-based models and deserve careful analysis.

Paper I described an approach to large-scale Galactic structure using a standard candle analysis that allows the information to be reconstructed and possibly corrected in the observer's coordinate system before being translated to a Galactocentric system. Unfortunately, this translation approach is natural only if the coverage is complete and has suffered in application to the IRAS sample because of spatial gaps due to an incomplete second full-sky epoch. Here, we present the results of a different approach to the problem: the direct density estimation by maximum likelihood. A Bayesian density estimation has the advantage of directly incorporating selection effects and missing data.

The number of ongoing surveys that bear on Galactic structure—SDSS, 2MASS, DENIS—that at various stages will have surveyed parts of the sky is a second motivation for this study; there is a need for a systematic method suited to inferential studies using possibly incomplete data from many wave bands. Recent analyses (e.g., Bahcall & Soneira 1980 in the optical; Wainscoat et al 1992 in the infrared) have modeled the Galactic components with standard profiles and structural parameters chosen to provide a match to star count data. To explore the structural parameters themselves, we propose a Bayesian density estimation technique to treat data from scattered fields during the survey and to incorporate easily data from wave bands. Conceptually, this approach is midway between a classical inversion and modeling.

The first part of the paper describes and characterizes the method. More specifically, § 2 reviews the IRAS selection procedure described in Paper I and motivates the approach. The new analysis based on statistical density estimation is presented in § 3 and is precisely defined in § 4. The second part of the paper describes Monte Carlo tests and the results of applying the method to the IRAS data (§ 5). We explore density reconstructions with and without bisymmetry and describe the sensitivity of the results to certain assumptions—bisymmetry and the value of \( R_0 \) in particular. Qualitative morphological comparison still suggests a bar. Given this interpretation, we demonstrate that the assumption of bisymmetry does not appear to distort its characterization significantly. We conclude in § 6 with a summary and discussion.

2. IRAS SOURCE SELECTION

The analysis in Paper I was based on the variables selected in the IRAS Point Source Catalog (1988) by both color...
and $P_{\text{var}}$. Following the source selection procedure described in Paper I, we selected stars from the IRAS PSC with $F_{12} > 2$ Jy and variability flag $P_{\text{var}} \geq 98\%$. A low flux limit is desirable to ensure sufficient depth of the sample. However, the analysis of IRAS sources shows that below a certain level, the census is affected by a blending of sources. The 2 Jy flux limit reduces the confusion in source identification toward the center of the Galaxy but also restricts the sensitivity to distant sources. See § 5.2 for additional discussion. The limiting distance to a star ($d$) is estimated using a simple exponential layer with vertical scale height $h$ and midplane extinction coefficient $K_{12}$:

$$m = M + 5 \log d - 5 + K_{12} h (1 - e^{-d \sin |b|/\sin |b|}).$$

(1)

For a typical AGB star ($L = 3500 \ L_\odot$, see Appendix A) and $K_{12} = 0.18 \ \text{kpc}^{-1}$, the limiting distance in the plane is $R_{\text{sun}} = 7 \ \text{kpc}$. We assume that the extinction is dominated by the molecular gas, that $h = 100 \ \text{pc}$, and that the extinct layer is horizontally isotropic. The true extinction toward the inner Galaxy is most likely dominated by the molecular ring and nuclear region given the molecular gas distribution. However, a precise estimate of the true distribution is not available, and a horizontally isotropic model will adequately represent its systematic effect on the photometric distances.

Of the more than 158,000 good-flux-quality sources listed in the IRAS PSC, 5736 satisfy both the flux-limit and variability criteria. Their spatial distribution is shown in Figures 1 and 2. To obtain variability data, at least two epochs are needed. Unfortunately, IRAS's multiple epochs did not have complete sky coverage. Most of the coverage (77% in the Galactic plane) was achieved in HCON 2 and HCON 3 separated by roughly 7.5 months on average. The rest of the Galactic plane is poorly sampled (shaded regions in Fig. 1). For this analysis, all the data in the poorly sampled sectors have been excised, reducing the size of the sample to 5500 stars.

3. METHOD OVERVIEW

All the selection effects, but especially data incompleteness, greatly complicate the analysis. Bayesian techniques are ideally suited to parameter estimation over data with general but well-defined selection criteria and underlie both the maximum entropy and maximum likelihood procedures. Below, we will parameterize the source density by an exponentiated orthogonal series with unknown coefficients $A_{ij}$ and $B_{ij}$ (cf. eq. [14]). In this context, the basic theorem of the Bayesian theory reads

$$P([A_{ij}], [B_{ij}] | D, I) = \frac{P([A_{ij}], [B_{ij}] | D) \cdot P(D | [A_{ij}], [B_{ij}], I)}{P(D | I)},$$

(2)

The probability $P([A_{ij}], [B_{ij}] | D, I)$ is the conditional (or posterior) probability of the coefficients of the source density provided the data $(D)$ and information $(I)$. The information may describe the incompleteness of the data and/or any previous independent knowledge about the source density. The probability $P([A_{ij}], [B_{ij}] | I)$ is the prior probability (or simply prior) of the coefficients provided only the information. Following Bretthorst (1990), we assign the prior using the maximum entropy principle. In our case it is constant, implying that all coefficient values are equally likely initially. The function $P(D | [A_{ij}], [B_{ij}], I)$ is the direct probability that describes the likelihood of data given the coefficients. Finally, $P(D | I)$ is a normalization constant that may be omitted provided that the posterior probability is normalized.

With these definitions, it follows that

$$P([A_{ij}], [B_{ij}] | D, I) = \text{const} \cdot P(D | [A_{ij}], [B_{ij}], I),$$

(3)

or in words, the posterior probability is proportional to the likelihood function. Therefore, the best estimate of posterior...
probability is obtained for the set of coefficients that maximize the likelihood function.

4. LIKELIHOOD FUNCTION

The likelihood is the joint probability of the observed stars given a source density. We may then consider the probability of observing a star with intrinsic luminosity in the range \((L, L + dL)\) to be detected in the distance interval \((s, s + ds)\), in the azimutal interval \((l, l + dl)\), in the galactic latitude interval \((b, b + db)\), and with magnitude in the range \((m, m + dm)\). Assuming a normal distribution of intrinsic luminosities \(L\) and a normal error distribution for the apparent magnitudes \(m\), this becomes

\[
P_n(s, l, b, m, L | \sigma_m, \sigma_L, K_{12}, h, R_0) ds \, dl \, db \, dm = \frac{1}{\sqrt{2\pi \sigma_m^2 \sigma_L^2}} \exp\left[-\frac{(L - \mu_L)^2}{2\sigma_L^2} - \frac{(m - \mu_m)^2}{2\sigma_m^2}\right] \cos b \, dl \, db \, dm . \tag{5}
\]

Here \(s, l, b\) are coordinates about the observer’s position; \(r, \phi, z\) are coordinates about the center of the Galaxy; \(L\) is the normalization constant; \(\Sigma(r, \phi, z)\) is the source density at Galactocentric radius \(R_0\); \(L\) and \(\sigma_L\) are the mean intrinsic luminosity and the dispersion of the sample; \(\mu_m\) is the measurement error in magnitudes; and \(m = \mu(s, b)\) is given by equation (1). Alternatively, we may replace luminosity by absolute magnitude:

\[
P_n(s, l, b, m, M | \sigma_m, \sigma_M, K_{12}, h, R_0) ds \, dl \, db \, dM \, dm = \frac{1}{\sqrt{2\pi \sigma_m^2 \sigma_M^2}} \exp\left[-\frac{(M - \mu_M)^2}{2\sigma_M^2} - \frac{(m - \mu_m)^2}{2\sigma_m^2}\right] \cos b \, dl \, db \, dM \, dm , \tag{6}
\]

where \(M\) and \(\sigma_M\) correspond to \(L\) and \(\sigma_L\). The Gaussian distributions in \(L\) or \(M\) in the above two equations can be generalized to an arbitrary luminosity function for traditional star count applications. Although we will not give the general expressions below, the development is parallel.

Since the convolution of two Gaussians is a new Gaussian whose variance is the sum of the two individual variances,

\[
\sigma^2_{m, eff} = \sigma^2_m + \sigma^2_M , \tag{7}
\]

equation (5) can be rewritten as

\[
P_n(s, l, b, m | \sigma_{m, eff}, k, h, R_0) ds \, dl \, db \, dm = \frac{1}{\sqrt{2\pi \sigma_{m, eff}^2}} \exp\left[-\frac{(m - \mu_m)^2}{2\sigma_{m, eff}^2}\right] \cos b \, dl \, db \, dm \tag{8}
\]

after integrating over the unmeasured absolute magnitude \(M\). For notational clarity, we will omit the subscript “eff” and write simply \(\sigma_m\). The constant \(C\) is determined from the normalization condition:

\[
C \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi \sigma_{m, eff}^2}} \exp\left[-\frac{(m - \mu_m)^2}{2\sigma_{m, eff}^2}\right] \cos b \, dl \, db \, dm = 1 . \tag{9}
\]

The integration over \(l\) runs over the entire circle except missing azimuthal sectors, explicitly accounting for missing data at particular ranges in azimuth. The limiting distance \(s_{\text{max}}\) in the \(l, b\) direction incorporates the 2 Jy flux limit.

In a standard star count analysis, no explicit distance information is provided, and \(s\) is eliminated from analysis by integration, yielding

\[
P_n(l, b, m | \ldots) \cos b \, dl \, db \, dm = \int_0^{s_{\text{max}}(b)} \Sigma(r, \phi, z) \cos b \, dl \, db \, dm . \tag{10}
\]

For our relatively small sample of \(IRAS\) stars, sensitivity to vertical structure will be poor. This motivates replacing the general unknown three-dimensional disk density with a density that depends on radial position and azimuth alone: \(\Sigma(r, \phi, z) = \Sigma(r, \phi)\).

Finally, the joint probability of observing \(N\) stars selected from the \(IRAS\) PSC is

\[
L \equiv P_{\text{total}} = \prod_{n=1}^{N} P_n(l, b, m | \ldots) . \tag{11}
\]

Expressing the likelihood function in logarithmic form, our desired solution is the set of parameters that maximize

\[
\log L = \sum_{n=1}^{N} \log P_n(l, b, m | \ldots) . \tag{12}
\]

This and nearly all star count analyses reduce to the standard problem of density estimation: find the density function \(f(x)\), which satisfies nonnegativity constraint

\[
f(x) \geq 0 \tag{13}
\]

and integral constraint

\[
\int f(x) dx = 1 , \tag{14}
\]

which best describes the observed data distribution. Both parametric and nonparametric estimation techniques have been used to solve this problem (e.g., Silverman 1986; Izenman 1991). For inhomogeneous multidimensional data, the positivity constraint is cumbersome. However, searching for the unknown function \(f(x)\) in the form of an exponentiated orthogonal series (Clutton-Brock 1990) guarantees positivity. A candidate stellar surface density is

\[
\Sigma(r, \phi) = \exp \left[ \sum_{i=1}^{i_{\text{max}}} \sum_{j=0}^{j_{\text{max}}} (A_{ij} \cos j \phi + B_{ij} \sin j \phi) J(jk^i | r) \right] , \tag{15}
\]

where \(J(jk\, | \, r)\) is Bessel function of the \(j\)th order and \(k^i\) is the \(i\)th root of Bessel function of the \(j\)th order and are chosen to produce a complete orthogonal set over the disk of radius \(R_{\text{max}}\). The coefficients \(A_{ij}, B_{ij}\) are the parameters to be determined. There is no loss of generality in taking the Fourier-Bessel series, although the choice is arbitrary.

5. RESULTS

5.1. Sensitivity to Incompleteness

A major advantage of the approach presented here over that presented in Paper I is that the significance of inferred structure is robustly quantified. In particular, we can test the sensitivity of selection effects to the detection of a bar. To test the presence of the coverage gaps, we generated four sample disks of 1000 stars each using the source density (eq. [14]) with \((A_{ij}^2 + B_{ij}^2)^{1/2} = 1\) for \(j = 0, 2\) and zero otherwise and the following bar position angles: 0°, ±45°, and 90°. The root sum square of the coefficients \(A_{ij}\) and \(B_{ij}\) rep-
resents the strength of the \( \ell \)th radial component for the \( j \)th polar harmonic. Figure 3 shows the restored strength of a harmonic \( (A^2_{ij} + B^2_{ij})^{1/2} \) as a function of the position angle of the bar. Insensitivity of these strengths to bar position angle suggests that missing azimuths will not obscure the inference of a true bar. The computed values are consistent with the expected value of unity. Conversely, regions of missing data can produce non-axisymmetric distortions and, in principle, suggest the existence of a bar in an initially axisymmetric sample. However, analysis of a simulated axisymmetric disk \((A_{10} = A_{20} = 1; \text{all others } = 0)\) and the same azimuthal incompleteness as in the real sample shows that the power in the non-axisymmetric harmonics is about 3\% of the axisymmetric contribution. Together, these tests suggest that the mis-identification of a bar due to missing azimuthal sectors alone is unlikely.

5.2. Application to IRAS Data

The formalism developed in §4 requires the distance to Galactic center \( R_0 \), extinction in the plane \( K_{12} \), and average luminosity of the AGB stars, \( L \). We take \( R_0 = 8.0 \) kpc, \( K_{12} = 0.18 \) mag kpc\(^{-1}\), and \( L = 3500 \) \( L_\odot \). The method can be straightforwardly modified for complex models (e.g., patchy or nonuniform extinction and an arbitrary luminosity function). The only limitation here is the CPU available and sufficient data to attain a satisfactory measure of confidence.

Choosing the truncation of the series in equation (14) poses a problem common to many nonparametric density estimations: because too few terms result in large bias and too many terms increase variance, \( i_{\text{max}}, j_{\text{max}} \) would be best determined by jointly minimizing the bias and the variance. However, this approach is computationally prohibitive because of the integral in equation (9) and the normal-

![Figure 3](image-url)  
**Figure 3.** The amplitude of harmonic coefficients as functions of the position angle of the bar. Open triangles, \( i = 1, j = 0 \); open squares, \( i = 1, j = 2 \); filled triangles, \( i = 2, j = 0 \); filled squares, \( i = 2, j = 2 \). The symbols are slightly offset along the x-axis for clarity.

Therefore, a heuristic approach was adapted in selecting \( i_{\text{max}}, j_{\text{max}} \) based on the increase in the likelihood function when a particular term or set of terms is added. Significance could be quantified in terms of the likelihood ratio (Wilks 1962), but we have not done this here. In addition, the hardware available to us makes it impossible to sample the parameter space beyond \( i_{\text{max}} = 4, j_{\text{max}} = 4 \) (see Appendix B for computational details). Nevertheless, up to that limit, the space was sampled thoroughly, with some of the solutions shown in Figure 4 along with the corresponding offsets of the likelihood function (the lowest value of likelihood is set to 0 for ease in comparison). Some of the features feature the ghost peaks due to the absence of data beyond the Galactic center or in missing azimuthal sectors (cf. Figs. 1 and 2). The likelihood analysis may attempt to place a nonexistent source density peak in that region, providing it will increase the overall score. We will pursue penalizing the likelihood function and other procedures for choosing an alternative prior (dropping the assumption that all coefficients in eq. [14] are equally likely initially) in future work.

All reconstructions in Figure 4 imply a jetlike feature in the first quadrant. The depth of our sample (estimated to correspond to a mean distance of 7 kpc in the plane) prevents ascertaining whether this feature corresponds to a bisymmetric bar or is a lopsided distortion. A deeper sample is needed to resolve this question. Unfortunately, lowering the flux limit to 1 Jy increases the depth of the sample only slightly (up to 9 kpc at \( b = 0^\circ \)); more important, the Bayesian analysis with a lower flux limit is difficult to implement since the likelihood function must account for possible blending of the sources and additional parameters need to be introduced to model it. As shown in Paper I, decreasing the flux limit to 1 Jy leads to detection of similar feature on the far side of the Galaxy, suggesting a real bar. The apparent similarity of the source density (Fig. 4) and the isophotal maps of external barred galaxies (Ohta et al. 1990) also advocates bisymmetric reconstruction. We use this knowledge as the independent information in equation (2) and modify our prior to include only even azimuthal harmonics. For inferences that follow, we will use the reconstruction in Figure 4 when possible. However, when the artificial absence of stars beyond the Galactic center leads to unrealistic conclusions, we will assume bisymmetry. The reconstruction with enforced bisymmetry is shown in Figure 5. Here, the corresponding prior assigns zero values to coefficients of odd azimuthal order. The likelihood value (the origin is the same as in Fig. 4) has dropped because the resulting density lacks data support beyond the Galactic center, although ghost peaks in the regions of missing data are eliminated. Figures 4 and 5 plot relative density, and therefore the true values of density in the two figures are different (in fact, the central peak value varies from panel to panel even within the same figure). For example, it appears that the feature in the first Galactic quadrant has smaller size in the nonbisymmetric case (e.g., 50\% contour levels in Fig. 4 are systematically inside the same level contours of Fig. 5). If one compares absolute contour levels, the agreement in the size of the feature is excellent.

There is no widely used procedure for quantifying the morphology of a bar (e.g., Ohta et al. 1990 Fourier-decomposed their photometry; Athanassoula et al. 1990 fitted to generalized ellipses). We chose to fit elliptical isodensity contours to our reconstruction and parameterize...
the bar length by the scale length of the semimajor axis profile. As a check, the shoulder in this profile nearly coincides with the location of the peaks in the ratio of non-axisymmetric to axisymmetric density. The logarithm of a suitable likelihood function for estimating the semimajor axes, eccentricity, and position angle is

$$\log L = \sum_{i=1}^{M} [\Sigma_{rec}(r_i, \phi_i) - C]^2,$$

where $\Sigma_{rec}(r, \phi)$ is the reconstructed density function and $C$ is isodensity level. The summation runs over equally spaced points on an ellipse. For a given ellipse, a grid of semimajor axis values are specified, and the surface density $C$, position angle $\phi_0$, and eccentricity $e$ that maximizes $\log L$ are found. The results are presented in Figures 6 and 7.

Figure 6 indicates that the density profile drops to half of its central value at about 4 kpc. Perhaps a better measure of the bar size would be the position of the inflection point along the radial profile of the bar at approximately 3 kpc. The half-length would then be about 4 kpc, in good agreement with the value obtained in Paper I. If we take this value as the size of the major axis of the bar, then the axis ratio varies from 2.2 in the central regions to 2.7 in the outer regions of the bar. The value of the position angle for the entire extent of the bar (out to 4 kpc) is $\approx 19^\circ$. The accuracy of the position angle determination can be quantified in terms of confidence interval, making use of the fact that in the limit of large number of sources $N$, the likelihood in $n$ dimensions is distributed as $\chi^2/2$ with $n$ degrees of freedom (e.g., Lehmann 1959). We analyzed the likelihood as the function of a single variable—orientation angle of the bar in the plane. The analysis gives the uncertainty of $1^\circ$ at the 3 $\sigma$ level.
Another way to determine the parameters of the bar is to look at the map of the ratio of nonaxisymmetric to axisymmetric components of the density. The ratio displays two peaks at $3.3 \pm 0.1$ kpc located on the opposite sides from the center; the line connecting them has the position angle of $\sim 24^\circ \pm 2^\circ$. The peak ratio, the relative strength of the bar, is 0.73. This implies the existence of a strong bar in the intermediate-age population responsible for the AGB stars.

5.3. Disk Scale Length

Having calculated the source density, we are in a position to characterize the parent population of the IRAS variables. In Paper I, we assumed that these variables represented a disk population based on their flux distribution, but several colleagues have suggested in discussion that the IRAS variables are more likely to be bulge stars. Here, we determine the scale length of the population in the Galactic plane. For comparison, we fit our reconstruction by an oblate spheroid model (the G0 bulge model from the DIRBE study by Dwek et al. 1995):

$$\Sigma_{G0}(x, y) = \Sigma_0 e^{-0.8r^2}, \quad (16)$$

with $r^2 = (x^2 + y^2)/r_0^2$. The scale length $r_0$ is found by minimizing the following cost function while simultaneously satisfying the overall normalization constraint for $\Sigma_{G0}$ (eq. [13]):

$$\text{cost} = \int d^2r (\Sigma_{\text{rec}} - \Sigma_{G0})^2. \quad (17)$$

To estimate the value of $r_0$, we used the covariance matrix from the likelihood analysis used to determine $\Sigma_{\text{rec}}$ to make...
5000 Monte Carlo realizations of the source density. The ensemble of realizations, then, have $\Sigma_{\text{rec}}$ as their mean. For each realization, we found $r_0$ by minimizing the cost function (eq. [17]), and the resulting distribution of scale lengths is shown in Figure 8. Our result of $r_0 = 4.00 \pm 0.55$ kpc indicates that the IRAS variables have the scale length of the old disk population. This value is in good agreement with the scale length of 4.5 kpc reported by Habing (1988), derived from an analysis of a color-selected IRAS sample.

An alternative way to estimate the scale length of the population is the parametric likelihood analysis with the source density of the G0 model, which yields $r_0 = 4.6$ kpc. Dwek’s value obtained by analyzing bulge emission was $r_0 = 0.91 \pm 0.01$ kpc. The factor of 5 difference between the scale lengths suggests that the IRAS bar and the bulge bar belong to distinct populations. Our analysis does not preclude the existence of the smaller bulge bar within 2 kpc from the Galactic center, which Dwek et al. deduced from COBE data. The Milky Way may have both a bulge bar and a disk bar, but our IRAS sample has insufficient depth and size to distinguish between the two. We expect that 2MASS-based star count analyses will allow us to characterize the features in both components and on both scales. In addition, longitudinal asymmetries in gravitational microlensing rates will provide independent limits on the length scale of observed features (see § 5.4).

In § 5.1 we demonstrated that the Bayesian analysis is robust; i.e., it accurately restores the Fourier amplitudes even with the incomplete data. However, the possibility that the assumption of the bisymmetry could effectively stretch or reinforce the bar is a serious concern. The most likely source of a stretch caused by imposing bisymmetry is inaccuracy in the Galactic center position, $R_0$. To test sensitivity to this, we perform a series of parametric density restorations with source density given by equation (16) and by varying $R_0$. The results are shown in Figure 9 and indicate that the scale length is not strongly affected by possible error in $R_0$. A visual comparison of Figure 4 (top right) and Figure 5 (right) suggests similar bar lengths and comparable isophotal shapes, although the asymmetric reconstruction has a steeper central profile.

5.4. Optical Depth Due to Microlensing

Originally proposed as a test for dark matter in the Milky Way halo (Paczyński 1986), gravitational microlensing was later shown (Griest et al. 1991; Paczyński 1991) to be potentially useful for extracting information about the inner regions of our Galaxy. Three groups (OGLE, MACHO, and EROS) are monitoring stars in the Galactic bulge for gravitational microlensing and have found higher event rates than most theoretical estimates. Udalski et al. (1994) derived an lensing optical depth of $\tau = (3.3 \pm 1.2) \times 10^{-6}$ toward Baade’s window ($l = 1^\circ$, $b = -3^\circ9$) based on the analysis of the OGLE data, and the MACHO group reported $\tau = 3.9^{+0.8}_{-1.2} \times 10^{-6}$ (Alcock et al. 1995a) estimated from the sample of clump giants, while theoretical estimates give optical depths in the range $0.5 \sim 2.0 \times 10^{-6}$ (e.g., Alcock et al. 1995a; Evans 1994). Following Paczyński et al.’s (1994) suggestion that a bar with a small inclination angle could enhance the optical depth, Zhao et al. (1995) have developed a detailed bar model and found $\tau = (2.2 \pm 0.5) \times 10^{-6}$. Here, we estimate the optical depth using two of our density reconstructions: bisymmetric ($i_{\text{max}} = j_{\text{max}} = 4$) and nonbisymmetric ($i_{\text{max}} = 4$, $j_{\text{max}} = 3$), assuming that our AGB sample represents the entire stellar disk.

The lensing optical depth is defined as the probability of any of the sources being lensed with magnification factor of $A > 1.34$, with

$$ A = \frac{u^2 + 2}{u\sqrt{u^2 + 4}} = \frac{r}{R_E} \quad \text{(18)} $$

(Refsdal 1964), where $r$ is the distance between the projected position of the source and the lensing mass and $R_E$ is the
radius of Einstein ring. Kiraga & Paczyński (1994) derived
\[
\tau = \frac{4\pi G}{c^2} \int_0^\infty \left\{ \int_0^L \rho [D_d(D_s - D_d)/D_d]D_d \rho D_d^{2+2\beta} dD_d \right\} ,
\]
where \( D_s \) is the distance to the sources, \( D_d \) is the distance to the deflectors, and the free parameter \( \beta \) accounts for detectability of sources in a flux-limited survey. The reasonable range is \(-3 \leq \beta \leq -1\), and we take \( \beta = -1 \) following Evans (1994) and Kiraga & Paczyński (1994). The density \( \rho = \rho_{\text{bulge}} + \rho_{\text{disks}} \), with \( \rho_{\text{bulge}} \) given by equation (1) of Kent (1992) and
\[
\rho_{\text{disk}} = C\Sigma_{\text{rec}}(r, \phi)e^{-|z|/h} ,
\]
where \( \Sigma_{\text{rec}} \) is the reconstructed surface density of our model and \( h = 0.325 \text{ kpc} \) is the scale height. The simple axisymmetric model of the bulge was chosen because we try to rectify the effect of the disk on the total optical depth. We explored two possible normalization prescriptions: (1) Assign a local column density of \(~50 M_\odot \text{ pc}^{-2} \) ("canonical disk") following Kuijken & Gilmore 1989; Gould 1990). The mass of the bisymmetric disk in this case is \( M_{\text{disk}} = 1.95 \times 10^{10} M_\odot \), while the mass of nonbisymmetric disk is \( M_{\text{disk}} = 1.65 \times 10^{10} M_\odot \). (2) Assign the total disk mass of \( M = 6 \times 10^{10} M_\odot \) (Bahcall & Soneira 1980). The second normalization ("maximal disk" of Alcock et al. 1995b) gives a local column density of approximately 100 \( M_\odot \text{ pc}^{-2} \) for a bisymmetric disk (130 \( M_\odot \text{ pc}^{-2} \) for a nonbisymmetric disk). We prefer the latter normalization because the optical depth estimate depends on the global mass distribution rather than the local density. In addition, there are some indications that the variation of the column density with Galactic longitude may be quite significant—a factor of 2–3 (Rix & Zaritsky 1995; Gnedin, Goodman, & Frei 1995). The mass of the bulge is \( M_{\text{bulge}} = 1.65 \times 10^{10} M_\odot \).

For the bisymmetric disk, the total lensing optical depth at Baade’s window for canonical normalization is \( 1.1 \times 10^{-6} \) (0.50, 0.76), where the notation \( \tau(x, y) \) means that the fraction \( x \) of the total optical depth \( \tau \) is the contribution of disk lenses and the fraction \( y \) is due to lensing of bulge sources. If the disk is maximal, the optical depth is \( 1.6 \times 10^{-6} \) (0.68, 0.59). The results for the nonbisymmetric disk are quite similar: canonical normalization produces the optical depth of \( 1.1 \times 10^{-6} \) (0.57, 0.78), while the maximal scenario gives total optical depth of \( 1.8 \times 10^{-6} \) (0.78, 0.57).

For all scenarios, optical depth is a function of the orientation of the bar. Experiments show that both symmetric and nonbisymmetric models can enhance the optical depth toward Baade’s window by 30%–40% as compared with the axisymmetric disk \( (\rho \propto e^{-r/3.5 \text{ kpc}}e^{-z/0.325 \text{ kpc}}) \) of corresponding mass. The largest enhancement is produced by orienting the asymmetric feature along the line of sight.

Current generation optical-band lensing surveys have concentrated on low-extinction bulge-centered windows to maximize the lensing event rate. An infrared-band lensing microlensing survey would be less constrained by extinction and therefore a more efficient probe of the overall structure of the Galaxy. In particular, any bar that is not perfectly aligned along the Sun–Galactic center axis will produce an asymmetry in the optical depth. We describe this asymmetry by the ratio of the difference in optical depths at positive and negative longitudes to their arithmetic mean. This ratio is shown in Figure 10 for a bisymmetric disk and Figure 11 for a nonbisymmetric disk. One can clearly see the difference in the asymmetry patterns caused by different models of the disk: the asymmetry at \( l \approx 10^\circ \) is much larger when the disk is not bisymmetric. This is probably due to the compensating effect of the lenses on the far side of the bar. At the same time, the negative asymmetry at \( |l| \approx 30^\circ \) is much more prominent in the bisymmetric case and is most likely due to a local increase in the surface density at negative longitudes close to the observer (cf. Figs. 5 and 4). Comparison with the Bahcall & Soneira (1980) model suggests that \( \beta \approx -1 \) is a fair approximation of the high-

---

**Fig. 10.—Bisymmetric “maximal” disk.** Solid line, dashed line, and dotted line represent cuts with \( b = 0^\circ, 2^\circ, \) and \( 4^\circ, \) correspondingly. **Left panel:** Asymmetry in the microlensing optical depth. **Right panel:** Average optical depth as a function of the Galactic longitude.
luminosity end of the disk luminosity function. Therefore, equation (19) also applies at large $|l|$ where both lenses and sources are disk members. More important than the details of the asymmetry is the suggestion that a pencil-beam microlensing survey in the infrared would be sensitive to global asymmetries in the stellar disk component. Confusion is not a limitation at $b = 0^\circ$ for larger values of $|l|$ (Figs. 10 and 11), and the optical depth has a magnitude similar to Baade’s window.

6. SUMMARY AND DISCUSSION

This paper explores a model-independent Bayesian estimation of the stellar density from star counts, rigorously accounting for incomplete data. The general approach can incorporate multiple colors and even different databases. The usual high dimensionality and topological complexity of the posterior distribution, however, complicates both optimization algorithms and subsequent moment analyses. We propose here a hybrid downhill plus directed-search Monte Carlo algorithm (Appendix B); the former speeds convergence, and the latter facilitates the location of the global extremum. Other similar and potentially more efficient techniques that can bypass the extremization step altogether (such as general Markov chain Monte Carlo) are worth careful consideration.

Application of the technique to the variability-selected sample described in Weinberg (1992), assumed to be AGB stars, confirms the presence of a strong nonaxisymmetric feature in the first Galactic quadrant. By imposing bisymmetry on the source density, a clear signature of a bar is obtained. The size and shape of density isophotes suggest a bar semimajor axis of approximately 3 kpc and a position angle of $\phi_0 = 18^\circ \pm 2^\circ$ at the outer edge of the bar. The analysis of the scale length for the AGB-candidate distribution gives $r_0 = 4.00 \pm 0.55$ kpc, indicating that these objects are part of the old disk population.

Finally, we use our estimate for a nonaxisymmetric Galactic disk to explore the dependence of optical depth to gravitational microlensing by bulge and disk stars. The nonaxisymmetric disk does enhance the optical depth toward Baade’s window by 30%–40%, but the overall value is still roughly a factor of 2 below the MACHO result of $q = 3.9 \pm 1.3 \times 10^{-6}$.

Of interest for future microlensing surveys is the finding that a bar in the disk produces a different asymmetry pattern from that of a lopsided distortion. The peak asymmetry can be as large as 40%–50% at $b = 0^\circ$, with similar values of $r$ to the Baade’s window field. Such a survey might best be carried out in the infrared to take advantage of the low interstellar extinction and colors of the late-type giants. At $|l| \gtrsim 30^\circ$, confusion should not be a limitation at $b = 0^\circ$.

We would like to thank Steve Price and Mike Skrutskie for comments and anonymous referee for helpful suggestions. This work was supported in part by NASA grant NAG 5-1999 and the Alfred P. Sloan Foundation.

APPENDIX A

LUMINOSITIES OF AGB STARS

The luminosities of AGB variables and the inference of their progenitor masses play a role in constraining the stellar evolution history of the Galaxy and has received some attention. Investigations based on a theoretical approach (Iben & Renzini 1983) and observations of sources close to the Galactic center (Jones & Hyland 1986) placed the luminosities somewhere between a few $\times 10^3 L_\odot$ and $6 \times 10^4 L_\odot$. Van der Veen & Habing (1990) revised the results of Jones & Hyland based on the analysis of a larger sample of OH/IR stars and found that the luminosities are in the range of $10^3$–$10^4 L_\odot$, with
the peak of the distribution at \( \sim 5000 - 5500 \, L_\odot \). They suggested the variability of the sources \( \Delta m \leq 2^m \) and possible selection effects as main reasons for higher limits of Jones & Hyland. They also noted that as many as 20% of the stars may be in the low-luminosity tail of the distribution, but only 2% or fewer can exceed the upper limit. Kastner et al. (1993) obtained kinematic luminosities based on the radial velocities of circumstellar envelopes with respect to the LSR and distances derived from the Galactic rotation curve. They found the range of \( 1.3 \times 10^4 - 2 \times 10^4 \, L_\odot \) with average uncertainty of factor of 2. The theoretical estimate was recently revised by Groenewegen et al. (1995), who obtained luminosity functions for carbon- and oxygen-rich stars based on the synthetic evolution. They found a mean luminosity for Galactic carbon- and oxygen-rich AGB stars to be 7050 \( L_\odot \) and 3450 \( L_\odot \), respectively. They stated, “The luminosity of a typical Galactic AGB star is in any case less than the \( 10^4 \, L_\odot \) often assumed.” Habing (1988) reported the average luminosity of 4000 \( L_\odot \) for a color-selected sample from the \textit{IRAS} PSC catalog. Finally, analysis of a sample of oxygen Miras using the \( P-L \) relation established on the observations of LMC Miras (Feast et al. 1989) places their average luminosity at \( L = 3900 \pm 450 \, L_\odot \). Unfortunately, we can not use the \( P-L \) relation, since \textit{IRAS} had insufficient temporal coverage to reliably constrain periods. Rather, we approximate the source density by an axisymmetric distribution at \( R_0 = 8 \, \text{kpc} \) and choose the average luminosity that maximizes the likelihood function. The results for different number of radial terms are shown in Figure 12. For 10 terms, the maximum likelihood of this axisymmetric density is achieved when \( L \approx 3000 \, L_\odot \). We adopt \( L = 3500 \, L_\odot \), which is the low end of published results, and interpret our statistical analysis as a consistency check.

APPENDIX B

COMPUTATIONAL NOTES

Likelihood maximization is the rate-limiting step in inferring the surface density from a source catalog. The cost of computing the likelihood is proportional to the sample size, so analyses of very large data sets will be technically challenging. Our “workhorse” algorithm for locating the maximum of the likelihood function is the conjugate gradient method, which is thoroughly discussed in the literature (e.g., Press et al. 1988). We have adopted an implementation by Shanno & Phua (1976; CONMIN). The algorithm has good convergence properties but requires a good initial approximation. Near the expected quadratic maximum, the convergence should be extremely rapid.

However, the likelihood function may have a large number of extrema, limiting the use of the standard downhill technique. In such cases, the simulated annealing (SA) algorithm (Metropolis et al. 1953; Otten & van Ginneken 1989) has the advantage. It places no restrictions on continuity and easily incorporates arbitrary boundary conditions and constraints. Adaptive simulated annealing (ASA; Ingber 1989)—a faster version of the SA algorithm—proved to be effective in narrowing the domain of the search to the comparatively small region in parameter space. However, in the vicinity of the extremum, it converges slowly.
The complementary features of the two techniques suggest the following two-step hybrid scheme:

1. Use a directed search algorithm (ASA) to isolate the global maximum. Although the SA class of algorithms converges slowly, there is a probabilistic guarantee of convergence: the probability of finding the maximum is inversely proportional to the total number of iterations to some power (e.g., Shu & Hartley 1987; Ingber 1993).

2. After either a limiting number of steps or a significant drop off in convergence, use the current ASA solution as input to conjugate a gradient scheme. This is motivated by our expectation that the true maximum of the likelihood function will be a quadratic form in the unknown variables.

This sequence can be repeated again, in case step 2 fails to find a well-defined maximum. The scheme is difficult to analyze but appears to work well in practice and is potentially useful for large parameter space and complex geometry (boundary conditions, irregular likelihood function) cases.

The entire computation time scales as the number of coefficients $M$ (total number of $A_{ij}$ and $B_{ij}$ in the sum in eq. [14]) and the sample size $N$: $N(2M + 1)$. Computation of the Hessian matrix requires CPU time proportional to $M^3N$. For large $M$, this is the bottleneck. However, the algorithm is straightforwardly parallelized by partitioning the data.

REFERENCES

Alcock, C., et al. 1995a, preprint

Allen, L. E., Kleinmann, S. G., & Weinberg, M. D. 1993, ApJ, 411, 188

Athanassoua, E., Morin, S., Wozniak, H., Puy, D., Pierce, M. J., Lombard, J., & Bosma, A. 1990, MNRAS, 245, 130

Bahcall, J. N., & Soneira, R. M. 1980, ApJS, 44, 73

Brethorst, G. L. 1990, in Maximum Entropy and Bayesian Methods, ed. P. F. Fougere (Dordrecht: Kluwer), 53

Clutton-Brock, M. 1990, J. Am. Stat. Assoc., 85, 411, 760

Dwek, E., et al. 1995, ApJ, 445, 716

Evans, N. W. 1994, ApJ, 437, L31

Feast, M. W., Glass, I. S., Whitelock, P. A., & Catchpole, R. M. 1989, MNRAS, 241, 375

Gnedin, O. Y., Goodman, J., & Frei, Z. 1995, AJ, 110, 1105

Gould, A. 1990, MNRAS, 244, 25

Griest, K., et al. 1991, ApJ, 372, L79

Groenewegen, M. A. T., van den Hoek, L. B., & de Jong, T. 1995, A&A, 293, 381

Habing, H. J. 1988, A&A, 200, 40

Harmon, R., & Gilmore, G. 1988, MNRAS, 235, 1025

Iben, I., Jr., & Renzini, A. 1983, ARA&A, 21, 271

Ingber, L. 1989, Math. Comput. Modelling, 12, 967

Izenman, A. J. 1991, J. Am. Stat. Assoc., 86, 413, 205

Joint IRAS Science Working Group. 1988, IRAS: Explanatory Supplement (Washington: GPO)

Jones, T. J., & Hyland, A. R. 1986, AJ, 92, 805

Kastner, J. H., Forveille, T., Zuckerman, B., & Omont, A. 1993, A&A, 275, 163

Kent, S. M. 1992, ApJ, 387, 181

Kiraga, M., & Paczyński, B. 1994, ApJ, 430, L101

Kuijken, K., & Gilmore, G. 1989, MNRAS, 239, 571

Lehmann, E. L. 1959, Testing Statistical Hypotheses (New York: Wiley)

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. 1953, J. Chem. Phys., 21, 1087

Metzger, M. R., & Schechter, P. L. 1994, ApJ, 420, 177

Ohta, K., Hamabe, M., & Wakamatsu, K. 1990, ApJ, 357, 71

Otten, R. H. J. M., & van Ginneken, L. P. P. 1989, The Annealing Algorithm (Boston: Kluwer)

Paczynski, B. 1986, ApJ, 304, 1

Paczynski, B., Stanek, K., Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., Mateo, M., & Krzemiński, W. 1994, ApJ, 435, L113

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1988, Numerical Recipes in C (New York: Cambridge Univ. Press)

Press, S. 1964, MNRAS, 128, 295

Rix, H.-W., & Zaritsky, D. 1995, ApJ, 447, 82

Shamino, D. F., & Phua, K. H. 1976, ACM Trans. Math. Software, 2(1), 87

Shu, H., & Hartley, R. 1987, Phys. Lett. A, 122(3–4), 157

Silverman, B. W. 1986, Density Estimation for Statistics and Data Analysis (New York: Chapman)

Udalski, A., et al. 1994, Acta Astron., 44, 165

van der Veen, W. E. C. J., & Habing, H. J. 1990, A&A, 231, 404

Wainscoat, R. J., Cohen, M., Volk, K., Walker, H. J., & Schwarz, D. E. 1992, ApJS, 83, 111

Weinberg, M. D. 1992, ApJ, 384, 81 (Paper I)

Weinberg, M. D. 1994, in IAU Symp. 169, Unsolved Problems of the Milky Way (Dordrecht: Kluwer), 11

Wilks, S. S. 1962, Mathematical Statistics (New York: Wiley)

Zhao, H. S., Spergel, D. N., & Rich, R. M. 1995, ApJ, 446, L13