LJMU Research Online

Whitaker, JS, Jackson, JM, Rathborne, JM, Foster, JB, Contreras, Y, Sanhueza, P, Stephens, IW and Longmore, SN

 MALT90 Kinematic Distances to Dense Molecular Clumps

http://researchonline.ljmu.ac.uk/id/eprint/7296/

Article

For more information please contact researchonline@ljmu.ac.uk

MALT90 Kinematic Distances to Dense Molecular Clumps

J. Scott Whitaker[1] , James M. Jackson[2,3] , J. M. Rathborne[4] , J. B. Foster[5] , Y. Contreras[6] , Patricio Sanhueza[2,7] ,
Ian W. Stephens[2] , and S. N. Longmore[8]

[1] Physics Department, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA; scott@bu.edu
[2] Institute for Astrophysical Research, Boston University, Boston, MA 02215, USA
[3] School of Mathematical and Physical Sciences, University of Newcastle, University Drive, Callaghan NSW 2308, Australia
[4] CSIRO Astronomy and Space Science, P.O. Box 76, Epping NSW 1710, Australia
[5] Department of Astronomy, Yale University, P.O. Box 28101, New Haven, CT 06520-8101, USA
[6] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands
[7] National Astronomical Observatory of Japan, National Institute of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[8] Astrophysics Research Institute, Liverpool John Moores University, Egerton Wharf, Birkenhead CH41 1LD, UK
Received 2017 March 27; revised 2017 August 1; accepted 2017 August 10; published 2017 September 11

Abstract

Using molecular-line data from the Millimetre Astronomy Legacy Team 90 GHz Survey (MALT90), we have estimated kinematic distances to 1905 molecular clumps identified in the ATLASGAL 870 μm continuum survey over the longitude range $295° < l < 350°$. The clump velocities were determined using a flux-weighted average of the velocities obtained from Gaussian fits to the HCO^+, HNC, and N_2H^+ (1–0) transitions. The near/far kinematic distance ambiguity was addressed by searching for the presence or absence of absorption or self-absorption features in 21 cm atomic hydrogen spectra from the Southern Galactic Plane Survey. Our algorithm provides an estimation of the reliability of the ambiguity resolution. The Galactic distribution of the clumps indicates positions where the clumps are bunched together, and these locations probably trace the locations of spiral arms. Several clumps fall at the predicted location of the far side of the Scutum–Centaurus arm. Moreover, a number of clumps with positive radial velocities are unambiguously located on the far side of the Milky Way at galactocentric radii beyond the solar circle. The measurement of these kinematic distances, in combination with continuum or molecular-line data, now enables the determination of fundamental parameters such as mass, size, and luminosity for each clump.

Key words: ISM: clouds – stars: distances – stars: formation

Supporting material: machine-readable table

1. Introduction

Although high-mass stars (stars with mass $>8 M_\odot$) dominate the luminosity, chemical enrichment, and mechanical energy input of galaxies, their formation remains poorly understood. Studying the formation of high-mass stars is therefore important in the quest to understand the self-organization of matter and energy and the consequences for the structure and evolution of galaxies. Estimation of the distances to regions of high-mass star formation is an essential part of this study. In this paper, we use the measured molecular-line velocities of dense clumps in the Millimetre Astronomy Legacy Team 90 GHz Survey (MALT90) survey to estimate their kinematic distances. We attempt to resolve the kinematic distance ambiguity using H I data from the Southern Galactic Plane Survey (SGPS; McClure-Griffiths et al. 2005; Haverkorn et al. 2006). Section 2 describes the principles and challenges of the kinematic distance method. Section 3 briefly describes the observational data from MALT90 and SGPS. Section 4 discusses the location of the MALT90 clumps on the Galactic longitude–velocity (l–v) diagram, the Galactic rotation models used in this study, and the distribution of galactocentric radii. Section 5 describes the algorithm used to resolve the kinematic distance ambiguity, including a novel technique to resolve the kinematic distance ambiguity for clumps not associated with continuum emission, and our analysis of the probability of a correct near/far assignment and errors in the distance. Section 6 reports the results of the kinematic distance analysis and presents Galactic "face-on" views of the MALT90 sources

and their possible relationship to the Galactic spiral arms. Finally, Section 7 presents a summary of the conclusions.

2. Motivation and Methods

High-mass star formation is deeply embedded within dense, dusty molecular clumps with size scales of $\sim$1 pc. To minimize extinction effects, observational studies concentrate on infrared, far-infrared, submm/mm, and radio wavelengths. Large-scale Galactic plane continuum surveys—e.g., GLIMPSE (Benjamin et al. 2003; Churchwell et al. 2009) and MIPSGAL (Carey et al. 2009) in the infrared; Hi-GAL (Molinari et al. 2010), BGPS (Aguirre et al. 2011), and ATLASGAL (Schuller et al. 2009) in the far-infrared, submm, and mm; and CORNISH (Hoare et al. 2012) in the radio regime—have now identified thousands of high-mass star-forming regions.

Although the continuum surveys are valuable, one of their major limitations is their lack of kinematic information. By measuring line-of-sight velocities from molecular lines, for example, one can separate distinct clumps that happen to lie along the same line of sight. Moreover, by revealing internal motions, the velocity fields show the dynamical state of the clump. Perhaps most important, the velocities allow the determination of the kinematic distances to the clouds. Without distance determinations, basic parameters such as mass, luminosity, and Galactic location cannot be deduced. Since high-mass star formation occurs almost exclusively in spiral arms, accurate distance determinations of high-mass star-forming regions are a critical means to determine the spiral

structure of the Milky Way. For all of these reasons, it is important to determine distances to high-mass star-forming clumps. In practice, distance measurements are difficult. Although some new techniques have been recently proposed using mid-infrared data on stellar extinction (e.g., Foster et al. 2012; Ellsworth-Bowers et al. 2013), the two most common methods are (1) kinematic distances and (2) maser parallax distances.

In the kinematic distance method, the velocity of a molecular clump is measured from its associated molecular-line emission, typically from molecular rotational transitions or ammonia inversion transitions (e.g., Dunham et al. 2011; Wienen et al. 2015). With the assumption of a rotation curve, the function that describes the average circular orbital velocity of stars around the Galactic center as a function of galactocentric radius, the measured radial velocity is simply related to the distance. In the inner Galaxy, i.e., within the solar circle, the same radial velocity corresponds to two distances, one on the near side of the Galaxy and one on the far side. This ambiguity can often be resolved by assessing the presence or absence of absorption or self-absorption lines in the 21 cm H I spectrum (cf. Kolpak et al. 2003; Roman-Duval et al. 2009; Green & McClure-Griffiths 2011; Jones & Dickey 2012; Urquhart et al. 2012, 2014; Wienen et al. 2015).

The second main method, the maser parallax technique, uses radio very long baseline interferometry (VLBI) to measure the precise positions of astronomical masers (typically CH_3OH or H_2O) associated with molecular clumps (e.g., Reid et al. 2009, 2014). With better than milliarcsecond positional accuracy, the annual parallax due to the Earth's orbital motion around the Sun can be measured and an accurate parallax distance established. The maser parallax method is the most accurate way to determine distances to clumps. Unfortunately, this method requires the presence of a bright maser and a substantial amount of observing time to monitor source positions. Many molecular clumps do not harbor masers. Indeed, in the sample of infrared dark clouds examined by Chambers et al. (2009), only 16% of the "quiescent" clumps (those without enhanced 4.5 μm emission or compact 24 μm emission) and about 50% of the "active" clumps (those with enhanced 4.5 μm emission or compact 24 μm emission) had detected H_2O masers. Moreover, since centimeter-wave VLBI networks are much more sensitive in the northern hemisphere, accurate determination of parallax distances for most molecular clumps south of decl. $\sim-30°$ is difficult. In practice, with current facilities, the distances to perhaps a few hundred molecular clumps can be measured with the parallax method.

Because the determination of a kinematic distance requires only the measurement of the clump's velocity and solution of the near/far kinematic distance ambiguity, kinematic distances can be found for thousands of high-mass star-forming clumps. To date, the largest database of star-forming molecular clumps with measured molecular-line velocities is MALT90 (Foster et al. 2011, 2013; Jackson et al. 2013). MALT90 employed the Australia Telescope National Facility (ATNF) 22 m Mopra radio telescope to observe 16 90 GHz lines and mapped several relatively abundant molecular species that trace high densities in the interstellar medium (ISM) toward a large number of probable high-mass star-forming cores in a range of evolutionary states. Due to the small solid angle occupied by the cold dense cores in which high-mass stars are believed to form, a fully sampled 90 GHz survey of a significant portion of the

Galactic plane was not feasible. Consequently, a targeted study of molecular clumps was undertaken. As reported in Foster et al. (2011), a pilot study was conducted to compare several alternative source-selection strategies. Among the options considered, the largest target catalog with high detection efficiency was the ATLASGAL 870 μm survey (Contreras et al. 2013; Urquhart et al. 2014). ATLASGAL identified over 10,000 bright, compact dust sources; these high-column-density sources are expected to be dense molecular clumps in which high-mass star formation may be occurring. MALT90 sources were selected from the ATLASGAL catalog with Galactic longitude l in the range $295° < l < 15°$ and Galactic latitude b in the range $-1° < b < 1°$.

MALT90 measured molecular-line velocities (Rathborne et al. 2016) toward over 3000 ATLASGAL clumps, allowing kinematic distance estimation for an unprecedentedly large sample. Their detection by ATLASGAL at 870 μm virtually ensures easy detection at submm wavelengths by the *Herschel* Hi-GAL survey (Molinari et al. 2010). Thus, nearly every MALT90 clump has a well-determined submm/mm spectral energy distribution (SED), and, together with the MALT90 kinematic distances, these submm/mm SEDs will allow for the determination of the masses, dust temperatures, and bolometric luminosities (e.g., Rathborne et al. 2010; Guzmán et al. 2015; Contreras et al. 2017). Indeed, analyses of the SEDs and molecular lines have already been used in conjunction with these kinematic distances in other MALT90 studies (Hoq et al. 2013; Stephens et al. 2015, 2016).

2.1. Determining Kinematic Distances

The line-of-sight velocity of a Galactic source relative to the Local Standard of Rest (LSR) can be interpreted in the context of a rotation curve $V(R)$ for the average circular motion of the Galaxy's disk to allow the determination of a distance—the kinematic distance to the source. For purely circular motions, the radial velocity with respect to the local standard of rest V_{LSR} indicates a unique galactocentric radius R for the object's orbit given by

$$R = R_0 \sin l \frac{V(R)}{V_{LSR} + V_0 \sin l}.$$

Here R is the galactocentric radius, R_0 is the distance from the Sun to the Galactic center, l is the Galactic longitude, $V(R)$ is the rotation curve (the Galactic average orbital velocity V at galactocentric radius R), V_{LSR} is the measured radial velocity with respect to the local standard of rest, and V_0 is the orbital velocity at the Sun's radius ($V_0 = V(R_0)$). For objects interior to the solar circle ($R < R_0$), the distance from the Sun d is then determined from geometry by

$$d = R_0 \cos l \pm \sqrt{R^2 - (R_0 \sin l)^2}.$$

The fact that this function is double-valued means that there are two possible kinematic distances for each measured radial velocity: a "near" distance that uses a minus sign and a "far" distance that uses a plus sign. Essentially, the line of sight intersects the circular orbit of radius R at two points, each of which has the same projected radial velocity. The fact that the kinematic distances are double-valued is called the near/far kinematic distance ambiguity. We address the resolution of the near/far kinematic distance ambiguity in Section 5.

The near/far ambiguity aside, the determination of kinematic distances is complicated by several factors. First, there is no

clear consensus on the best rotation model $V(R)$ for the Galaxy; some models, in fact, posit corrections to the not-always-standard local standard of rest, introducing a longitude-dependent variation between models. Further, if the rotation curve $V(R)$ is complicated, and especially if it is double-valued or multivalued, multiple distance solutions can exist for the same V_{LSR}. Finally, the method implicitly assumes purely circular motions. Thus, deviations from purely circular motion introduce errors in the kinematic distances. Giant molecular clouds, within which stars are formed, exhibit a velocity dispersion that is scale-dependent but of order $5 \, km \, s^{-1}$ (Stark 1984; Clemens 1985). At certain longitudes and velocities, such random motions can lead to pathologically erroneous distances (Xu et al. 2006). Deviations from circular motion are particularly pronounced near the Galactic center, where the gravitational potential is greatly modified by the presence of a central bar. Shocks at spiral arms can also induce noncircular motions of some tens of $km \, s^{-1}$ (Roberts 1969). Further, sources at small longitude have small radial velocities and normal velocity dispersion results in large distance errors. For these reasons, we have restricted our kinematic distance analysis to the Galactic longitude range $295° < l < 350°$.

Nevertheless, since in many models the rotation curve is essentially flat for galactocentric radii $R > 2 \, kpc$, for the bulk of the Galaxy the kinematic distance method provides reasonably accurate distances, apart from a few pathological combinations of l and V_{LSR}. The primary virtue of the kinematic distance method is that it provides an estimate of distance when few other options are available.

3. Observations

This study employs the molecular-line velocities measured by the MALT90 survey and described by Rathborne et al. (2016). MALT90 used the ATLASGAL 870 μm survey to select targets likely to be dense clumps and used the ATNF Mopra 22 m telescope to map a $4' \times 4'$ region around these targets; the central $3' \times 3'$ portion of each map has superior noise characteristics due to the on-the-fly mapping process employed by MALT90. The pixel size is $9''$, the angular resolution is $38''$, and the spectral resolution is $0.11 \, km \, s^{-1}$. Additional ATLASGAL targets falling within the mapped regions were added to our source list: in a total of 2014 MALT90 maps, we identified 3246 ATLASGAL targets.

We calculated the systemic velocity for each target as the flux-weighted average of the velocities of the HNC (1–0), HCO$^+$ (1–0), and N$_2$H$^+$ (1–0) rotational lines derived from MALT90 spectra averaged over a 3×3 pixel block centered on the peak of the ATLASGAL emission. The median 1σ brightness sensitivity of the averaged spectrum is 0.18 K on the T_A^* scale. Approximately 5% of the targets (175 out of 3246) had no line signal passing our detection threshold; we retained these as null observations. In approximately 9% of the targets, we found two distinct velocity components separated by more than $15 \, km \, s^{-1}$ resulting in two distinct MALT90 source entries in the MALT90 catalog. The final catalog contains 3556 entries.

To address the kinematic distance ambiguity through the H I absorption and self-absorption techniques, we use data from the SGPS (McClure-Griffiths et al. 2005; Haverkorn et al. 2006). We primarily use the ATCA+Parkes H I data cubes with continuum emission, which combined observations from the ATNF Parkes telescope and the ATNF Australia Telescope Compact Array to map 21 cm H I line and continuum emission
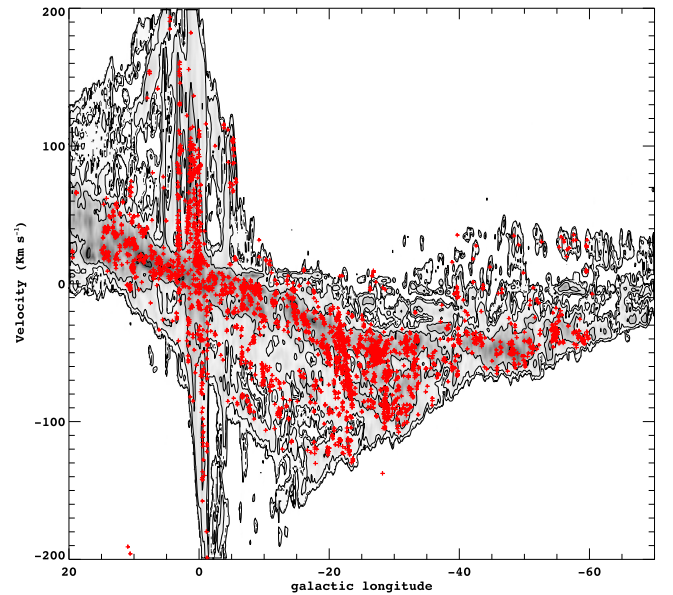
Figure 1. LSR velocity (y-axis) vs. Galactic longitude (x-axis) for all MALT90 targets from Rathborne et al. (2016). The red crosses represent the MALT90 targets. The background gray scale represents the emission from CO 1–0, integrated from Galactic longitude $-1° < b < +1°$, from the CfA-Columbia survey (Dame et al. 1987, 2001).
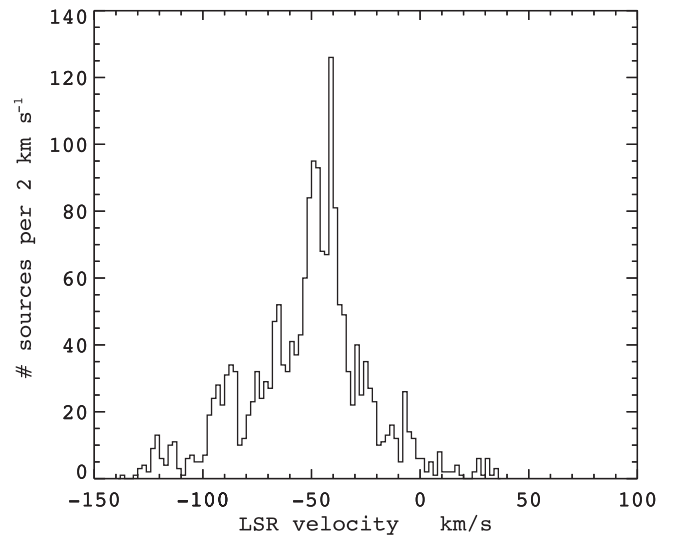
**Figure 2.** LSR velocities for sources in the longitude range $295° < l < 350°$.

in the southern Galactic plane with $253° < l < 358°$ at $\sim120''$ angular resolution and $0.82 \, km \, s^{-1}$ spectral resolution. Detailed properties of the SGPS data set that are important for our analysis are discussed below.

Of the total of 3556 MALT90 sources, 2029 fall in the fourth quadrant at longitude less than 350°. Of that group, 1908 have line detections with well-measured velocities, and, out of that sample, 1905 have SGPS data available. These 1905 sources constitute the source selection for this kinematic distance analysis.

Figure 1 displays the full set of MALT90 sources on a longitude–velocity (l–v) diagram as red crosses superposed on the CO 1–0 emission integrated from $-1° < b < +1°$ from the Columbia-CfA survey (Dame et al. 1987, 2001). The dense clumps plotted on this diagram are detected in one or more of three major dense-gas-tracing lines in MALT90: HCO$^+$ (1–0),
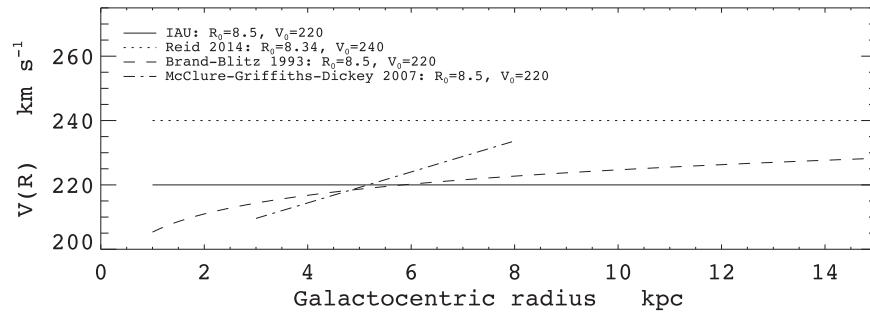
Figure 3. Several frequently used rotation curves: circular velocity as a function of galactocentric radius in kpc.
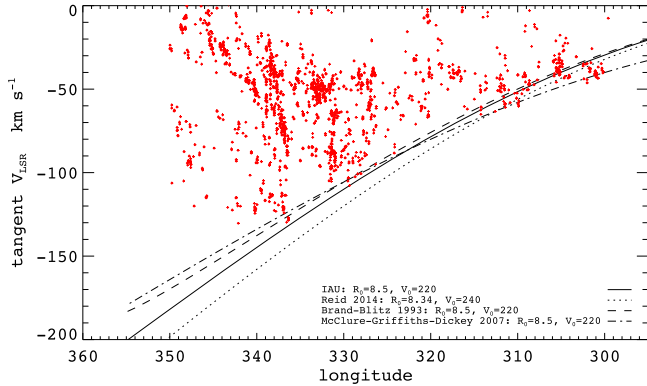
Figure 4. Tangent velocities as a function of Galactic longitude in the fourth quadrant for each rotation model. The red symbols indicate positions of the MALT90 sources in longitude–velocity space.
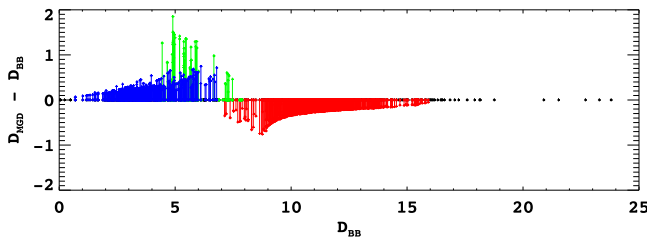

Figure 5. Differences between select distances (representative of the MALT90 ensemble) calculated with the McClure-Griffiths–Dickey model and the Brand–Blitz model, plotted against the Brand–Blitz distance. Blue indicates near kinematic distances, red indicates far distances, and green indicates distances associated with velocities beyond the tangent velocity (and so assigned the tangent-point distance) in the Brand–Blitz model. Distances corresponding to galactocentric radii outside the validity of the McClure-Griffiths–Dickey limits are shown in black.
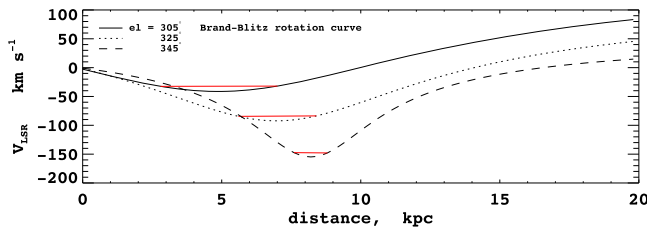
Figure 6. The V_{LSR} as a function of distance for several longitudes using the Brand–Blitz rotation curve. The horizontal red lines indicate the range of distances over which the velocity falls within 10 km s^{-1} of the tangent velocity.

N_2H^+ (1–0), and HNC (1–0). The distribution of velocities for sources in the longitude range $295° < l < 350°$ is shown in Figure 2.

All of the MALT90 sources lie within or near the regions defined by significant CO 1–0 emission. (We note two outliers
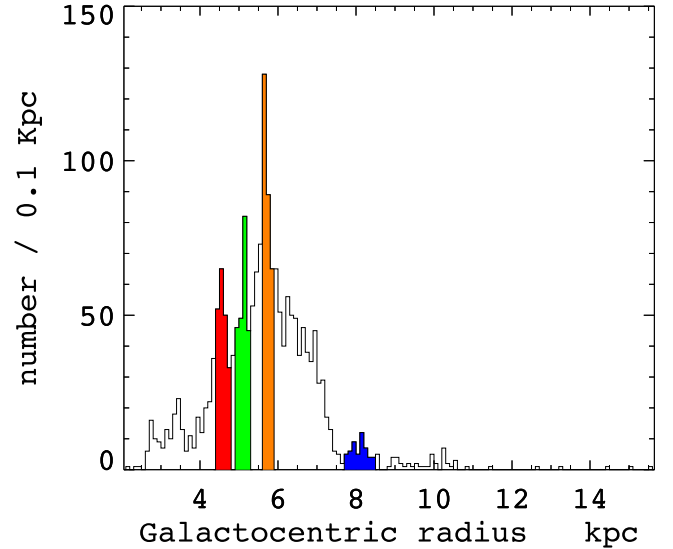

Figure 7. Galactocentric radii for fourth-quadrant MALT90 sources. Colors indicate local peaks that are likely associated with spiral arms; see text for discussion.

near 10° in longitude and -190 km s^{-1} in velocity; these are weak detections, likely due to interloping lines or spurious noise peaks, but we have not identified a likely molecular or recombination line at a suitable rest frequency.) Thus, the dense clumps are associated with the more diffuse molecular gas traced by CO 1–0. Moreover, the velocity range of the dense clumps matches that of the brightest CO emission. This indicates that the dense clumps are associated with all of the major kinematic features, usually interpreted as spiral arms, identified by CO 1–0 emission.

4. Rotation Models

Line-of-sight relative velocities provide galactocentric radii through the use of an assumed rotation curve $V(R)$. The IAU standard is a constant rotation speed of $V_0 = 220$ km s^{-1} and distance to the galactic center $R_0 = 8.5$ kpc. Many other rotation curves have been proposed, e.g., Reid et al. (2014), McClure-Griffiths & Dickey (2007), and Brand & Blitz (1993). Figure 3 illustrates these rotation curves as functions of galactocentric radius. Figure 4 plots the corresponding tangent-point velocities as a function of longitude; MALT90 sources are indicated by the data points in this longitude–velocity space. For longitudes less than 315°, the McClure-Griffiths & Dickey (2007) rotation curve is most consistent with our expectation that star-forming regions should be found at velocities up to and perhaps exceeding the tangent velocity
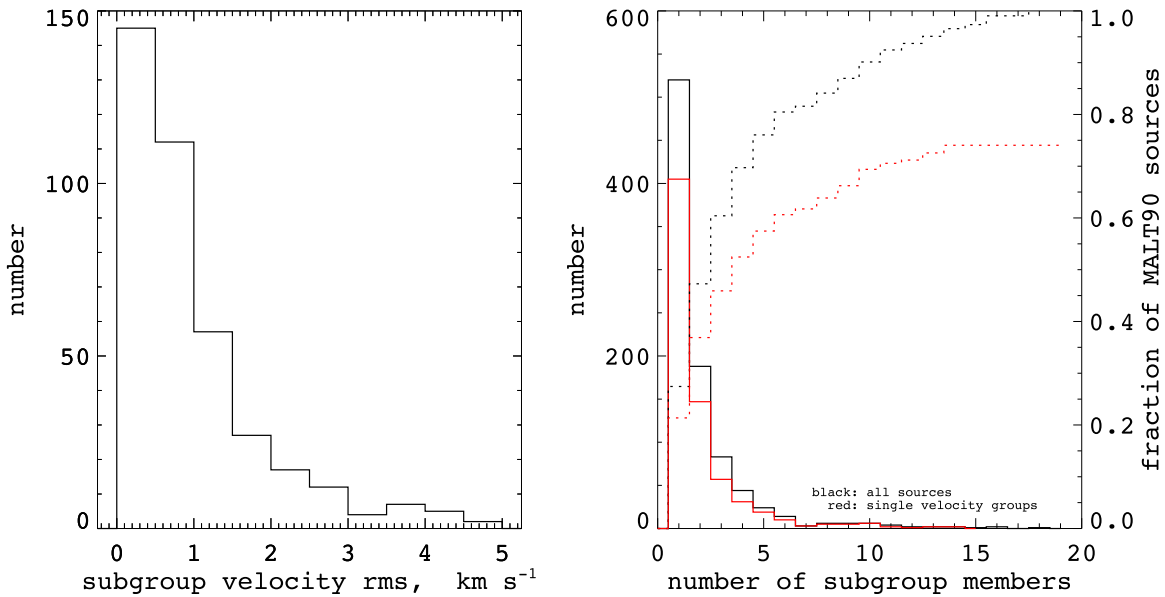
Figure 8. Left panel: subgroup velocity dispersions. The clustering algorithm applied a maximum rms deviation of 5 km s^{-1}; however, most of the spatially correlated sources fall into subgroups that have velocities within $\sigma \sim 1$ km s^{-1}. Right panel: the solid histograms show the number of subgroups with the number of members indicated on the x-axis. Black shows all subgroups, while red shows those subgroups with no others co-located (i.e., single-velocity groups with all group members in one subgroup). The dotted lines show (right axis) the cumulative fraction of sources that are in subgroups of less than or equal to the x-axis value; black is for all subgroups, and red is for single-velocity groups.
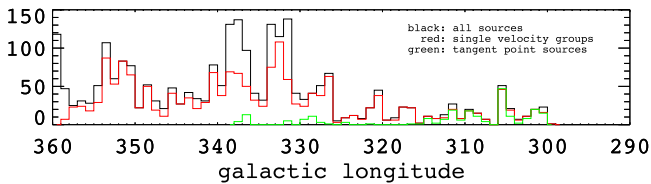
Figure 9. Histogram of the number of MALT90 sources per degree of longitude in the fourth quadrant. The total number is indicated by the black line, the red line indicates those sources that fall into single-velocity subgroups, and the green line indicates those sources assigned to the tangent distance.

by a random velocity of order 5 km s^{-1}. However, the McClure-Griffiths & Dickey (2007) model is restricted to the inner Milky Way, 3 kpc $< R <$ 8 kpc. Among our 1905 selected MALT90 sources, 43 have $R < 3$ kpc and 85 have $R > 8$ kpc. These sources will be of particular interest for studying Galactic structure.

We generally use the Brand & Blitz (1993) rotation curve to calculate galactocentric radii and near and far kinematic distances for the MALT90 clumps, but we compare with results from the McClure-Griffiths & Dickey (2007) model when looking at implications for galactic structure. For each model, we calculate error estimates for the near and far kinematic distances by varying the source radial velocity by ± 7 km s^{-1} (Reid et al. 2009). Each model presents the issue of what value of the solar rotational velocity V_0 to employ. The Brand–Blitz model is not constrained to yield V_0 at $R = R_0$; in fact, when setting $R_0 = 8.5$ kpc and $V_0 = 220$ km s^{-1}, the model yields $V(R_0) = 223.25$ km s^{-1}. The 3.25 km s^{-1} offset is just 1.5% of the nominal V_0 and less than half of the 7 km s^{-1} by which we vary the radial velocities for error estimation; however, it does affect some distance calculations where the observed radial velocity is near zero. For consistency, we have used the Brand–Blitz $V(R_0)$ for the solar rotation velocity when calculating kinematic distances. This is not an option with the McClure-Griffiths & Dickey (2007) model due the range

restrictions on R; applying the model anyway would yield $V(R_0) = 235.8$ km s^{-1}. Consequently, for this model, we have used the nominal V_0 of 220 km s^{-1} for the solar rotation. The 15.8 km s^{-1} difference between V_0 and $V(R_0)$ is outside the peak-to-peak residual range of 10 km s^{-1} of the McClure-Griffiths–Dickey fit to the SGPS data in the range 3 kpc $< R <$ 8 kpc; this discontinuity might indicate non-circular motions or streaming effects as discussed in McClure-Griffiths & Dickey (2007), or it may indicate limitations of the linear functional form used in the model.

Figure 5 shows the difference between distances calculated with the McClure-Griffiths–Dickey model and the Brand–Blitz model for representative MALT90 source velocities. The blue symbols show the difference for select near kinematic distances, red symbols for select far kinematic distances, green symbols for sources found to be near or beyond the tangent velocity in the Brand–Blitz model, and black symbols for those sources outside the 3 kpc $< R <$ 8 kpc range of the McClure-Griffiths–Dickey model. Generally, the difference between the two distance estimators is less than 10%. The greatest impact of the rotation model choice is for sources lying near the tangent velocity in the range of 300°–315° longitude; when we discuss galactic structure, we will return to the differences between the two models. We caution the reader that the rotation curve may well be refined in future work.

To accommodate the characteristic velocity dispersion of molecular clouds, it is customary for sources with velocities close to the tangent velocity to be placed at the tangent point, where "close" is typically 10–25 km s^{-1} for this "tangent-point-masking" approach. Figure 6 plots V_{LSR} as a function of distance from the LSR for three longitudes using the Brand–Blitz model; the red lines indicate the range of distances at each longitude over which the velocity is within 10 km s^{-1} of the tangent velocity. This illustrates the well-known inner-Galaxy feature of "velocity crowding," where emission from a considerable path length will appear near the tangent velocity. A consequence is that absorption by intervening clouds of
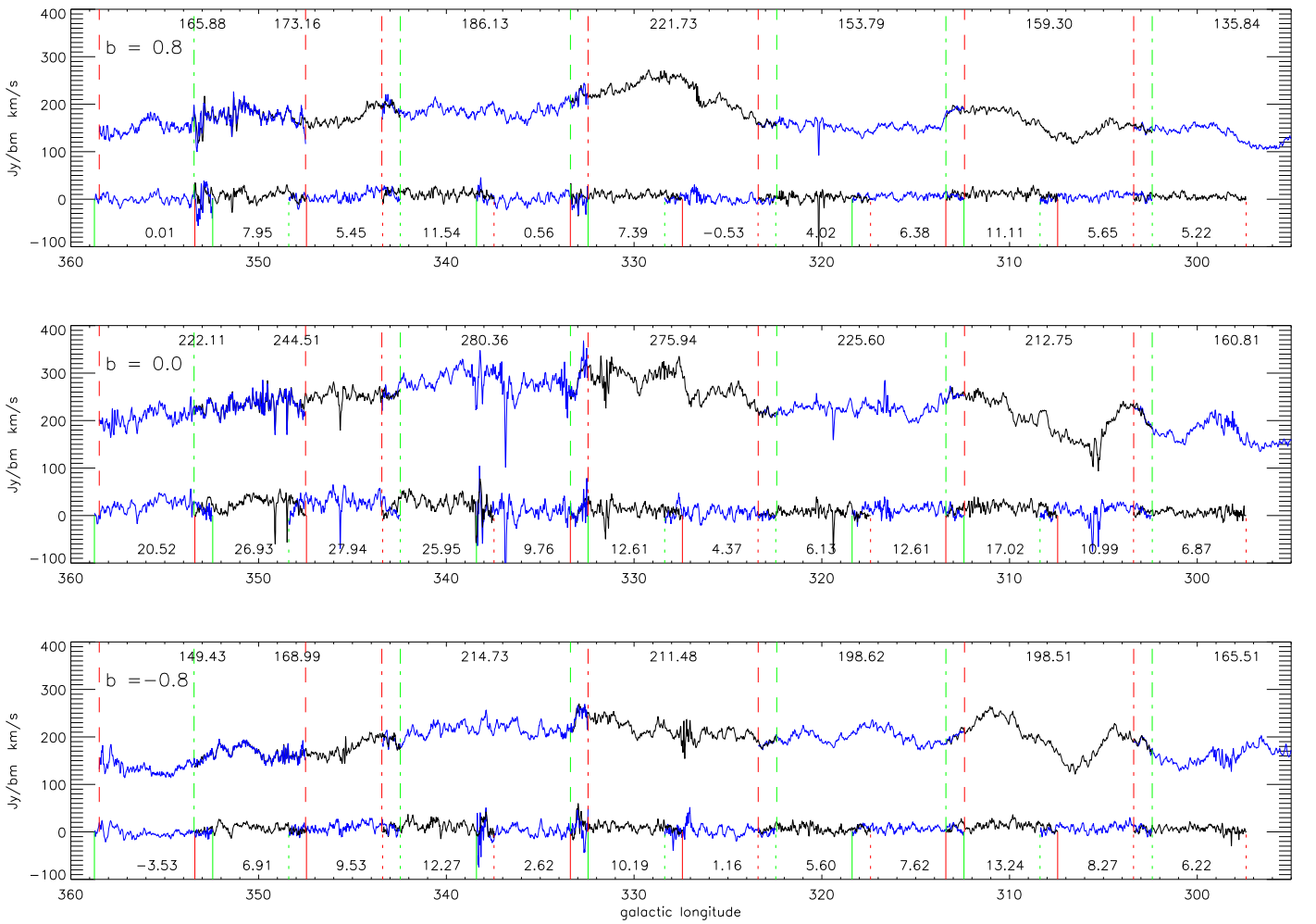
Figure 10. Integrated continuum-subtracted line signal intensity (Jy beam^{-1} km s^{-1}) as a function of longitude for three latitudes (upper left corner in each panel). The pairs of solid or dotted red or green vertical lines mark the longitude limits of the individual data files. The upper traces are the ATCA+Parkes-continuum-subtracted CS data; the lower traces are the ATCA+Parkes-with-continuum WC data with the continuum fitted and subtracted in our analysis.

continuum emission associated with far sources is likely to appear near the tangent velocity. We will exploit this in our analysis. We have not applied tangent-point masking in our analysis; the values published in this paper include the tangent-point distances and velocities so that users may apply their preferred approach.

Figure 7 shows the distribution in galactocentric radii for the sources under study. The distribution is clearly nonuniform with a few sharp peaks, notably at 4.6, 5.3, and 6.3 kpc, and a broad peak at 8.0 kpc. These peaks are color-coded for reference in later discussion. Only a small fraction of the sample is located at galactocentric radii beyond 8 kpc. As we will discuss below, the sharp peaks probably represent spiral arms and/or large nearby star-forming complexes, and the lack of sources at large galactocentric radii probably arises from sensitivity effects.

To look for possible velocity pileup and confusion effects in our analysis, we have studied the clustering of MALT90 sources in longitude–latitude–velocity space. We assigned each clump to a "group" based on its proximity in velocity and position to other clumps. The clustering analysis was done first in longitude–latitude by iteratively adding a source to a group when its angular separation from any member of the group was less than a specified maximum value, a procedure sometimes referred to as the "friends-of-friends" algorithm (e.g., Wienen

et al. 2015, and references therein). We used a maximum angular separation of 0°.04, roughly the size of the SGPS beam, as our goal was to understand the correlations of the near/far analysis for closely spaced sources. Within a group thus selected by l–b proximity, we sought subgroups with common velocities. We used a "casting-out" approach: if the rms deviation of the velocities of the group members exceeded a set level, the greatest outlier from the mean velocity was cast out iteratively until the rms criterion was satisfied. The retained sources constituted the first subgroup; the outcasts were then processed with the same approach until all group members were assigned to a subgroup. We used a maximum rms velocity deviation of 5 km s^{-1} for this analysis. The 1905 sources were divided into 801 groups; 703 of these groups had a single subgroup, i.e., a common velocity for all group members, and 410 groups had a single member. In Figure 8, the left panel shows the distribution in rms velocity dispersion for subgroups; the typical value is of order 1 km s^{-1}, and only a few subgroups are being defined by our 5 km s^{-1} cutoff. The right panel shows the number of subgroups of different sizes (solid histogram and left axis) and the normalized cumulative distribution of sources (dotted histogram and right axis); the black lines show all sources, and the red lines show the distributions for sources in the 703 single-velocity groups. As indicated by the dotted red cumulative curve for the single-velocity groups, approximately
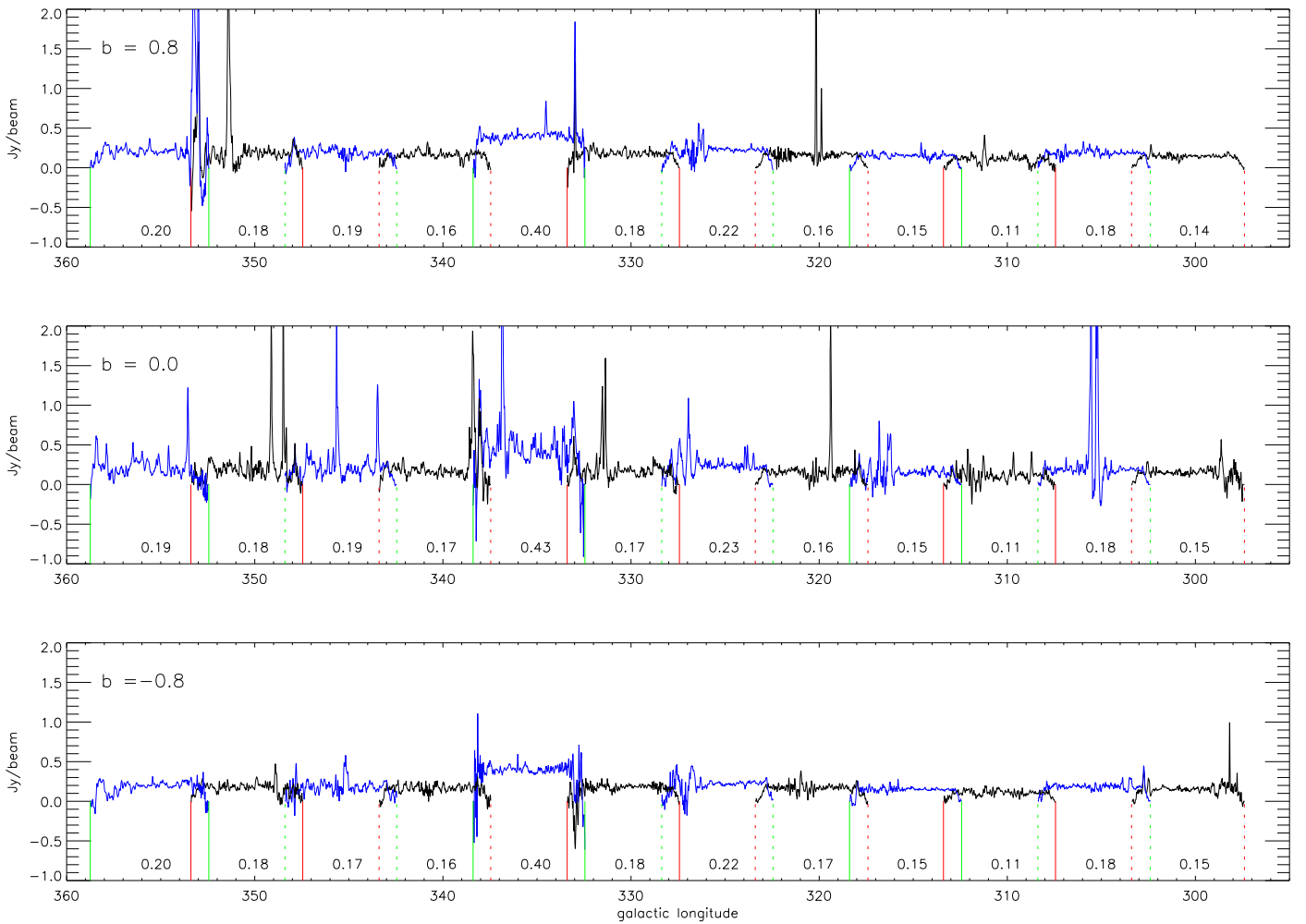
Figure 11. Continuum emission (Jy beam^{-1}) as a function of galactic longitude for three different galactic latitudes. The pairs of solid or dashed red or green vertical lines mark the longitude limits of the individual data files. The file centered on 335° exhibits a markedly higher typical value, indicating that a different spatial filtering was applied for this subset of the data.

75% of our sources are isolated, on the angular scale of the SGPS beam, from MALT90 sources at other velocities.

Figure 9 displays these data by longitude: the black histogram includes all sources, the red includes those sources that are in single-velocity groups, and the green includes sources assigned to the tangent point. Sources with near neighbors at other velocities are concentrated in a few regions, in particular near $l \approx 332°$ and $l \approx 338°$; the first is near tangency to the Norma Arm, and the second is near the start of the Perseus Arm or the end of the Long Bar (see, e.g., Vallée 2014). We will show that we see evidence for pileup and distance confusion among the 22% of sources that have near neighbors at different velocities, illustrating an intrinsic vulnerability of the method.

5. An Algorithm to Resolve the Near/Far Kinematic Distance Ambiguity

The technique we have used to resolve the near/far kinematic distance ambiguity relies on additional information provided by 21 cm H I data. This approach has been pursued by many studies, and various methods are described in detail by, e.g., Kolpak et al. (2003), Anderson & Bania (2009), Roman-Duval et al. (2009), Brown et al. (2014), and Wienen et al.

(2015) and references therein. The technique relies on the fact that molecular clumps contain cold atomic hydrogen. If a clump is located in front of a bright-enough background source of 21 cm radiation, either from radio continuum or from warm atomic hydrogen-emitting line radiation at the same velocity, this cold atomic hydrogen will absorb the background emission. Thus, the presence of an absorption feature in the H I spectrum indicates that a molecular clump is in front of the background emission, and the absence of an absorption feature indicates that it is behind what is now considered foreground emission. This additional information helps to resolve the near/far kinematic distance ambiguity. Because the technique differs between clumps that are associated with a 21 cm continuum source and clumps that are not, each case (with or without continuum) is described separately. In this study, we attempt to provide a resolution of the kinematic distance ambiguity for each clump in our sample. The reliability of this resolution is difficult to estimate, particularly for clumps with low 21 cm continuum flux intensity; the tabulated results include several variables that may be useful indicators of the robustness of the resolution.

A critical assumption in our approach is that 21 cm continuum emission observed toward a source is assumed to arise from an H II region associated with the clump. We
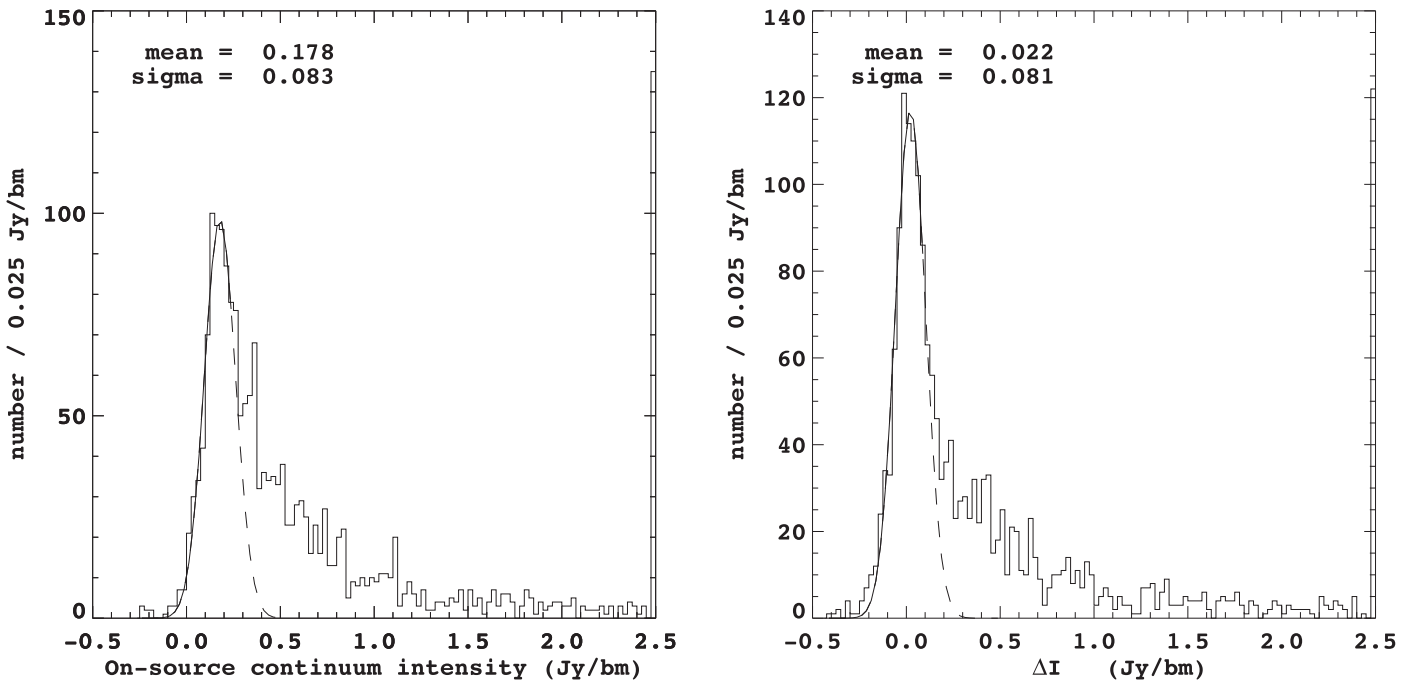
Figure 12. Left panel: distribution in on-source continuum intensity. The smooth curve is a Gaussian fit to the data below $I = 0.25$ Jy beam^{-1}, continued as the dashed curve. Right panel: distribution in ΔI, the difference between the on-source continuum intensity and the average of the off-position intensities. The curve is a Gaussian fit to the data below $\Delta I = 0.10$ Jy beam^{-1}, continued as the dashed line. The text insert in each panel reports the mean and σ of the Gaussian curve.

performed an ON/OFF continuum subtraction (described below) to mitigate the contributions of large-scale continuum emission and identify the continuum emission localized to our sources. We checked for significant positive velocity absorption features toward all clumps associated with 21 cm continuum emission. Their presence would indicate that the continuum source is outside the solar circle and possibly extragalactic. We see no persuasive indications. This method, however, would not detect cases where the 21 cm continuum source is an accidental feature along the line of sight. The best way to verify the association of 21 cm radiation with the molecular clump of interest is to search for recombination-line emission at the same velocity as the molecular gas. When the Southern H II Region Discovery Survey (Brown et al. 2017) is completed, it will provide valuable information for this purpose.

### 5.1. Characteristics of the SGPS Data Set

The SGPS data set (McClure-Griffiths et al. 2005; Haverkorn et al. 2006) includes two sets of files: ATCA +Parkes-with-continuum files, which we will call WC data and use to search for absorption against the continuum, and ATCA +Parkes-continuum-subtracted files, which we will call CS data and which reveal line emission. Figure 10 shows the integrated H I line intensity at three different Galactic latitudes. The upper bound to the integration was at $+50$ km s^{-1}, since the WC spectra near 300° longitude exhibit a discontinuity of roughly 0.2 Jy beam^{-1}near that velocity. The vertical dotted and solid lines mark the boundaries of the individual data cubes. The upper trace is derived from the CS data; the lower trace is derived from the WC data, where we have fit and subtracted the continuum (see below). Comparison of the two traces shows that the WC data have been spatially filtered, suppressing most of the line emission and consequently also the diffuse continuum emission. Our ON/OFF analysis

suggests that the angular scale of this filtering is larger than a degree or so (see below). Figure 11 shows continuum emission (Jy beam^{-1}) as a function of galactic longitude for three different galactic latitudes, where the continuum intensity was measured by averaging the WC amplitude over the velocity interval -207 to -186 km s^{-1} (channels 25–50). The consistency of the continuum level versus both longitude and latitude sustains our conclusion that in this data set, much of the diffuse continuum emission has been filtered out, as otherwise we would expect significantly stronger continuum emission toward the galactic center and spiral arm tangencies and significant latitude dependence (Duncan et al. 1995). The data cube centered on 335° exhibits a significantly stronger typical continuum emission; from comparison of the continuum level with the emission in the overlapping neighboring cubes, we conclude that this represents a different spatial filtering from that of the other cubes rather than a scale factor difference. Our ON/OFF analysis appears to remove the effect of this difference.

The SGPS H I+ continuum spectrum $I_{on}(v)$ toward each MALT90 source position was extracted from the WC data cubes. The spectrum has 599 channels covering the velocity range -227.5 to $+266.3$ km s^{-1}. We model the spectrum as the superposition of two components: $I_{on}(v) = I_{line}(v) + I_C(v)$, where $I_{line}(v)$ represents H I line emission and absorption features appearing primarily near the center of the spectral range and $I_C(v)$ is continuum emission that we assume to be slowly varying across the spectral range. We estimated $I_C(v)$ by a linear fit to the average amplitudes over velocity intervals at the beginning (-207 to -186 km s^{-1}, channels 25–50) and end (226–247 km s^{-1}, channels 550–575) of the spectrum; we expect that over these ranges, all of the 21 cm radiation will be due to continuum. (SGPS spectra for sources in the longitude range $295° < l < 302°.5$ exhibited a discontinuity between the left and right halves of the spectrum; for these, we
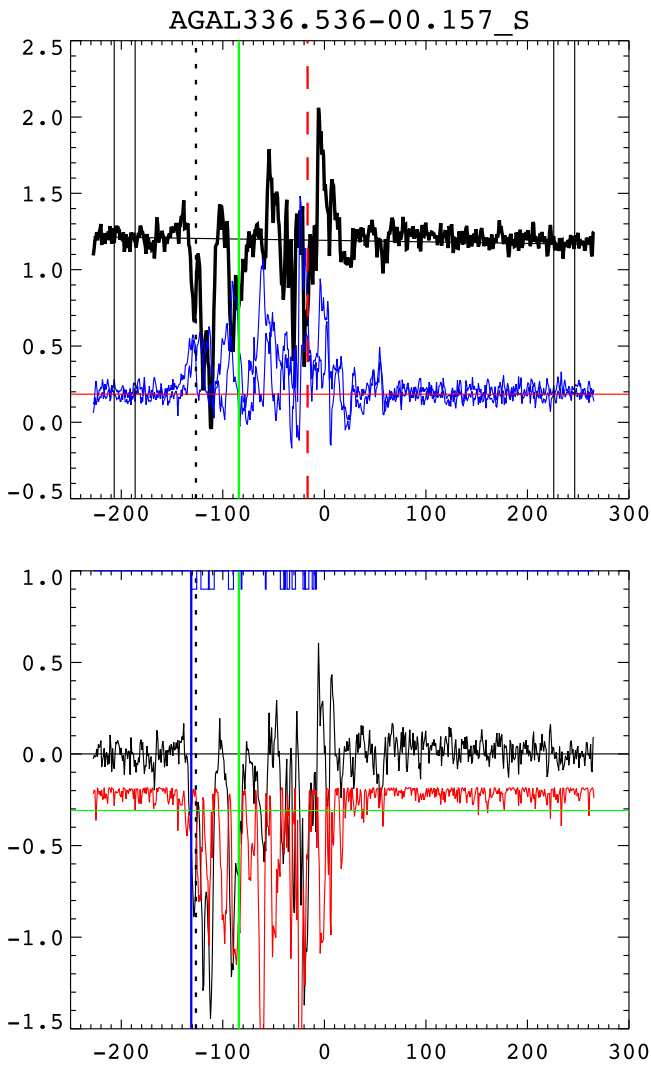
Figure 13. Source with significant continuum emission that exhibits absorption near the tangent velocity and is therefore assigned the far kinematic distance. See text for an explanation of the panels in this and subsequent similar figures.
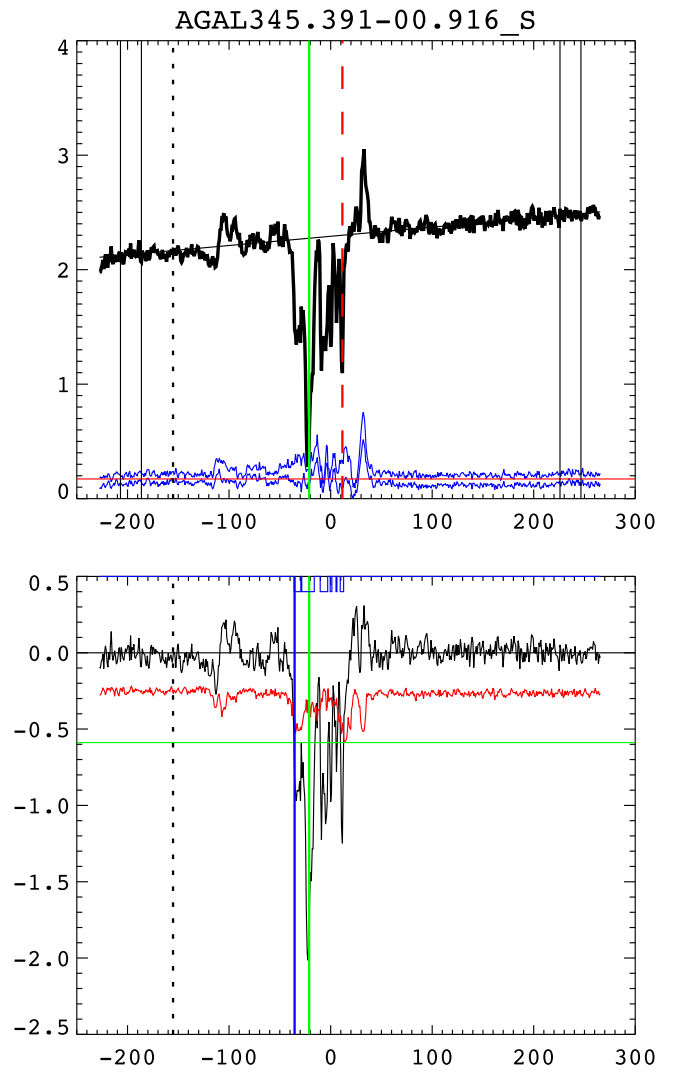
Figure 14. Source with significant continuum emission that does not exhibit absorption near the tangent velocity and is therefore assigned the near kinematic distance.

used a flat line (constant offset) at the low-velocity amplitude.) To characterize the source continuum intensity with a single number, we defined the typical source continuum intensity I_{on_0} as the average intensity over the low-velocity interval, which lies at a more negative velocity than the tangent velocity for the sources in this study. The uncertainty in I_{on_0} was calculated as the rms deviation of the spectrum over the indicated velocity range divided by the square root of the number of channels. The left panel of Figure 12 shows the distribution in I_{on_0}. We considered a spectrum with $I_{on_0} > 0.35$ Jy beam^{-1} ($\sim 2\sigma$ above the mean of a Gaussian fit to the low end of the distribution) to show continuum emission sufficient to pursue the search for intervening absorption against the continuum.

The search for absorption features or a self-absorption signal must take into account the background H I line signal arising from spatially extended sources. To isolate absorption signals arising from a small region around the clump, we need to subtract from the spectrum at the clump position a larger-scale spatial average of nearby extended H I emission. To estimate the background, we extracted the spectra at eight cardinal points at $(l_0 \pm \Delta\theta, b_0)$ and $(l_0, b_0 \pm \Delta\theta)$, where l_0 and b_0 are the longitude and latitude of the clump position and $\Delta\theta = 0°.11$ or

$0°.22$. We will address the choice of these offset values in Section 5.3.

For each cardinal point, we determine a continuum intensity I_{CP_j} using the same procedure as that for the source position. We estimated the OFF spectrum $I_{off}(v)$ of this large-scale background signal by averaging the spectra of all cardinal points for which $0 < I_{CP_j} < 0.35$ Jy beam^{-1}. This criterion is imposed to avoid spurious contamination by an absorption signal against a continuum source in the "off" position. Usually at least three cardinal-point spectra were used, and often all eight; in a few cases of widespread continuum emission, the limits were expanded until at least two cardinal-point spectra passed. The background continuum intensity I_{off0} is the average of the continuum intensities of the participating cardinal points. The difference spectrum is $\Delta I(v) = I_{on}(v) - I_{off}(v)$. We calculate the continuum difference $\Delta I_0 = I_{on0} - I_{off0}$ and its uncertainty $\sigma(\Delta I_0)$ by propagating errors in the usual way. The right panel of Figure 12 shows the distribution in ΔI_0.

The 1σ brightness sensitivity is reported (Table 4 in McClure-Griffiths et al. 2005) as 1.3–2.6 K over the longitude range $295° < l < 350°$ of this study. The spectral intensity for the fourth-quadrant SGPS data is given in units of Jy beam^{-1};
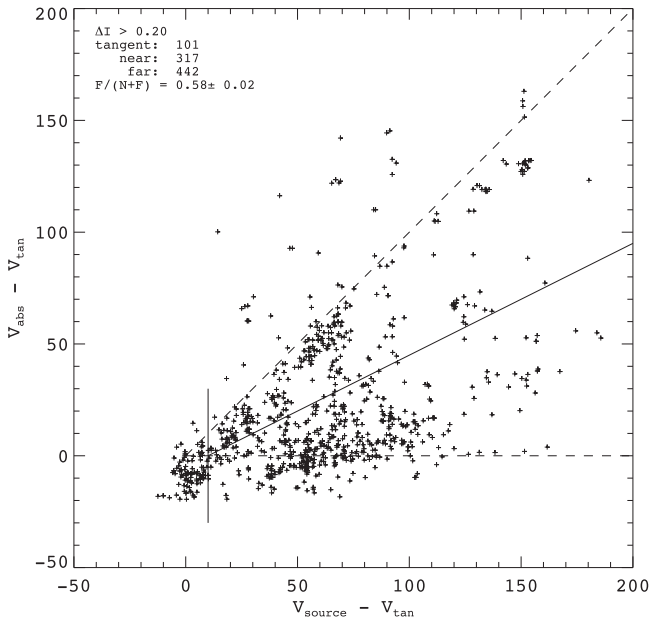
Figure 15. Velocity separation of absorption from the tangent velocity vs. the velocity separation of the source velocity from the tangent velocity. Sources near the horizontal dashed line have absorption near the tangent velocity and are likely far; sources near the diagonal dashed line have absorption near the source velocity and are likely near. Sources near the vertical solid line have velocities near the tangent velocity and are likely near the tangent distance. Sources below/above the diagonal solid line are assigned the far/near kinematic distance in our algorithm.

our analysis was conducted in intensity rather than converting to an antenna temperature scale. The channel-to-channel fluctuations in the data indicated an rms noise level of ~ 0.04 Jy beam^{-1}, consistent with the cited rms brightness temperatures and beam sizes.

We define a continuum source as one having $\Delta I_0 > 0.20$ Jy beam^{-1} and $\Delta I_0/\sigma(\Delta I_0) > 3$; the kinematic distance ambiguity for these sources will be addressed by searching for absorption toward the continuum at velocities between the source velocity and the tangent velocity. Sources failing those requirements are defined to be noncontinuum sources; their kinematic distance ambiguity will be addressed by looking for H I self-absorption at the source velocity. Out of the total of 1905 selected sources, 873 are classified as continuum sources and 1032 are noncontinuum sources.

5.2. 21 cm Continuum Sources

For a clump containing a continuum source that lies at the far distance, the probability is high for detecting an absorption signal against the continuum between the clump velocity and the tangent-point velocity. In fact, if absorbing clouds are sufficiently numerous, an absorption signal should be found at a velocity very near the tangent-point velocity. In contrast, if a clump containing a continuum source is located at the near distance, absorption features should only exist at velocities between the source velocity and $V_{LSR} = 0$. As noted particularly in Anderson & Bania (2009), as an H II region evolves, it expands, and its ionization front moves outward. Unless the H II region has fully emerged from its natal cloud, there should be sufficient neutral atomic hydrogen just beyond the ionization front such that for both near and far sources, absorption should be observed near the source velocity, with a negative offset of up to 10 km s^{-1}. Thus, near and far clumps can in principle be

separated by comparing the velocity of the absorption feature nearest the tangent-point velocity with the tangent and source velocities.

Following the method of Kolpak et al. (2003; see also Wienen et al. 2015), we start from 20 km s^{-1} below the tangent velocity and look for the first absorption feature at higher velocity (i.e., toward the source velocity). The first absorption feature is determined by the first two consecutive channels with absorption that exceeds both 30% of ΔI and 4 $\delta\Delta I$. These criteria were selected to diminish the chance of noise yielding a false absorption indication and to reduce the sensitivity to fluctuations in the OFF cardinal-point intensities. The velocity at which the first absorption occurs is designated as V_{abs}.

Figure 13 illustrates the analysis of a source with significant continuum emission and absorption observed near the tangent velocity, leading to the assignment of the far kinematic distance to this source. Figure 14 illustrates the analysis of a source with significant continuum emission and no absorption observed near the tangent velocity, leading to the assignment of the near kinematic distance to this source. These figures have a common format as follows. The top panel shows the on-source H I + continuum spectrum in bold black and the $\pm 1\sigma$ bounds on the average OFF spectrum in blue. The nearly horizontal black line shows the interpolated continuum emission $I_C(v)$ as measured in the two regions marked by vertical black solid lines. The vertical green line marks the source velocity, the vertical black dotted line marks the tangent velocity, and the vertical red dashed line marks the maximum absorption velocity $V_{max-abs}$. The lower panel plots the continuum-subtracted difference spectrum $(I_{on}(v) - I_C(v)) - (I_{off}(v) - I_{off0})$. The red line marks the -4σ level on this quantity, and the horizontal green line indicates the absorption criterion of reduction in continuum intensity by at least 30% of the on-source continuum intensity. The crenellations in the horizontal blue line mark channels where the absorption criteria are met; the vertical blue line marks V_{abs}, the lowest velocity where there are two consecutive channels showing an absorption signal.

In Figure 15, we plot $V_{abs} - V_{tan}$ on the y-axis versus $V_{source} - V_{tan}$ on the x-axis, i.e., the velocity distance of the absorption feature from the tangent velocity versus the velocity distance of the source from the tangent velocity. Near sources will have absorption near the source velocity and so will lie near (and likely slightly below) the diagonal dashed line. Far sources will have absorption near the tangent velocity, if cold H I clumps of sufficient optical depth are sufficiently numerous along this particular line of sight, and so will lie near the horizontal dashed line; since we are finding the leading edge of the first absorption feature, our method will often place V_{abs} slightly below V_{tan}. (McClure-Griffiths & Dickey (2007) also noted the width of tangent-point H I features due to thermal spread as well as random bulk motions.) Sources with a velocity less than 10 km s^{-1} greater than the tangent velocity (vertical black line) are not measurably separated from the tangent point. The solid diagonal line marks where V_{abs} would be midway between V_{tan} and V_{source}-10 km s^{-1}. We use this line as the near/far dividing line: sources above the line are designated near, and sources below the line are designated far. While there is considerable scatter, we see clear clustering of sources in the vicinity of each dashed line. Some sources can confidently be considered near and some far.

Many effects, however, might confound this analysis. Two leading candidates are signal-to-noise limitations and
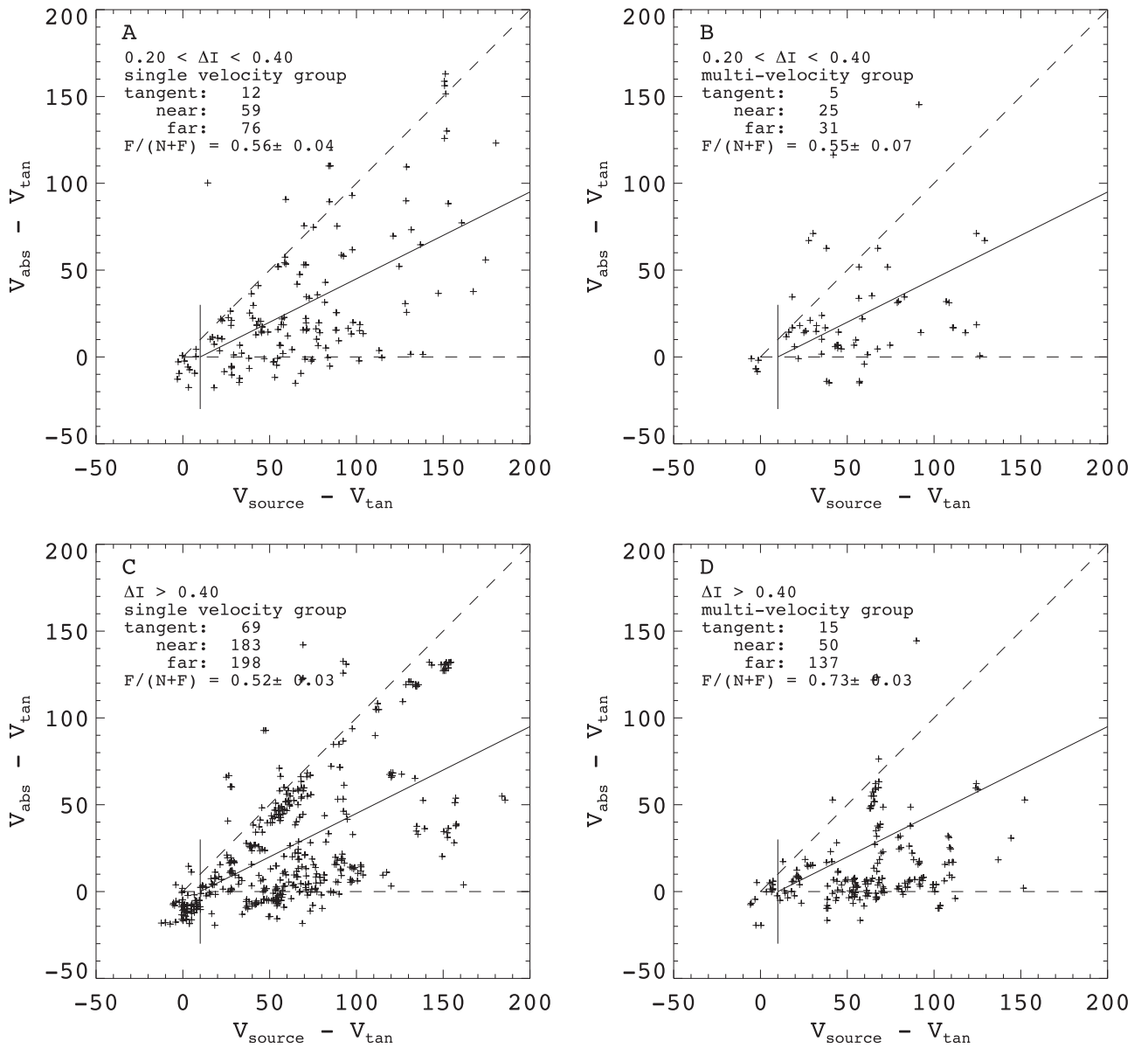
Figure 16. Velocity distance of absorption from the tangent velocity vs. the velocity distance of the source velocity from the tangent velocity for four cuts in continuum intensity and proximity to off-velocity sources.

confusion due to pileup of sources along the line of sight. The absorption-intensity uncertainty $\delta\Delta I(v)$ is the quadrature sum of the error in the ON signal, for which we use the rms measured in the low-velocity line-free region and the rms variation of the cardinal-point OFF intensities. The requirement that observed absorption exceed $4\,\delta\Delta I$ means that for sources with lower continuum emission, clouds with higher optical depth are required; since the frequency of absorbing clouds drops with increasing optical depth (see, e.g., Payne et al. 1983), the probability of an absorption peak falling near the tangent velocity diminishes and the scatter of points in the velocity-differences plot increases. Further, if multiple clumps containing continuum sources are piled up along the line of sight, then two problems may arise: the continuum emission adds and therefore each source appears to have higher continuum emission, and absorption features due to one source may alter the near/far determinations for others. For example, if a source with continuum emission is near the

tangent point and produces a local absorption feature, that feature may cause all other sources along the line of sight, independent of their true individual continuum emission, to be considered continuum emitters with absorption near the tangent point, i.e., to be tagged as far sources.

To grapple with these possibilities, we have used our grouping analysis and divided the data into four subsets, shown in Figure 16. Panel (A) includes sources that fall into single-velocity groups and have ON/OFF continuum differences in the range $0.2 < \Delta I < 0.4$; panel (B) shows sources in multiple-velocity groups in the same ΔI range; panel (C) has sources that fall into single-velocity groups and have $\Delta I > 0.4$; and panel (D) shows sources that are in multiple-velocity groups and have $\Delta I > 0.4$. In each panel, the ratio of far to near+far determinations is indicated. In the lower-intensity cuts, panels (A) and (B), we see greater scatter consistent with our expectations. Comparing the two higher-intensity sets, the multiple-velocity set shows a $>4\sigma$ higher far fraction. While
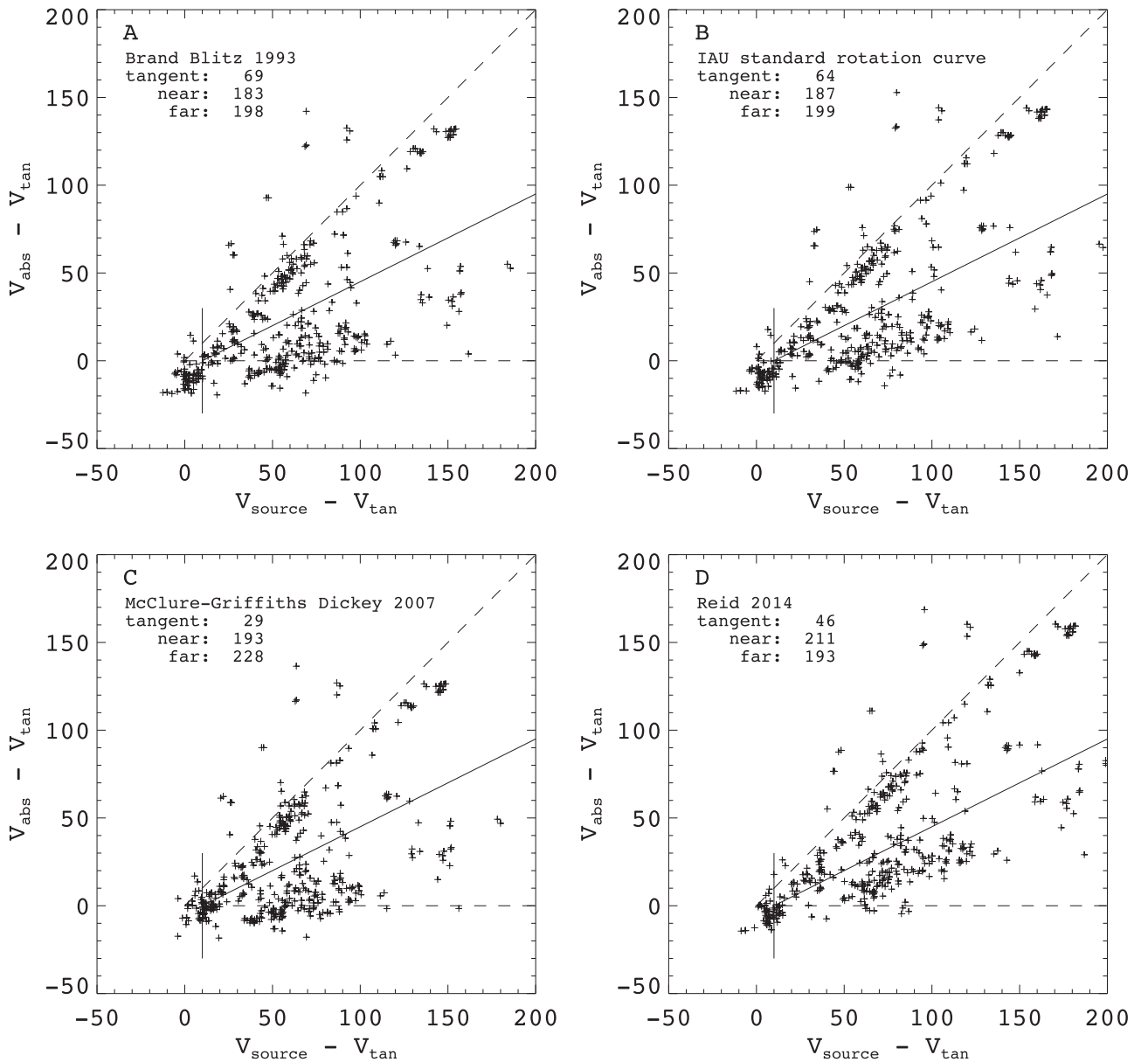
Figure 17. Plots of (absorption velocity–tangent velocity) (y-axis) vs. (source velocity–tangent velocity) (x-axis) in four different rotation models for sources that fall into single-velocity groups and have $\Delta I > 0.4$ Jy beam^{-1}. Panel (A) repeats panel (C) from Figure 16 for ease of comparison.

some of this difference could be real, specific cases show clear evidence for confusion due to pileup. We will use the degree of proximity of the absorption feature to the tangent or source velocity and the presence of nearby off-velocity sources as factors in our estimation of the quality or reliability of the distance ambiguity resolution for each source.

Figure 17 repeats the analysis of the high-intensity, single-velocity fourth-quadrant data sample for three rotation models in addition to the Brand–Blitz model. Panel (A) repeats panel (C) from the previous figure for easy comparison. The distribution of sources is most consistent with the model expectations, i.e., that for far or tangent-point sources, the absorption velocity should lie near or below the tangent velocity for the Brand & Blitz (1993) model or the McClure-Griffiths & Dickey (2007) model. As noted previously, the McClure-Griffiths & Dickey (2007) rotation curve is valid only for 3 kpc < R < 8 kpc; for this reason, we have used the Brand & Blitz (1993) model in this analysis.

5.3. Non–21 cm Continuum Sources

In the absence of strong continuum sources, the spectrum is dominantly due to H I line emission from warm atomic hydrogen that is ubiquitous in the galactic plane modulated by absorption by cold H I clumps. A cold clump at the near kinematic distance will resonantly absorb at the source velocity the line emission arising from warm H I at the far position, and the width of the absorption feature will be comparable to the width of the molecular-line emission from the source. A cold clump at the far kinematic distance will not create such an absorption signal; in fact, emission from warm H I at the near position will fill in the H I emission profile. To separate clumps at the near distance from those at the far distance, then, we look for a significant and relatively narrow H I self-absorption (HISA) signal at the source velocity. We expect that a cold clump at the near position will be spatially localized, so we look for absorption in comparison to the line signal at positions slightly offset in longitude and latitude. We
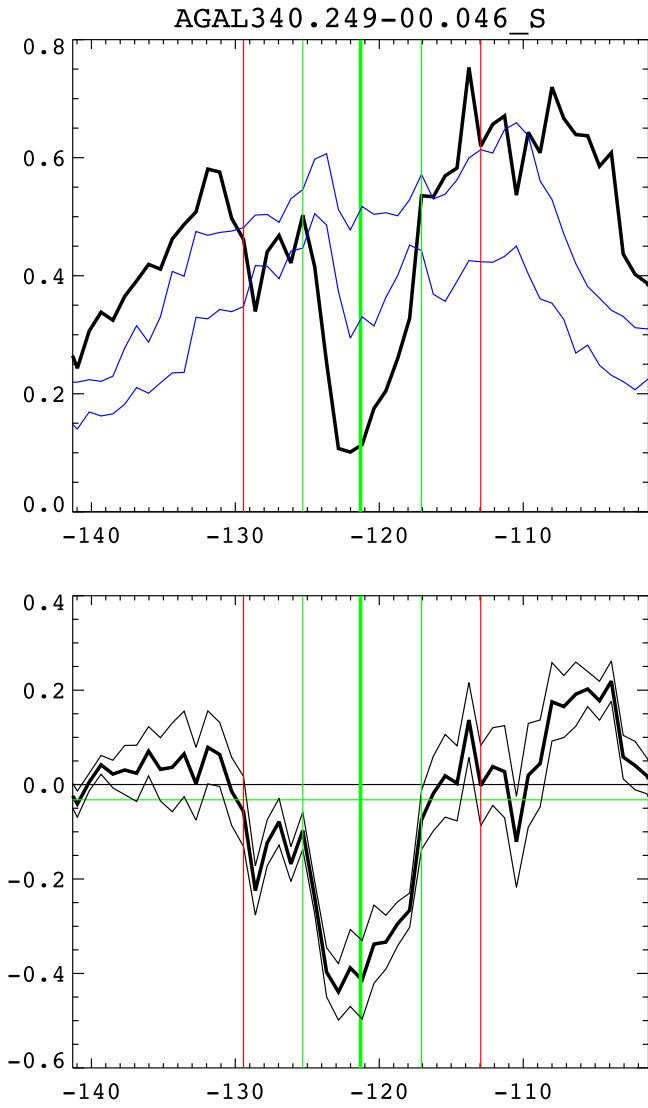
Figure 18. Example of a source that does not have significant continuum emission and is identified as near due to a strong self-absorption dip at the source velocity.

work with the differential line intensity calculated as the (ON intensity minus ON continuum) minus the average (OFF intensity minus OFF continuum). See Kavars et al. (2003) for a discussion of the radiative transport issues. We define the "HISA measure" as the average intensity per channel in a narrow window centered on the source velocity subtracted from the differential line intensity per channel averaged over two "wing" windows abutting the on-source velocity windows. The units of the HISA measure are consequently Jy beam^{-1} channel^{-1}, and a negative HISA measure indicates the presence of self-absorption. The width of the on-source velocity and wing windows over which the HISA measure is calculated is adapted to the line width of the observed molecular-line emission: the on-velocity average is taken over channels $j_V \pm N_{hw}$, where j_V is the channel in which the source velocity falls, and the wing average is taken over the next $N_{hw} + 1$ channels on each side, with

$$N_{hw} = \frac{\Delta V}{2\delta V},$$
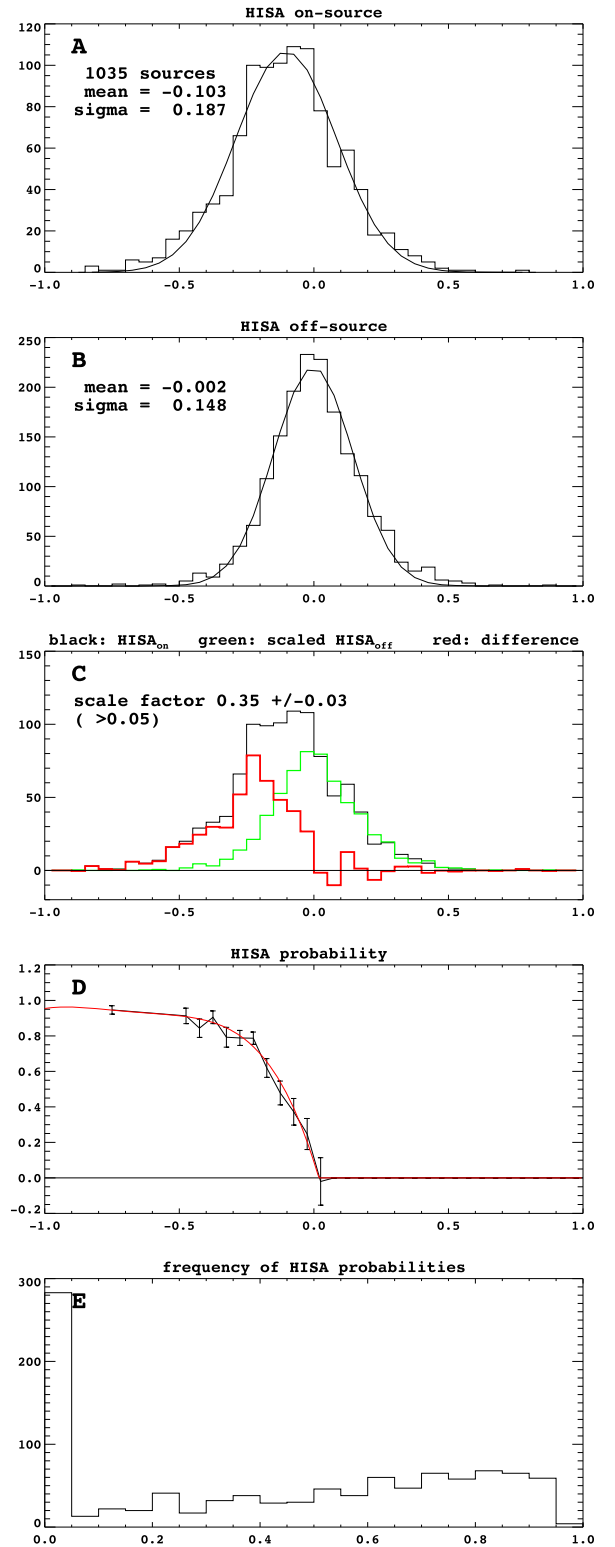
Figure 19. A: Distribution in HISA measure observed on-source. B: Distribution of HISA measure observed when source longitudes are offset by $-1°0$. C: Modeling the on-source distribution (A; black histogram) as the sum of a no-absorption component due to far sources, using the longitude-offset HISA distribution as a proxy (green histogram; scaled to match (A) beyond HISA = 0.05) and a residual distribution (red histogram) due to near sources. D: Derived near probability as a function of HISA value. The red curve is a polynomial fit used to assign a HISA probability to each source. E: Distribution in near probabilities for the noncontinuum source sample.

where δV is the SGPS channel width ($0.82\,\mathrm{km\,s^{-1}}$) and ΔV is the maximum of the velocity FWHM of the HCO^+ (1–0), HNC (1–0), and N_2H^+ (1–0) lines (or the HCN (1–0) value if none of those was detected). We constrained $2 \leqslant N_{hw} \leqslant 10$, corresponding to a window of width from 4 to $17\,\mathrm{km\,s^{-1}}$ within which the HISA signature was sought. An example of a source with a strong self-absorption signature is shown in Figure 18. As before, the top panel shows the on-source H I + continuum spectrum in bold black and the $\pm 1\sigma$ bounds on the average OFF spectrum in blue; the lower panel plots the continuum-subtracted difference spectrum ($I_{\mathrm{on}}(v) - I_C(v)) - (I_{\mathrm{off}}(v) - I_{\mathrm{off0}})$ in bold black with the $\pm 1\sigma$ bounds in light black. Now, however, this is a zoom into the region about the source velocity, marked by the bold green vertical line. The region over which the on-velocity average is taken is marked by the thin green lines; the wings over which the off-velocity average is taken are bounded by the thin green and red lines.

The distribution in HISA measure for on-source difference spectra is shown in panel (A) of Figure 19. We model this distribution as the sum of two components: the distribution of HISA measure for far sources, which we expect to average zero and have a width reflecting spectral fluctuations and noise, and the distribution of HISA measure for near sources, where the average value should be negative, since all near sources should have at least some absorption. To get a handle on the two components, we measure the HISA distribution when all source longitudes were offset by $-1°0$; this simulates a measurement of the HISA measure at essentially random positions in the Galactic plane. The offset HISA distribution is plotted in panel (B) of Figure 19. Assuming that far sources will have the same spectral fluctuations and noise as the offset sample, this gives us the shape of the HISA measure distribution for far sources. We scale this distribution to the signal HISA distribution for HISA > 0.05 and subtract it to derive the HISA distribution for the near sources. This is illustrated in panel (C) of Figure 19, where the black histogram is the signal HISA distribution as in the top panel, the green histogram is the scaled offset histogram, and the red histogram is the residual distribution; we interpret the red histogram as the distribution in HISA measure for near sources. (Note that the scale factor of 0.35 does not mean that 35% of the signal distribution should be considered far; the offset analysis is done for all source positions, but there is a small fraction of continuum sources in the offset sample, so more sources pass through the noncontinuum analysis path than the on-source analysis.)

We also tried to measure the no-absorption HISA distribution by using a control region offset in velocity by $20–30\,\mathrm{km\,s^{-1}}$ from the signal region in each spectrum. In the offset longitude analysis, both the signal region HISA distribution and the control region HISA distribution were centered at zero, but the width of the control distribution was significantly less than the width of the on-velocity distribution. We concluded that the longitude–velocity correlations of our sources make the velocity offset approach unreliable, while the good agreement within statistics between the positive tails of the HISA distributions for the on-source and offset longitude samples offers some support for using spatial offsets instead.

We used the ratio of the mean of the signal HISA distribution divided by the Gaussian standard deviation of the offset distribution as a figure of merit for tuning the analysis parameters. Methodical exploration of the parameter space led us to the best-choice set of cardinal-point offsets and HISA width parameters described previously. If we moved the cardinal points closer, or narrowed the velocity range for the HISA measure, then we were subtracting signal, and the figure of merit decreased. If we moved the cardinal points farther or increased the velocity range for the HISA measure, then we were including more noise, and the figure of merit again decreased.

We have now decomposed the HISA distribution to plausible far and near components, represented by the green and red histograms in panel (C) of Figure 19. For a source with a particular HISA measure, then, the probability P_{HISA} that the source has a significant self-absorption signature and therefore is near will be the ratio of the residual (red histogram) amplitude to the full (black histogram) amplitude at that HISA value. The P_{HISA} probability as a function of HISA measure is plotted in panel (D) of Figure 19, where the red line is a polynomial fit used to interpolate the data points. Panel (E) shows the distribution in P_{HISA} for the 1032 sources passing through this analysis path. When P_{HISA} is greater than 50%, we consider the source to be at the near kinematic distance, and when the probability is less than 50%, we consider it to be far. The difference in the probability from 50% will contribute to our estimation of the quality or reliability of the distance ambiguity resolution.

6. Assignment of Kinematic Distances

For each source with a valid velocity determination, kinematic distances were calculated per the equations given in Section 2.1 using both the Brand–Blitz and the McClure-Griffiths–Dickey rotation curves. Further, because many sources have velocities lying near the tangent velocity, we have calculated distances with and without masking velocities near the tangent velocity and assigned those sources to the tangent point. The four panels in Figure 20 show the source positions in the Galactic plane for these four analysis tracks. The red symbols indicate the nominal source position, and the green lines indicate the variation in distance from the LSR when the source velocity was dithered by $\pm 7\,\mathrm{km\,s^{-1}}$. Sources for which the galactocentric radius lay beyond the range of validity (3–8 kpc) of the McClure-Griffiths–Dickey model were omitted in the lower two panels. In some cases, the dithering of the source velocity to calculate higher and lower distance ranges resulted in a galactocentric radius that fell outside the McClure-Griffiths–Dickey validity range; in those cases, the higher distance error bar extends off the panel.

Figure 21 again displays MALT90 sources on a longitude–velocity (l–v) diagram superposed on the CO 1–0 emission integrated from $-1° < b < +1°$ from the Columbia-CfA survey (Dame et al. 1987, 2001). The sources for which a near kinematic distance is preferred are shown in blue, those for which a far kinematic distance is preferred are shown in red, and those assigned to the tangent point are shown in green.

Figure 22 presents another view of the MALT90 sources, this time plotted in longitude and latitude with symbol shapes indicating which analysis track was used and color indicating the distance assignment: green for the tangent point, blue for near, and red for far. Sources that could not be assigned a distance, either because their velocity was incompatible with
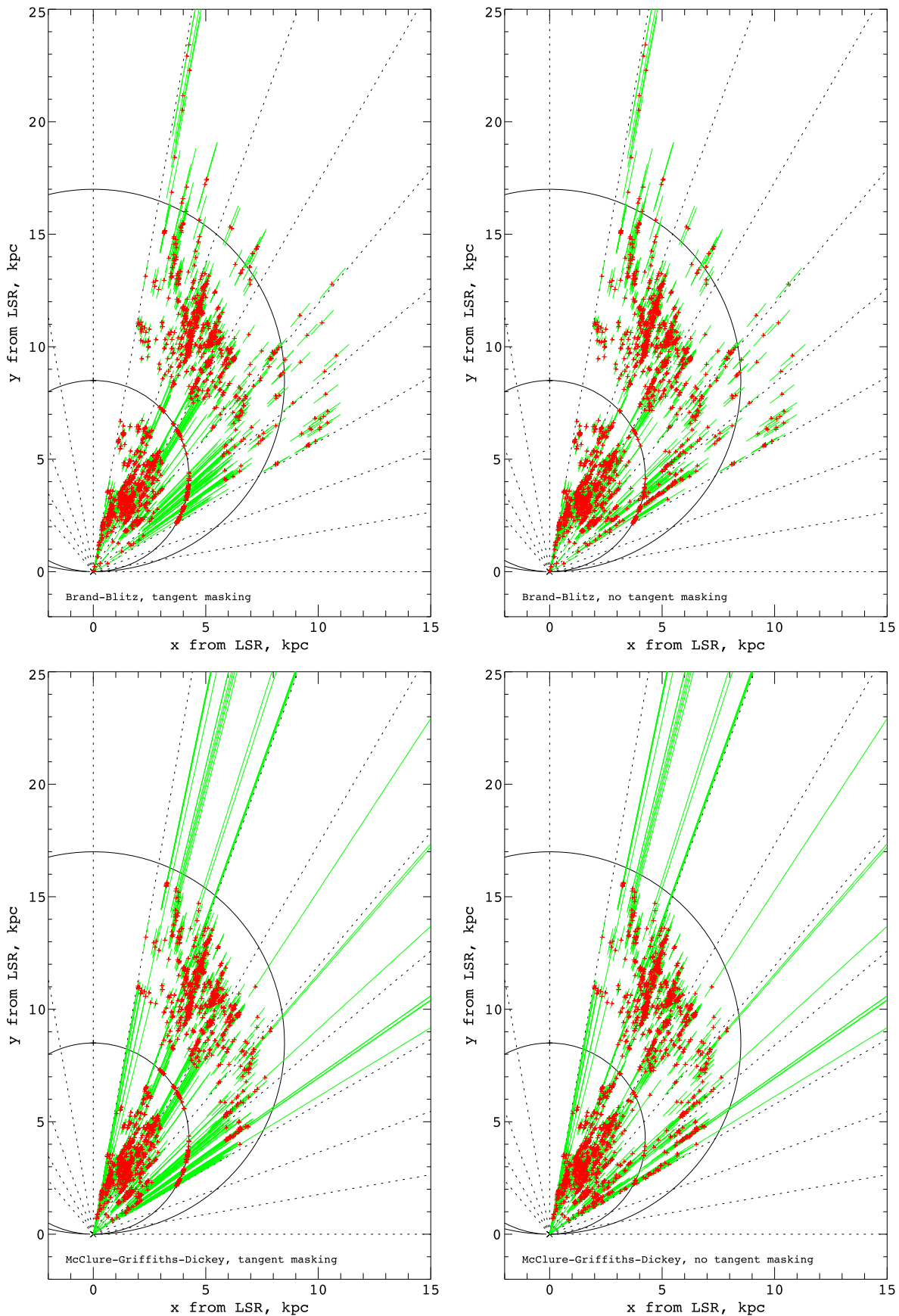
Figure 20. Plots of source locations in the Galactic plane. The top panels show the source locations calculated with the Brand–Blitz rotation curve. In the left panel, sources with velocities within 10 km s^{-1} of the tangent velocity are assigned to the tangent point, while in the right plot, no such masking is done. The lower two panels show locations determined with the McClure-Griffiths–Dickey rotation curve with (left) and without (right) masking near the tangent velocity. See the text for discussion of the error bars.
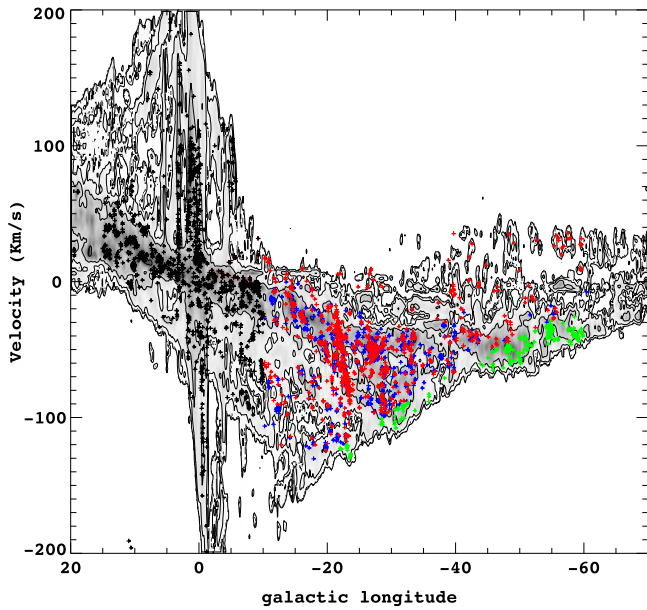
Figure 21. LSR velocity (*y*-axis) vs. Galactic longitude (*x*-axis) for all MALT90 targets from Rathborne et al. (2016). The points represent the MALT90 targets. The background gray scale represents the emission from CO 1–0 integrated from Galactic longitude $-1° < b < +1°$ from the CfA-Columbia survey (Dame et al. 1987, 2001). Red points indicate sources believed to lie at the far kinematic distance, green points are assigned to the tangent distance, blue points indicate sources more likely to be at the near kinematic distance, and black points are sources where no kinematic near/far disambiguation has been made.

the rotation model or because no H I data was available at that position, are shown in black.

6.1. Distance Catalog

The results of this analysis are presented in Table 1, where the Brand–Blitz rotation model has been used. These distances were calculated without applying any tangent-point masking, but since the tangent velocity and distance are included for each source, users can apply their own preferred tangent-masking algorithm. The upper and lower ranges for the near and far kinematic distances were calculated by varying the source velocity by $\pm 7 \, \mathrm{km \, s^{-1}}$; they correspond approximately to 1σ variations, except in those cases where the velocity variation passes beyond an algorithmic limit (such as the velocity becoming positive or exceeding the maximum possible velocity). All distances are capped at 33.33 kpc.

For each source, Table 1 gives the probability P_{HISA} (described in Section 5.3) that a narrow self-absorption signature was observed. The P_{HISA} algorithm was trained on noncontinuum sources. As noted in Section 5.2, in expanding H II regions, the local self-absorption may be shifted by as much as $-10 \, \mathrm{km \, s^{-1}}$ from the source velocity, which would diminish the HISA measure and the calculated value of P_{HISA}.

Table 1 also includes two variables for each source, "dcode" and "dflag," that report the analysis path and ambiguity resolution determination. The variable "dqual" is an integer from 0 to 9 with a larger value indicating higher confidence in the reliability of the near/far distance resolution. We consider the ambiguity resolution for sources with dqual less than 4 to be dubious.

6.2. Comparison with Other Distance Determinations

Many other groups have performed similar analyses to determine kinematic distances and resolve the kinematic distance ambiguity. The prior analysis that has the most significant overlap with the sources in this analysis is the work of Wienen et al. (2015; hereafter Wienen). That analysis started with ATLASGAL targets and collected velocity information from observations of ammonia, as well as from prior observations. In the fourth Galactic quadrant, Wienen used the SGPS H I data sets and studied H I absorption to address the kinematic distance ambiguity, as does the present analysis. Both analyses used the Brand–Blitz rotation curve to determine kinematic distances. One potentially significant difference between the analyses is that Wienen calculated tangent velocities for the range $300° < l < 340°$ using the McClure-Griffiths–Dickey rotation curve, while the present analysis determined tangent velocities from the Brand–Blitz curve.

For each MALT90 source, we searched for the closest Wienen source in angular separation; a separation of less than $0°.005$ was considered a positional match. We found matches for 557 of the 1895 MALT90 sources in our selected longitude range. The agreement in velocities was good: a Gaussian fit to the MALT90 velocity minus the Wienen velocity yielded a mean value of $-0.5 \, \mathrm{km \, s^{-1}}$ and σ of $0.4 \, \mathrm{km \, s^{-1}}$, so no further selection in velocity was made. The nonzero mean difference may be due to the fact that the velocities for the HCO$^+$ (1–0) and HNC (1–0) transitions used in the consensus velocity calculation are often slightly blue compared to the N$_2$H$^+$ (1–0) lines due to infall and optical depth effects. Such a blue–red asymmetry infall characteristic was suggested for MALT90 sources in Rathborne et al. (2016).

We considered the MALT90 and Wienen kinematic distances to be in agreement if the absolute value of the difference was less than 2 kpc. For the entire sample, we agreed in 70% of the cases. The MALT90 analysis found 43% of the sources to be at the far distance, while for Wienen, the far fraction was 29%. The match sample included 11 sources with positive velocities; the MALT90 analysis placed all of these at the far distance, while in the Wienen analysis, three of them were placed at the tangent or near distance. Forty-seven sources were found in the MALT90 analysis to have velocities exceeding the tangent velocity and so were placed at the tangent distance; all of these sources were placed by the Wienen analysis within 2 kpc of the tangent distance and so were considered to agree. We divide the remaining 499 cases into three categories. The first category includes 239 sources without associated continuum emission that followed the HISA analysis track. The remaining sources, those with associated continuum emission that were analyzed for absorption against the continuum, were divided into 203 that fell in a cluster with only one velocity subgroup in the longitude–latitude–velocity clustering analysis and 57 that fell in clusters with multiple-velocity subgroups. Agreement on the kinematic distance resolution was found in 63% of the HISA cases, 74% of the single-velocity continuum cases, and 65% of the multiple-velocity continuum cases. Requiring the distance quality estimator dqual to be greater than 4 reduced the HISA cases to 65 and increased the agreement to 70%, consistent with the difficulty in determining the presence of absorption in noisy difference spectra.

Expectations for the fraction of clumps found at the far kinematic distance depend on assumptions about the clump Galactic spatial distribution and the distribution of the clumps'
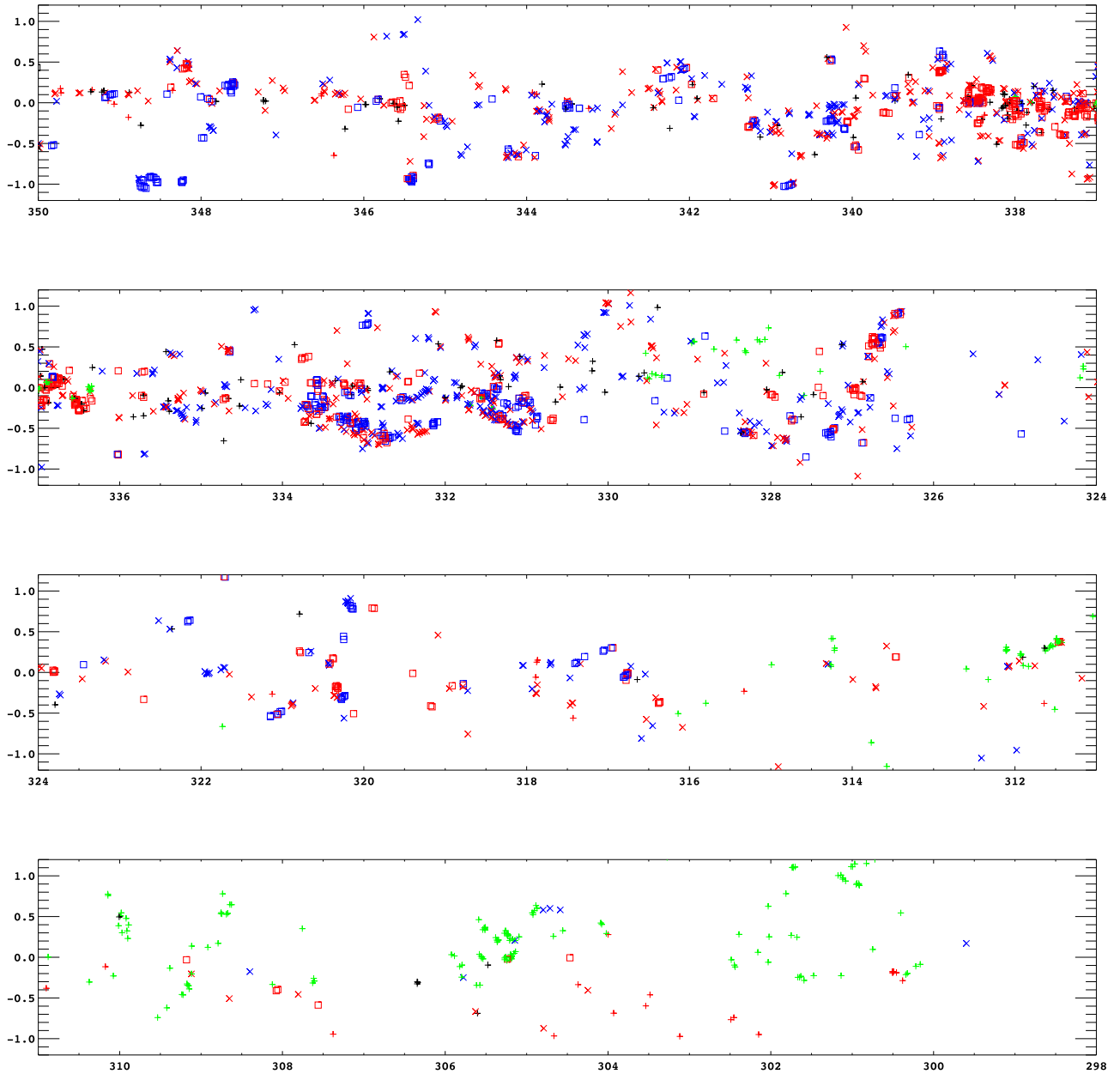
Figure 22. MALT90 sources by longitude and latitude for the longitude range 350° (upper left) to 298° (lower right). Sources analyzed as continuum emitters are indicated by open squares; those analyzed as noncontinuum emitters are indicated by crosses. Sources assigned to the tangent point are in green, those more likely to be at the near position are in blue, and those more likely to be at the far position are in red. Sources for which no kinematic distance ambiguity resolution was possible are in black.

870 μm luminosity, which in turn depends on the dust column density, temperature, and emissivity. At any longitude, the tangent point lies halfway to the solar circle; the locus of tangent points is a circle passing through the Galactic center, as shown in Figure 23. In any interval of longitude in the first or fourth Galactic quadrants, the area between the tangent circle and the solar circle (far distance) is three times the area within the tangent circle (near distance). If the sources were uniformly distributed within the solar circle (with a negligible fraction beyond the solar circle), the ATLASGAL survey was complete to the solar circle, and the subset of ALTASGAL clumps selected for the MALT90 survey was unbiased, then one would expect 75% of the sources to be far.

The selection of clumps from the ALTASGAL catalogs (Contreras et al. 2013; Urquhart et al. 2014) for the MALT90 survey was discussed in Rathborne et al. (2016). The ATLAS-GAL 870 μm peak flux sensitivity of 0.25 Jy corresponds to a mass of 200 $M_\odot$ for a clump with a dust temperature of 10 K at a distance of 10 kpc (Schuller et al. 2009). MALT90 used this flux sensitivity limit and additional criteria on the clump infrared emission to select clumps for the survey. If we assume that the MALT90 sources are uniformly distributed in the Galactic disk and consider only clumps with masses greater than 200 $M_\odot$, then the MALT90 sample is complete to 10 kpc (indicated by the dash-dotted line in Figure 23). For this 10 kpc limit, the ratio of areas suggests that roughly 54% of the sources should be far.

Realistic assumptions about the distributions in clump morphology, mass, and temperature confound this simplistic picture. Clumps with higher mass or higher temperature will be visible at a greater distance (Malmquist bias), increasing the

Table 1
Results of the Kinematic Distance Analysis

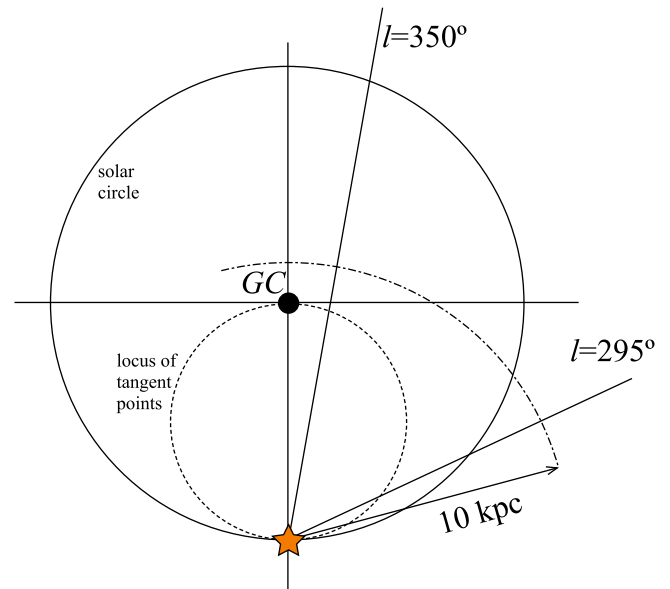| Name	Explanation
Name	Clump name
GLON	Galactic longitude
GLAT	Galactic latitude
Rvel	Radial velocity
Tvel	Tangent velocity
Rgal	Galactocentric radius
Kdt	Kinematic distance to tangent point
Kdn	Near kinematic distance
b_Kdn	Lower range to near kinematic distance
B_Kdn	Upper range to near kinematic distance
Kdf	Far kinematic distance
b_Kdf	Lower range to far kinematic distance
B_Kdf	Upper range to far kinematic distance
PHISA	Probability of a HISA absorption dip
Mabsv	Minimum velocity of significant absorption
Dcode	Analysis process flag
Dflag	Kinematic distance ambiguity resolution flag
Dqual	Integer flag 0–9 (higher = better) for near/far ambiguity resolution
Dbest	Best guess for kinematic distance
b_Dbest	Lower range of best guess for kinematic distance
B_Dbest	Upper range of best guess for kinematic distance

(This table is available in its entirety in machine-readable form.)

Figure 23. In this cartoon of the Galactic disk, the lines at 295° and 350° show the longitude interval addressed in this study. The dashed circle passing through the LSR (star) and the Galactic center (black dot) is the locus of the tangent points; clumps within this circle are at the near kinematic distance, while sources beyond it are at the far kinematic distance. The dash-dotted line indicates a distance of 10 kpc, within which the MALT90 source sample is complete for clumps with masses greater than 200 $M_\odot$.

fraction of far sources. In contrast, nearby clumps may be resolved into several constituent components, increasing the fraction of near sources. Furthermore, the 870 μm luminosity function for clumps is poorly constrained. Finally, since dense, star-forming clumps are almost certainly associated with spiral arms, they are unlikely to be uniformly distributed in the Galactic disk, and simple geometric arguments may not apply. Given all of these uncertainties, we conclude that our observed far fraction of 46% is reasonable based on the 10 kpc completeness limit for 200 $M_\odot$ clumps, and we anticipate that the results from this study will inform ongoing efforts to understand the structure of the Milky Way.

6.3. Implications for Galactic Structure

Many attempts have been made to model the spiral arm structure of the Milky Way based on atomic hydrogen, molecular-line data, dust continuum, and recombination-line data (e.g., Vallée 2014, and references therein). Many of these models suggest that the Milky Way's spiral structure is dominated by four major spiral arms: Norma, Perseus, Crux–Scutum, and Carina–Sagittarius. In the fourth-quadrant region discussed in this paper, only the Norma and Crux–Scutum arms will be seen in tangency: Norma at $l \sim -32°$ and Crux–Scutum at $l \sim -52°$, corresponding to galactocentric radii of $\sim$4.6 and 6.6 kpc (Vallée 2014). In Figure 22, the MALT90 sources assigned to the tangent-point distance in fact show distinct clusters at these longitudes, in satisfactory agreement with these models.

To assess the peaks in the histogram of galactocentric radius in Figure 7, we present in Figure 24 plots of galactocentric radius, latitude, and velocity versus longitude with the same color coding (as indicated in the inset). The histogram suggests that each of the peaks lies above a smooth underlying distribution of comparable number count, and we see in

Figure 24 that the tagged sources lie in fairly continuous bands with some denser aggregations that are likely to indicate individual complexes.

Because of the random velocity component of clouds, the overlap of the spiral arm features in l–v space due to Galactic rotation, and uncertainties in the Galactic rotation curve, it is difficult to definitively determine spiral arm structure from l–v diagrams. Since the MALT90 targets are dense clumps that will likely spawn high-mass stars, they should be excellent tracers of spiral arms, since such dense clumps presumably form by compression due to the passage of a spiral shock. A detailed analysis of the correspondence of the MALT90 clumps to the theoretical locations in l–v space for various spiral arm models awaits further analysis outside the scope of this paper.

The MALT90 data, however, do reveal some tantalizing hints about the spiral structure at large distances in the fourth quadrant. For example, for many of the MALT90 sources, positive LSR velocities are measured. In this case, no near-far ambiguity exists, and these sources are unambiguously located beyond the solar circle. These sources can be seen in the top two panels of Figure 20 in a cluster at $x \sim 10$ kpc and y ranging from $\sim$5 to 13 kpc. This cluster matches the expected location of the Carina–Sagittarius spiral arm. We illustrate this in Figure 25, where the full MALT90 sample with kinematic distances is superposed on the iconic representation of the Milky Way structure prepared by R. Hurt (NASA/JPL-Caltech/R. Hurt [SSC-Caltech]).

In addition, many MALT90 sources are placed at the far kinematic distance with high probability. Thus, if our near-far disambiguation is correct, a significant number of MALT90 targets are probing distant portions of the Galaxy, presumably lying on the far sides of the Crux–Scutum, Norma, and Carina–Sagittarius arms. Indeed, a few of these sources have positive velocities and may well represent the heretofore poorly
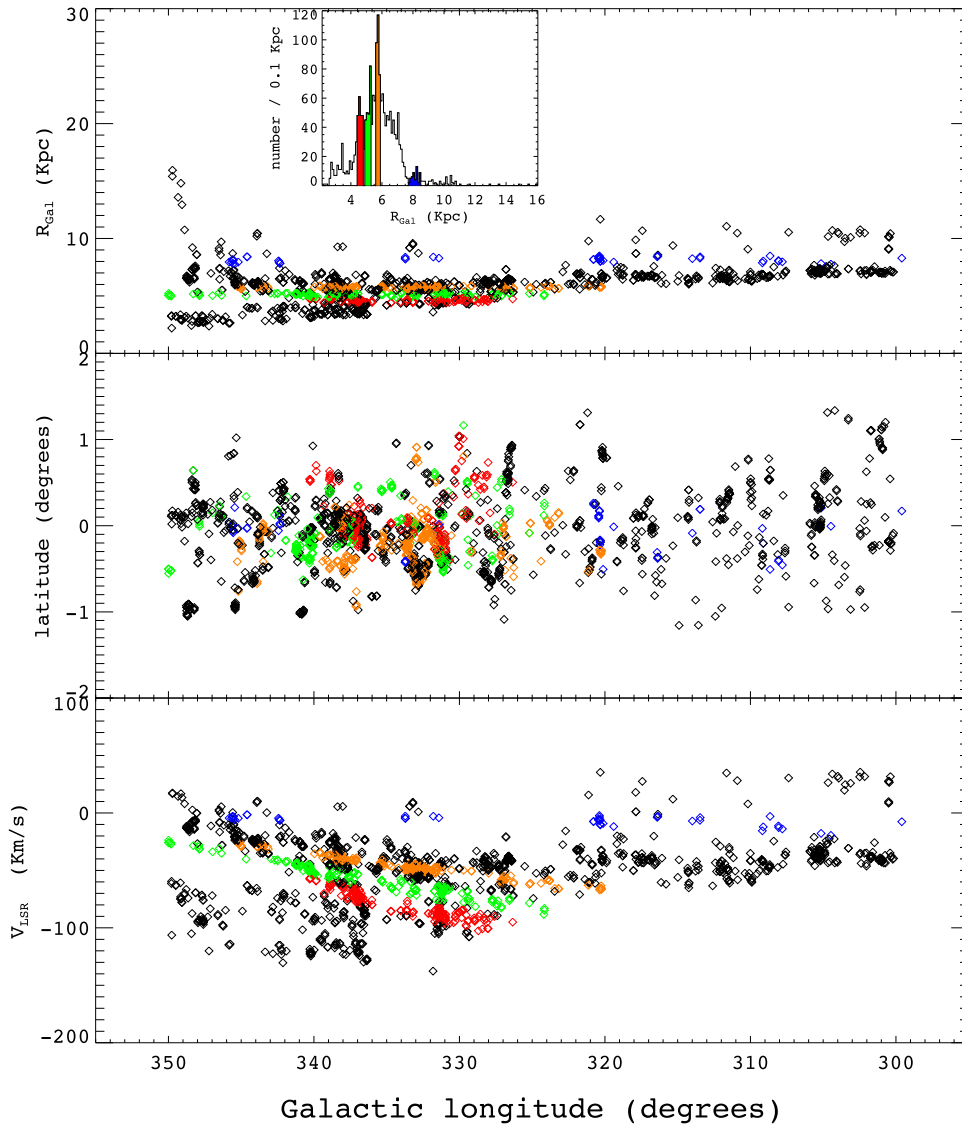
Figure 24. Scatter plots of galactocentric radius, latitude, and velocity vs. longitude. The symbol colors tag the peaks observed in the galactocentric radius histogram shown in Figure 7, which is repeated in the inset. See text for discussion.

explored far side of the Crux–Scutum spiral arm. To be detected at our sensitivity level, such sources must be extraordinarily luminous and probably quite massive. They may well represent interesting candidates for the progenitor clumps of young massive clusters. The search for such extreme clumps is discussed in a companion paper by Contreras et al. (2017). In Anderson et al. (2015), the WISE project reported the observation of H II regions in the outer Crux–Scutum arm.

7. Summary

Using molecular-line data from the MALT90 survey, we have established kinematic distances to 1905 dense molecular clumps. We first established a consensus velocity based on a flux-weighted average of the HCO^+ (1–0), HNC (1–0), and N_2H^+ (1–0) lines toward the central position of ATLASGAL 870 μm continuum targets. We then established the galactocentric radii using two rotation curves: Brand & Blitz (1993) and McClure-Griffiths & Dickey (2007).

The major difficulty in assigning kinematic distances is to resolve the near-far kinematic distance ambiguity. We use 21 cm atomic hydrogen data from the SGPS to resolve the ambiguity and assign a distance. Under the assumption that cold atomic hydrogen is associated with each molecular clump, the presence or absence of H I absorption or self-absorption signals can resolve the ambiguity. If the clump contains an H II region producing 21 cm free–free continuum, multiple absorption signals should arise from intervening clouds. In this case, if the clump is located at the near kinematic distance, the absorption signals must have velocities only between zero and the clump velocity. In contrast, if the clump is located at the far kinematic distance, absorption signals with velocities between the source velocity and the tangent velocity should also arise. To implement this test, we first determine which MALT90 sources are associated with significant 21 cm continuum emission, which we assume to arise from an embedded H II region associated with the source. For these sources, we then discriminate between the near and far distances by determining
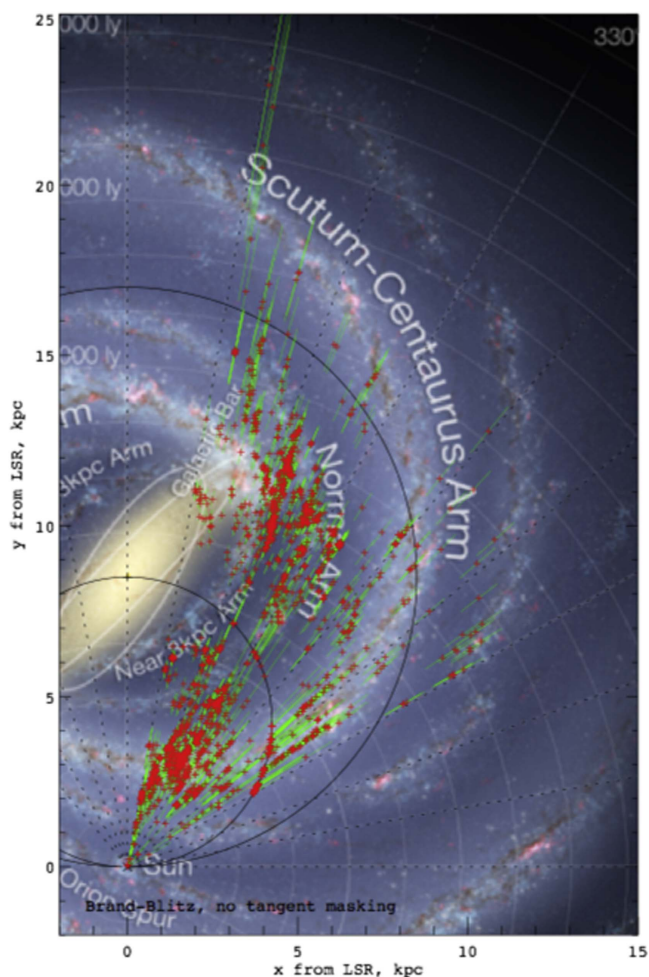
Figure 25. MALT90 sources with kinematic distances (upper left panel of Figure 20) superposed on the iconic Milky Way representation by R. Hurt (NASA/JPL-Caltech/R. Hurt [SSC-Caltech]).

the velocities of the absorption features nearest the tangent velocity and the source velocity.

For clumps without associated continuum emission, there will always be background emission at the source velocity from the ubiquitous warm H I at the far kinematic distance. If a clump lies at the near distance, its associated cold H I will produce a self-absorption signal, but if it lies at the far distance, it will not. Thus, for sources without associated continuum emission, we assign those with H I self-absorption to the near distance and those without to the far distance.

We have completed a thorough study of parameter space to select the best parameters for our algorithms. We have also used the 21 cm data at locations not associated with MALT90 targets to estimate the degree of false positives due to random fluctuations in the H I background signal. From these studies, we can estimate the probability of an accurate assignment of the near or far distance. Thus, we have built estimates of the degree of confidence for our assigned values into our distance determination. The results are presented in a tabulated catalog.

These distances should be useful to the community for establishing the general properties of molecular clumps and determining their locations, sizes, masses, and luminosities. Such an analysis, along with a search for possible progenitor clumps of young massive clusters, is presented in a companion

paper (Contreras et al. 2017). Moreover, the distances will be useful for studies of Galactic spiral structure. We have found, for example, good evidence that many MALT90 clumps seem to delineate distant portions of the Crux–Scutum, Norma, and Carina–Sagittarius spiral arms.

ORCID iDs

J. Scott Whitaker ⓘ https://orcid.org/0000-0001-9179-6312
James M. Jackson ⓘ https://orcid.org/0000-0002-3466-6164
J. M. Rathborne ⓘ https://orcid.org/0000-0002-3427-9092
J. B. Foster ⓘ https://orcid.org/0000-0001-7110-2192
Y. Contreras ⓘ https://orcid.org/0000-0002-6388-3635
Patricio Sanhueza ⓘ https://orcid.org/0000-0002-7125-7685
Ian W. Stephens ⓘ https://orcid.org/0000-0003-3017-4418
S. N. Longmore ⓘ https://orcid.org/0000-0001-6353-0170

References

Aguirre, J. E., Ginsburg, A. G., Dunham, M. K., et al. 2011, ApJS, 192, 4
Anderson, L. D., Armentrout, W. P., Johnstone, B. M., et al. 2015, ApJS, 221, 26
Anderson, L. D., & Bania, T. M. 2009, ApJ, 690, 706
Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953
Brand, J., & Blitz, L. 1993, A&A, 275, 67
Brown, C., Dickey, J. M., Dawson, J. R., & McClure-Griffiths, N. M. 2014, ApJS, 211, 29
Brown, C., Jordan, C., Dickey, J. M., et al. 2017, AJ, 154, 23
Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, PASP, 121, 76
Chambers, E. T., Jackson, J. M., Rathborne, J. M., & Simon, R. 2009, ApJS, 181, 360
Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, PASP, 121, 213
Clemens, D. P. 1985, ApJ, 295, 422
Contreras, Y., Rathborne, J. M., Guzman, A., et al. 2017, MNRAS, 466, 340
Contreras, Y., Schuller, F., Urquhart, J. S., et al. 2013, A&A, 549, A45
Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792
Dame, T. M., Ungerechts, H., Cohen, R. S., et al. 1987, ApJ, 322, 706
Duncan, A. R., Stewart, R. T., Haynes, R. F., & Jones, K. L. 1995, MNRAS, 277, 36
Dunham, M. K., Rosolowsky, E., Evans, N. J., II, Cyganowski, C., & Urquhart, J. S. 2011, ApJ, 741, 110
Ellsworth-Bowers, T. P., Glenn, J., Rosolowsky, E., et al. 2013, ApJ, 770, 39
Foster, J. B., Jackson, J. M., Barnes, P. J., et al. 2011, ApJS, 197, 25
Foster, J. B., Rathborne, J. M., Sanhueza, P., et al. 2013, PASA, 30, e038
Foster, J. B., Stead, J. J., Benjamin, R. A., Hoare, M. G., & Jackson, J. M. 2012, ApJ, 751, 157
Green, J. A., & McClure-Griffiths, N. M. 2011, MNRAS, 417, 2500
Guzmán, A. E., Sanhueza, P., Contreras, Y., et al. 2015, ApJ, 815, 130
Haverkorn, M., Gaensler, B. M., McClure-Griffiths, N. M., Dickey, J. M., & Green, A. J. 2006, ApJS, 167, 230
Hoare, M. G., Purcell, C. R., Churchwell, E. B., et al. 2012, PASP, 124, 939
Hoq, S., Jackson, J. M., Foster, J. B., et al. 2013, ApJ, 777, 157
Jackson, J. M., Rathborne, J. M., Foster, J. B., et al. 2013, PASA, 30, e057
Jones, C., & Dickey, J. M. 2012, ApJ, 753, 62
Kavars, D. W., Dickey, J. M., McClure-Griffiths, N. M., Gaensler, B. M., & Green, A. J. 2003, ApJ, 598, 1048
Kolpak, M. A., Jackson, J. M., Bania, T. M., Clemens, D. P., & Dickey, J. M. 2003, ApJ, 582, 756
McClure-Griffiths, N. M., & Dickey, J. M. 2007, ApJ, 671, 427
McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., et al. 2005, ApJS, 158, 178
Molinari, S., Swinyard, B., Bally, J., et al. 2010, A&A, 518, L100
Payne, H. E., Salpeter, E. E., & Terzian, Y. 1983, ApJ, 272, 540

Rathborne, J. M., Jackson, J. M., Chambers, E. T., et al. 2010, ApJ, 715, 310
Rathborne, J. M., Whitaker, J. S., Jackson, J. M., et al. 2016, PASA, 33, e030
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2014, ApJ, 783, 130
Reid, M. J., Menten, K. M., Zheng, X. W., et al. 2009, ApJ, 700, 137
Roberts, W. W. 1969, ApJ, 158, 123
Roman-Duval, J., Jackson, J. M., Heyer, M., et al. 2009, ApJ, 699, 1153
Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, A&A, 504, 415
Stark, A. A. 1984, ApJ, 281, 624

Stephens, I. W., Jackson, J. M., Sanhueza, P., et al. 2015, ApJ, 802, 6
Stephens, I. W., Jackson, J. M., Whitaker, J. S., et al. 2016, ApJ, 824, 29
Urquhart, J. S., Figura, C. C., Moore, T. J. T., et al. 2014, MNRAS, 437, 1791
Urquhart, J. S., Hoare, M. G., Lumsden, S. L., et al. 2012, MNRAS, 420, 1656
Vallée, J. P. 2014, AJ, 148, 5
Wienen, M., Wyrowski, F., Menten, K. M., et al. 2015, A&A, 579, A91
Xu, Y., Reid, M. J., Zheng, X. W., & Menten, K. M. 2006, Sci, 311, 54