



# A Census of Star Formation in the Outer Galaxy. II. The GLIMPSE360 Field

Elaine Winston , Joseph L. Hora , and Volker Tolls

Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge MA 02138, USA; [elaine.winston@cfa.harvard.edu](mailto:elaine.winston@cfa.harvard.edu)

Received 2020 March 26; revised 2020 May 20; accepted 2020 May 28; published 2020 July 16

## Abstract

We have conducted a study of star formation in the outer Galaxy from  $65^\circ < l < 265^\circ$  in the region observed by the GLIMPSE360 program. This Spitzer warm mission program mapped the plane of the outer Milky Way with IRAC at 3.6 and 4.5  $\mu\text{m}$ . We combine the IRAC, Wide-field Infrared Survey Explorer (WISE), and Two Micron All Sky Survey catalogs and our previous results from another outer Galaxy survey and identify a total of 47,338 young stellar objects (YSOs) across the field spanning  $>180^\circ$  in Galactic longitude. Using the DBSCAN method on the combined catalog, we identify 618 clusters or aggregations of YSOs having five or more members. We identify 10,476 class I, 29,604 class II, and 7325 anemic class II/class III YSOs. The ratio of YSOs identified as members of clusters was 25,528/47,338, or 54%. We found that 100 of the clusters identified have previously measured distances in the WISE H II survey. We used these distances in our spectral energy distribution (SED) fitting of the YSOs in these clusters, of which 96 had YSOs with  $<3\sigma$  fits. We used the derived masses from the SED model fits to estimate the initial mass function (IMF) in the inner and outer Galaxy clusters; dividing the clusters by galactocentric distances, the slopes were  $\Gamma = 1.87 \pm 0.31$  above 3  $M_\odot$  for  $R_{\text{Gal}} < 11.5$  kpc and  $\Gamma = 1.15 \pm 0.24$  above 3  $M_\odot$  for  $R_{\text{Gal}} > 11.5$  kpc. The slope of the combined IMF was found to be  $\Gamma = 1.92 \pm 0.42$  above 3  $M_\odot$ . These values are consistent with each other within the uncertainties and with literature values in the inner Galaxy high-mass star formation regions. The slopes are likely also consistent with a universal Salpeter IMF.

*Unified Astronomy Thesaurus concepts:* Young stellar objects (1834); Protoplanetary disks (1300); Protostars (1302); Stellar mass functions (1612); Star formation (1569); Star forming regions (1565); Milky Way Galaxy (1054); Circumstellar disks (235); Infrared excess (788)

*Supporting material:* figure set, FITS files

## 1. Introduction

The study of star formation has been revolutionized with the launch of the Spitzer Space Telescope (Spitzer; Werner et al. 2004). Much of the published work to date has focused on regions at galactocentric radii less than the Sun and in nearby clouds. The Spitzer Legacy programs,<sup>1</sup> executed early in the mission, provided a large data set for studies of star formation. The original GLIMPSE survey (Benjamin et al. 2003) and subsequent follow-on programs (Churchwell et al. 2009) mapped the inner Galactic plane ( $295^\circ < l < 65^\circ$ ) with IRAC at 3.6, 4.5, 5.8, and 8  $\mu\text{m}$ , and the MIPS GAL programs (Carey et al. 2009) mapped the regions with MIPS at 24 and 70  $\mu\text{m}$ . The c2d program (Evans et al. 2003) scanned large areas in nearby molecular clouds for low-luminosity sources to obtain a sample of nearby solar-type stars for debris disk studies. The FEPS program (Meyer et al. 2004) studied a large sample of young nearby solar-type stars to trace the evolution of circumstellar gas and dust from primordial planet-building stages in young circumstellar disks through to older collisionally generated debris disks. Later, Legacy and other large programs further contributed to our knowledge of star formation in nearby molecular clouds. The Gould’s Belt program (Allen et al. 2006) completed the observations of all prominent star-forming regions within 500 pc. Megeath et al. (2012, 2016) conducted studies of the Orion A and B clouds, identifying thousands of young stellar object (YSO) candidates in this nearby massive star-forming region. Other programs mapped massive star formation complexes such as Cygnus-X

(Hora et al. 2009) and Vela-Carina (Majewski et al. 2007) at distances of  $\sim 1\text{--}2$  kpc. These and many other individual programs have produced a wealth of data on nearby star-forming regions that have been utilized in thousands of papers and will continue to be mined for years to come.

The outer Galaxy is a distinctly different environment from that of the inner Galaxy, with conditions seemingly less likely to efficiently form stars. (In this paper, we will use the term “outer Galaxy” to refer to clusters with a galactocentric radius greater than  $\sim 8$  kpc and a Galactic longitude in the range of roughly  $65^\circ < l < 265^\circ$ .) The efficiency with which a molecular cloud forms stars is thought to be dependent on its density, temperature, and chemical abundances (e.g., Evans 1999). The metallicity of the Milky Way is believed to decline as a function of galactocentric radius (Rudolph et al. 1997). Average temperatures in molecular clouds are found to be lower (Mead & Kutner 1988), as is the cosmic-ray flux (Bloemen et al. 1984). Further, the volume density of molecular clouds in the outer Galaxy is lower, so interaction rates and incidence of spiral arm crossings will be lower compared to inner Galaxy regions over the star-forming lifetime of an individual cloud.

The Spitzer Mapping of the Outer Galaxy (SMOG) survey (Carey et al. 2008) was designed to help fill in our knowledge of star formation in the outer Galaxy by providing deep coverage of a field in the outer Galaxy in the IRAC and MIPS bands. Utilizing data from this survey, we presented in our previous paper (Winston et al. 2019, hereafter Paper I) an initial study of the outer regions of our Galaxy, where environmental factors may impact on the star formation occurring there. In this paper, we present a census of star formation across the entire

<sup>1</sup> <https://irsa.ipac.caltech.edu/data/SPITZER/docs/spitzermission/observingprograms/legacy/history/>

plane of our outer Milky Way galaxy using Spitzer’s GLIMPSE360 survey (Whitney et al. 2008; Whitney & GLIMPSE360 Team 2009). This survey allows us to identify the young stellar populations of these clusters to better answer the question of whether the colder, less dense, and lower-metallicity environment of the outer Galaxy affects the formation and evolution of young stars. With this work, we will expand on our previous study by applying the techniques outlined in Paper I to the outer Galaxy covered by the GLIMPSE360 Galactic plane survey from  $65^\circ < l < 265^\circ$  in a  $3^\circ$  wide strip that follows the warp in the outer Galactic disk. Here we identify YSOs across this  $\sim 600 \text{ deg}^2$  strip from their excess IR emission, locate clusters of YSOs indicating new regions of star formation, determine the evolutionary class of the YSOs to analyze the protostellar ratio of the clusters, and make a preliminary assessment of the initial mass function (IMF) across clusters with known distances in the outer Galaxy.

This paper is organized as follows. In Section 2 we discuss the origins of the IR catalogs. In Section 3 we discuss contamination removal and the identification of the YSOs and their evolutionary classification. We then discuss the spatial distribution of the young stars and the identification of stellar clusters in the outer Galactic fields in Section 4. We discuss the fits to the SEDs of YSOs in clusters with known distances in Section 5.1, and in Section 6 we compare our results to previous YSO catalogs constructed for the outer Galaxy. Finally, a brief summary is presented in Section 7.

## 2. Observations and Data Reduction

### 2.1. GLIMPSE360 Survey

The GLIMPSE360 survey completes the coverage of the Galactic plane that began with the earlier GLIMPSE surveys. The observations were taken as part of a Spitzer Warm Mission Exploration Science program<sup>2</sup> and performed using the two short-wavelength IRAC bands at 3.6 and 4.5  $\mu\text{m}$  (Fazio et al. 2004). The GLIMPSE360 point-source archive covering Galactic longitudes in the range  $65^\circ < l < 265^\circ$  was downloaded from the Infrared Science Archive (IRSA). This catalog did not contain the SMOG field data that were previously reported in Paper I ( $102^\circ < l < 109^\circ$ ) and observed for the cryogenic SMOG program. It also excludes the central field surrounding the Cygnus star-forming complex ( $76^\circ < l < 82^\circ$ ), which forms part of the Cygnus-X survey, though the flanking fields in the Galactic plane are included in the archive. Two types of catalog are available: a highly reliable point-source catalog and a highly complete point-source archive. The archive includes sources with spatial positions as close as  $0''.5$ , while the catalog excludes sources closer than  $2''$  in position. As in Paper I, the more complete archive was used in this paper to aid in the detection of fainter, more embedded YSOs.

The IRAC observations at 3.6 and 4.5  $\mu\text{m}$  were obtained in High Dynamic Range mode, which was comprised of three visits per mosaic position with 0.6 and 12 s integrations, similar to the SMOG survey. These data products were produced using the GLIMPSE team’s pipeline. The GLIMPSE360 catalog contained a total of 49,378,042 sources. The mid-IR photometry was supplemented by *J*-, *H*-, and *K*-band photometry

from the Two Micron All Sky Survey (2MASS) point-source catalog (Skrutskie et al. 2006). The photometric catalogs were merged using a  $1''.6$  matching radius. Documentation describing the GLIMPSE360 survey and the reduction process in detail is available on the IRSA website.<sup>3</sup>

### 2.2. WISE Catalog

The Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010) provides mid-IR photometry at 3.5, 4.6, 12, and 22  $\mu\text{m}$ . The AllWISE catalog is an all-sky survey combining the cryogenic WISE all-sky survey and the NEOWISE postcryogenic survey (Mainzer et al. 2011). The catalog is available via the NASA/IPAC IRSA archive.<sup>4</sup> Selection of all of the sources in the GLIMPSE360 field was made, resulting in a regional catalog of 14,483,596 point sources. The WISE satellite has considerably lower spatial resolution when compared to Spitzer, with a highest resolution of  $\sim 6''.1$  at 3.5  $\mu\text{m}$ . However, the astrometric accuracy and matching to the 2MASS catalog are to within  $1''.5$ , so we performed catalog matching to the GLIMPSE360 catalog at  $1''.5$ .

## 3. YSO Identification and Classification

The YSOs are most frequently identified by their excess emission at IR wavelengths. This emission arises from reprocessed stellar radiation in the dusty material of their natal envelopes or circumstellar disks. The IR identification of YSOs is carried out by identifying sources that possess colors indicative of IR excess and distinguishing them from reddened and/or cool stars (Allen et al. 2004; Gutermuth et al. 2004; Winston et al. 2007).

A full description of the criteria for identification and selection of YSO and non-YSO sources as applied to the data sets here is given in the Appendices. The following subsections outline the removal of background extragalactic objects, YSO selection methods for each data set, the complete YSO catalog, and the evolutionary classification of the YSOs based on their excess IR emission.

### 3.1. IRAC and 2MASS

#### 3.1.1. Contamination

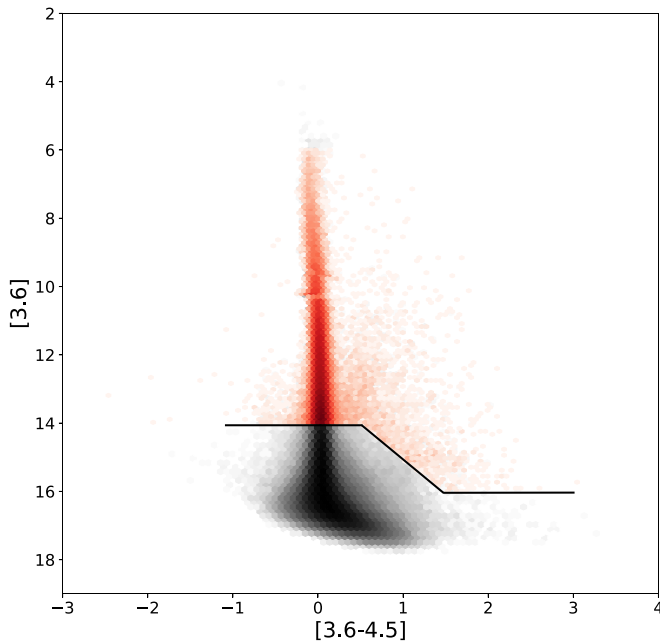
The GLIMPSE360 field covers the majority of the plane of the outer Galaxy, where the background sources suffer negligible extinction from the Galactic bar, and thus it is expected that many of the point sources detected will be active galactic nuclei (AGNs) or star-forming galaxies (polycyclic aromatic hydrocarbon (PAH) galaxies). Knots of emission in the structure of molecular clouds may also be mistaken for YSOs. A further source of confusion in the YSO sample comes from sources with photometric contamination of the apertures by PAH emission.

Such sources, which we will refer to as contaminants, were identified in the SMOG field as outlined in Paper I. The same methods could not be applied to the GLIMPSE360 data due to the lack of 5.8 and 8  $\mu\text{m}$  photometry. In order to constrain the photometric and color cuts to be applied to the GLIMPSE360 data to remove contaminants, the SMOG field data were utilized and cuts were determined from the characteristics of

<sup>2</sup> <https://irsa.ipac.caltech.edu/data/SPITZER/docs/spitzermission/observingprograms/es/>

<sup>3</sup> [https://irsa.ipac.caltech.edu/data/SPITZER/GLIMPSE/doc/glimpse360\\_dataprod\\_v1.5.pdf](https://irsa.ipac.caltech.edu/data/SPITZER/GLIMPSE/doc/glimpse360_dataprod_v1.5.pdf)

<sup>4</sup> <https://irsa.ipac.caltech.edu/Missions/wise.html>



**Figure 1.** Example of color–magnitude of the IRAC [3.6] vs. [3.6–4.5] contaminant selection in the 5° section of the SFOG field in the range  $180^\circ < l < 185^\circ$ . The red points represent the catalog after contaminants have been removed, and the gray points are the selected contaminants.

the contaminants in that field. A description of the applied cuts is given in Appendix A. Sources in the color/magnitude spaces where the contaminants were located were removed before the selection of YSOs was carried out.

Figure 1 shows the criteria for the removal of contaminants in the GLIMPSE360 data for sources in a section of the field in the Galactic longitude range  $180^\circ < l < 185^\circ$ . It was necessary to show only a subsection in order to enhance the clarity of the plot.

### 3.1.2. IRAC YSO Selection

The YSOs were selected using a combination of color–color diagrams (CCDs) with IRAC and 2MASS+IRAC colors, as described in Appendix A. Photometric uncertainties of  $<0.2$  mag and magnitudes fainter than the saturation limit were required in all bands used for a particular CCD to select YSOs.

The full GLIMPSE360 catalog contains 49,378,042 sources, of which 28,837 were identified as YSO candidates using the combined 2MASS and IRAC photometry. Hereafter, we will refer to the YSO candidates as “YSOs”; however, a definitive classification would require a more detailed analysis of the spectra and other characteristics of each object. Figure 2 shows the three source-selection CCDs used for the identification of YSOs for sources in the section of the field in the range  $180^\circ < l < 185^\circ$ .

The remaining contaminants in the YSO sample that are not accounted for here include galactic asymptotic giant branch (AGB) stars and highly/unusually reddened field stars that may be confused for anemic class II (class IIa) or class III objects. Such objects are expected to be scattered randomly over the field. This issue is discussed further in Section 4.3.3 on AGB contamination.

### 3.2. WISE Source Selection

The WISE catalog covering the GLIMPSE360 field comprised  $\sim 29\%$  of the number of detections as the IRAC catalog. The long edges of the GLIMPSE360 field exhibit a sawtooth pattern due to the IRAC mapping procedure. In our extraction of the WISE data from the all-sky catalog, we used a simple Galactic latitude cutoff that matches the largest extent of the IRAC data, so the edges of the WISE source distribution are smooth compared to the sawtooth edge of the GLIMPSE360 data. For this reason, a number of WISE sources fall into regions that are not included in the IRAC catalog and are identified solely by the WISE photometry on the edges of the field. Contaminant removal was undertaken using a procedure similar to the one used in Paper I and as outlined in Appendix B. The resulting catalog contained 543,457 point sources after the removal of spurious sources and extragalactic contaminants.

A search for YSOs was performed using the four-band WISE and 2MASS photometry; a detailed discussion is given in Appendix B. Of the 543,457 point sources, a total of 20,892 were identified as YSOs.

### 3.3. Combined IRAC and WISE Selection

Paper I compared the photometry in IRAC bands 1 and 2 and WISE bands 1 and 2 for sources matched to within  $1''.5$ . It was found that the photometric magnitudes matched well to within the uncertainties, which include measurement errors and possible variability of the sources themselves.

In the GLIMPSE360 field, we used photometry from the two available IRAC bands to replace the appropriate WISE band fluxes where the sources matched to within  $1''.5$ , and the WISE selection criteria were applied again to locate YSOs that may have reliable longer-wavelength detections in WISE but, due to the shallower survey coverage or source confusion, may not have been reliably detected at the shorter wavelengths. In these cases, the 2MASS selection criteria were not applied, since they replicate the selection made using the GLIMPSE360 data directly. The combined IRAC+WISE source list contained 318,588 objects. A total of 11,196 YSOs were identified following our methodology, as outlined in Appendix C.

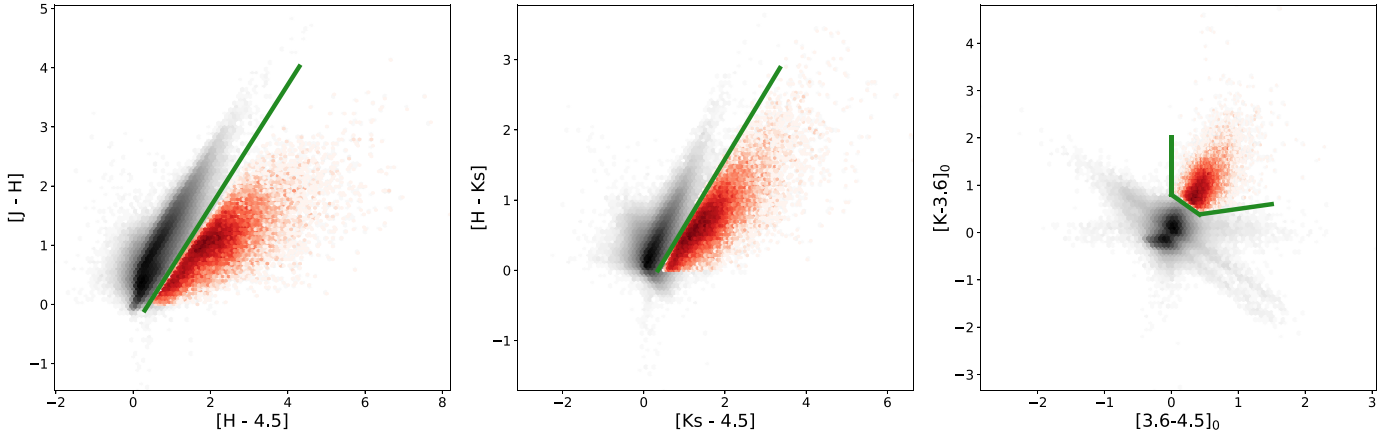
Figure 3 shows the two source-selection CCDs used in the IRAC+WISE identification of YSOs for sources in the section of the field in the range  $180^\circ < l < 185^\circ$ .

### 3.4. Combined SFOG YSO Catalog

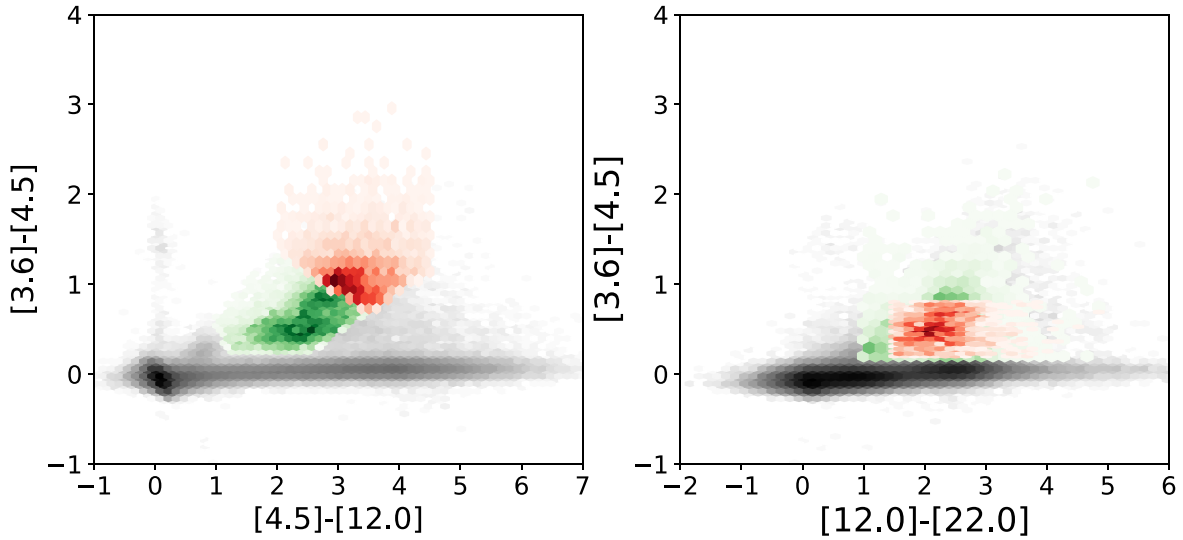
The three sets of YSO selections, 2MASS+IRAC (28,837 sources), IRAC+WISE (11,196 sources), and WISE (20,892 sources), were merged and common objects combined based on the source spatial position. Sources in different sets that were within  $1''.0$  of each other were assumed to be the same source and combined. The 2MASS+IRAC and IRAC+WISE catalogs also required a detection in the GLIMPSE360 catalog, which uniquely identifies each source. The unique list of YSOs in the GLIMPSE360 field contained 42,757 objects.

The GLIMPSE360 YSOs were then combined with the SMOG field YSOs to create the full catalog, which contained 47,405 candidate YSOs. We call this the Star Formation in the Outer Galaxy (SFOG) catalog, and we refer to the total field covered by all of the component surveys as the SFOG field. The SFOG catalog was then used for the following analysis.





**Figure 2.** Selection of CCDs of the near- and mid-IR 2MASS and IRAC YSO selection (red circles), overlaid on sources in the SFOG field (gray points) in the range  $180^\circ < l < 185^\circ$ . Left: 2MASS-IRAC  $[J-H]$  vs.  $[H-4.5]$ . Middle: 2MASS-IRAC  $[H-K_s]$  vs.  $[K_s-4.5]$ . Right: 2MASS-IRAC  $[K_s-3.6]$  vs.  $[3.6-4.5]$ .



**Figure 3.** Selection of CCDs of the IRAC+WISE YSO selection (red: excess; green: protostars) overlaid on sources in the SFOG field (gray points) in the range  $180^\circ < l < 185^\circ$ . Left: IRAC+WISE  $[3.6-4.5]$  vs.  $[4.5-12.0]$ . Right: IRAC+WISE  $[3.6-4.5]$  vs.  $[12.0-22.0]$ .

Table 1 lists a selection of the column identifiers of the photometry table for the full list of identified YSOs. This table is a stacked table of subsets from the standard IPAC tables available on the IPAC website for the GLIMPSE360, SMOG, and WISE data sets. It is available in its entirety in the online version of the paper.

The online SIMBAD catalog was searched for matches to the SFOG catalog within  $2''$ , with 7039 YSOs ( $\sim 17\%$ ) found to have a previous identification. Table 2 lists the object identifiers and positions and by which selection method the YSO was identified. Table 3 lists the YSOs that were found to have matches in the SIMBAD catalog along with their alternate identifications and object types. Figure 4 shows the spatial distribution of the identified YSOs over the outer plane of the Milky Way.

### 3.4.1. Evolutionary Classification

Young stars evolve through a number of broad stages from the embedded core phase, through the protostellar phase where stellar accretion is still dominant, to the circumstellar disk phase where the envelope has dissipated and processing of disk

material is ongoing, to the weak disk regime where the disk has dissipated and planets will have formed.

A general evolutionary classification of YSOs in the SFOG field was carried out by measuring the slope,  $\alpha$ , of the spectral energy distribution (SED) across the mid-IR bandpasses (Lada & Wilking 1984). In our SFOG catalog, many sources are detected with only the two shorter IRAC bands, while a smaller number also have the longer-wavelength WISE data. The SED slope was calculated based on the available photometric bands longward of the 2MASS  $K$ -band inclusive for each source by performing a least-squares polynomial fit to the data.

Protostellar objects (class 0 and I) have a rising slope,  $\alpha > 0$ ; class II sources are characterized by decreasing slopes in the range  $-1.6 < \alpha < 0$ , while class IIa sources lack optically thick emission from a disk and possess decreasing slopes  $\alpha < -1.6$ , consistent with a weak emission above a stellar photosphere. Truly diskless class III objects show slopes in the range  $-2.7 < \alpha < -2.0$  (Lada 1987; Lada et al. 2006). Many previous publications using Spitzer data combine these two categories as class III objects (e.g. Hora et al. 2009; Saral et al. 2017; Winston et al. 2019).

**Table 1**  
SFOG Field YSOs: Photometry Table Description

Column Number	Column ID	Description
0	designation	GLIMPSE ID
1	2mass designation	2MASS ID
2	2mass cnt	2MASS counter
3	l	Longitude
4	b	Latitude
7	ra	R.A.
8	dec	Decl.
12	mag_J	2MASS <i>J</i> band
13	dJ_m	2MASS <i>J</i> -band uncertainty
14	mag_H	2MASS <i>H</i> band
15	dH_m	2MASS <i>H</i> -band uncertainty
16	mag_K	2MASS <i>Ks</i> band
17	dKs_m	2MASS <i>Ks</i> -band uncertainty
18	mag3_6	IRAC band 1
19	d3_6m	IRAC band 1 uncertainty
20	mag4_5	IRAC band 2
21	d4_5m	IRAC band 2 uncertainty
22	mag5_8	IRAC band 3
23	d5_8m	IRAC band 3 uncertainty
24	mag8_0	IRAC band 4
25	d8_0m	IRAC band 4 uncertainty
78	designation_1	WISE identifier
79	ra_1	R.A.
80	dec_1	Decl.
93	w1mpro	WISE band 1
94	w1sigmpro	WISE band 1 uncertainty
97	w2mpro	WISE band 2
98	w2sigmpro	WISE band 2 uncertainty
101	w3mpro	WISE band 3
102	w3sigmpro	WISE band 3 uncertainty
105	w4mpro	WISE band 4
106	w4sigmpro	WISE band 4 uncertainty
150	tmass_key	WISE 2MASS ID
154	j_m_2mass	WISE 2MASS <i>J</i> band
155	j_msig_2mass	WISE 2MASS <i>J</i> -band uncertainty
156	h_m_2mass	WISE 2MASS <i>H</i> band
157	h_msig_2mass	WISE 2MASS <i>H</i> -band uncertainty
158	k_m_2mass	WISE 2MASS <i>Ks</i> band
159	k_msig_2mass	WISE 2MASS <i>Ks</i> -band uncertainty
179	mag_24	MIPS band 1
180	d24_m	MIPS band 1 uncertainty

**Note.** This table is available in its entirety in FITS format.

As we do not have ancillary data, such as X-ray observations, to separate young completely diskless class III YSOs from field stars, we cannot reliably differentiate between the weak disk-bearing class IIa YSOs identified here and truly mid-IR diskless class III YSOs (Winston et al. 2011, 2018). For this reason, we list the class of all objects with slope  $\alpha < -1.6$  as being class IIa/III in this paper.

The identified YSOs were assigned classes as follows: 10,476 class 0/I, 29,604 class II sources, and 7325 weak emission class IIa/III stars. Further, some of the WISE-identified sources were classified from the CCD as candidate transition disk sources. This classification was not used in the subsequent analysis, with only the classification based on SED slope reported.

Figure 5 shows the spatial distribution of the identified YSOs by evolutionary classification over the plane. The classification of each YSO is listed in Table 2.

### 3.4.2. Sample Completeness

The SFOG catalog is susceptible to the same difficulties in assessing incompleteness in the sample as the SMOG field, discussed in Paper I. In brief, the large spatial distribution and lack of data on the distances in the majority of the clusters means that estimates of the minimum mass YSOs detected are not possible. Incompleteness by evolutionary class is also difficult to quantify: because of the wavelength range of the 2MASS and IRAC data from 1.2 to 4.5  $\mu\text{m}$ , the survey is most sensitive to class II pre-main-sequence objects. The WISE long-wavelength bands at 12 and 22  $\mu\text{m}$  make it more sensitive than GLIMPSE360 for the detection of embedded protostars, but it is hampered by lower resolution and sensitivity at the shorter wavelengths.

## 4. Cluster Identification and Properties

### 4.1. Identification with DBSCAN

A cursory visual examination of the spatial distribution of the YSOs over the  $\sim 600$  deg<sup>2</sup> SFOG field shows evidence of clustering/clumping, as can be seen in Figure 4. Following the method described in Paper I, we used the DBSCAN (Ester et al. 1996) density-based algorithm to identify overdensities in the spatial distribution of the YSOs. The values of the two free parameters,  $\epsilon$  (the scaling size for clustering) and MinPts (the minimum number of points required to define a dense region), were determined following the approach used by Joncour et al. (2018) in the Taurus region.

Figure 6 (left) shows the one point correlation function, which gives the ratio of the cumulative distributions of the identified YSOs and a random distribution over the same field. From this, the value of  $\epsilon = 0.1$  was selected. Figure 6 (right) examines the cumulative distribution of the three nearest-neighbor distributions at the 8th, 9th, and 10th nearest neighbors for the random distribution showing that a probability of 0.001 occurs at 0.1 for a minimum cluster size with nine members.

### 4.2. Cluster Properties

With these values, 621 clusters were identified in the SFOG field. Figure 7 shows the spatial distribution of the identified clusters of YSOs over the plane. The black YSOs are unclustered, and the clustered YSOs are color-coded by identified cluster. The clustered YSOs represent 54% of the whole catalog, with 21,810 YSOs not clustered. The minimum cluster size was five members, and the largest cluster identified contained 1177 members. The median cluster size is 17 members. Of the 621 clusters, 133 have 10 members or less, 25 have 100–200 members, and 22 have more than 200 identified members.

Figure 8 shows the spatial distribution of the identified clusters overlaid on a schematic view of the Milky Way (Hurt 2008).

There is a possibility that some of the smaller clusters identified are not genuine stellar groupings but rather chance overdensities in the field. To determine the statistical likelihood that the identified groupings are true clusters, we repeated the DBSCAN analysis for 1000 iterations of randomly distributed points covering a similar area as the SFOG field. Using the same values of  $\epsilon = 0.1$  and MinPts = 9, we varied the number of fake “stars” as follows: 21,810, 29,604, 36,000, 43,000, and

**Table 2**  
SFOG Field YSOs: Classification Table Description

SF ID	Glimpse ID	R.A. Decl.	IRAC YSO <sup>a</sup>	WISE YSO	IRACWISE YSO	SMOG YSO	Cluster Num. <sup>b</sup>	Evolutionary Class <sup>c</sup>
SRC0	SSTGLMA G064.5244+01.1116	19 <sup>h</sup> 49 <sup>m</sup> 46 <sup>s</sup> .2437 +28 <sup>d</sup> 21 <sup>m</sup> 33 <sup>s</sup> .8616	1	0	0	0	−1	1
SRC1	SSTGLMA G064.5361+02.1303	19 <sup>h</sup> 45 <sup>m</sup> 47 <sup>s</sup> .3287 +28 <sup>d</sup> 52 <sup>m</sup> 59 <sup>s</sup> .394	1	0	0	0	−1	1
SRC2	SSTGLMA G064.5721+01.1438	19 <sup>h</sup> 49 <sup>m</sup> 45 <sup>s</sup> .2604 +28 <sup>d</sup> 25 <sup>m</sup> 00 <sup>s</sup> .6924	1	0	0	0	−1	1
SRC3	SSTGLMA G064.5808+01.1670	19 <sup>h</sup> 49 <sup>m</sup> 41 <sup>s</sup> .0153 +28 <sup>d</sup> 26 <sup>m</sup> 09 <sup>s</sup> .9204	1	0	0	0	−1	1
SRC4	SSTGLMA G064.5925+02.9442	19 <sup>h</sup> 42 <sup>m</sup> 41 <sup>s</sup> .1101 +29 <sup>d</sup> 20 <sup>m</sup> 13 <sup>s</sup> .8336	1	0	0	0	−1	1
SRC5	SSTGLMA G064.6126+01.2039	19 <sup>h</sup> 49 <sup>m</sup> 36 <sup>s</sup> .7589 +28 <sup>d</sup> 28 <sup>m</sup> 56 <sup>s</sup> .0388	1	0	0	0	−1	1
SRC6	SSTGLMA G064.6262+01.0862	19 <sup>h</sup> 50 <sup>m</sup> 06 <sup>s</sup> .311 +28 <sup>d</sup> 26 <sup>m</sup> 02 <sup>s</sup> .9688	1	0	0	0	−1	1
SRC7	SSTGLMA G064.6397+01.2037	19 <sup>h</sup> 49 <sup>m</sup> 40 <sup>s</sup> .5398 +28 <sup>d</sup> 30 <sup>m</sup> 19 <sup>s</sup> .8144	1	0	0	0	−1	2
SRC8	SSTGLMA G064.6532+02.0310	19 <sup>h</sup> 46 <sup>m</sup> 26 <sup>s</sup> .9479 +28 <sup>d</sup> 56 <sup>m</sup> 04 <sup>s</sup> .6608	1	0	0	0	−1	1
SRC9	SSTGLMA G064.6574+02.9303	19 <sup>h</sup> 42 <sup>m</sup> 53 <sup>s</sup> .2805 +29 <sup>d</sup> 23 <sup>m</sup> 11 <sup>s</sup> .598	1	0	0	0	−1	2

**Notes.** This table is available in its entirety in FITS format.

<sup>a</sup> A “1” in these YSO columns means the object was identified as a YSO based on this data set (see Section 3).

<sup>b</sup> The cluster number of the YSO is indicated in this column (see Table 4). A “−1” indicates that there was no cluster affiliation identified.

<sup>c</sup> The numerical YSO class, as described in Section 3.4.1.

**Table 3**  
SFOG Field YSOs: SIMBAD Matches within 2''

SF ID	Glimpse ID	R.A. (J2000)	Decl. (J2000)	SIMBAD ID	OTYPE
SRC9	SSTGLMA G064.6574+02.9303	295.722002	29.386555	“2MASS J19425328+2923114”	AGN_Candidate
SRC28	SSTGLMA G064.8321+01.3951	297.34215	28.768332	“HBHA 2703-38”	Em*
SRC31	SSTGLMA G064.8407+02.9476	295.808845	29.554165	“V* V1279 Cyg”	Mira
SRC35	SSTGLMA G064.8492+00.3496	298.375033	28.249132	“2MASS J19532999+2814568”	Candidate_YSO
SRC53	SSTGLMA G064.9585+02.4134	296.408345	29.390237	“IRAS 19436+2916”	Star
SRC62	SSTGLMA G064.9774+00.2589	298.538293	28.312292	“2MASS J19540918+2818443”	Candidate_YSO
SRC67	SSTGLMA G064.9890+00.2190	298.583948	28.301645	“IRAS 19523+2810”	Star
SRC69	SSTGLMA G064.9949+00.2704	298.537384	28.333228	“2MASS J19540897+2819594”	Candidate_YSO
SRC73	SSTGLMA G065.0172+00.0636	298.87522	28.179735	“OH 65.0-0.1”	OHIR
SRC77	SSTGLMA G065.0207+02.7296	296.129139	29.601918	“IRAS 19425+2928”	Star

**Note.** This table is available in its entirety in FITS format.

47,405. These are the numbers of unclustered and class II YSOs, two intermediate values, and the total number of YSOs in the SFOG catalog. This range was used to take into account the known clustering of objects, which reduces the overall density in the field and thus reduces the likelihood of a “fake” cluster being detected. The average number of fake clusters detected for each run was 0.04, 0.43, 1.9, 6.9, and 13.8. We can therefore assume that between zero and 14 of the 621 clusters found by DBSCAN may be misidentified random alignments. The fake clusters in our random trials all have  $\leq 10$  members, so if any clusters in our SFOG list in Table 4 are chance overdensities, rather than true clusters, they would most likely be in the subset of 133 clusters with  $\leq 10$  YSOs.

There is the further possibility that the unclustered YSOs are contaminants: foreground or background Galactic field dwarfs or AGB stars with either a high extinction or a small amount of dusty material surrounding them that leads to an excess of flux in the mid-IR or extragalactic sources. Spectroscopic analysis would be necessary to secure the identification of all sources.

The approximate area of each cluster was quantified by measuring the convex hull of the associated cluster members. The convex hull is the set of points whose vertices include all of the points in the set. Figures 9 and 10 show eight examples of clusters identified across the SFOG field, showcasing the range of sizes and environments over which clusters were identified. In many cases, it is clear that the clusters form part of a larger association

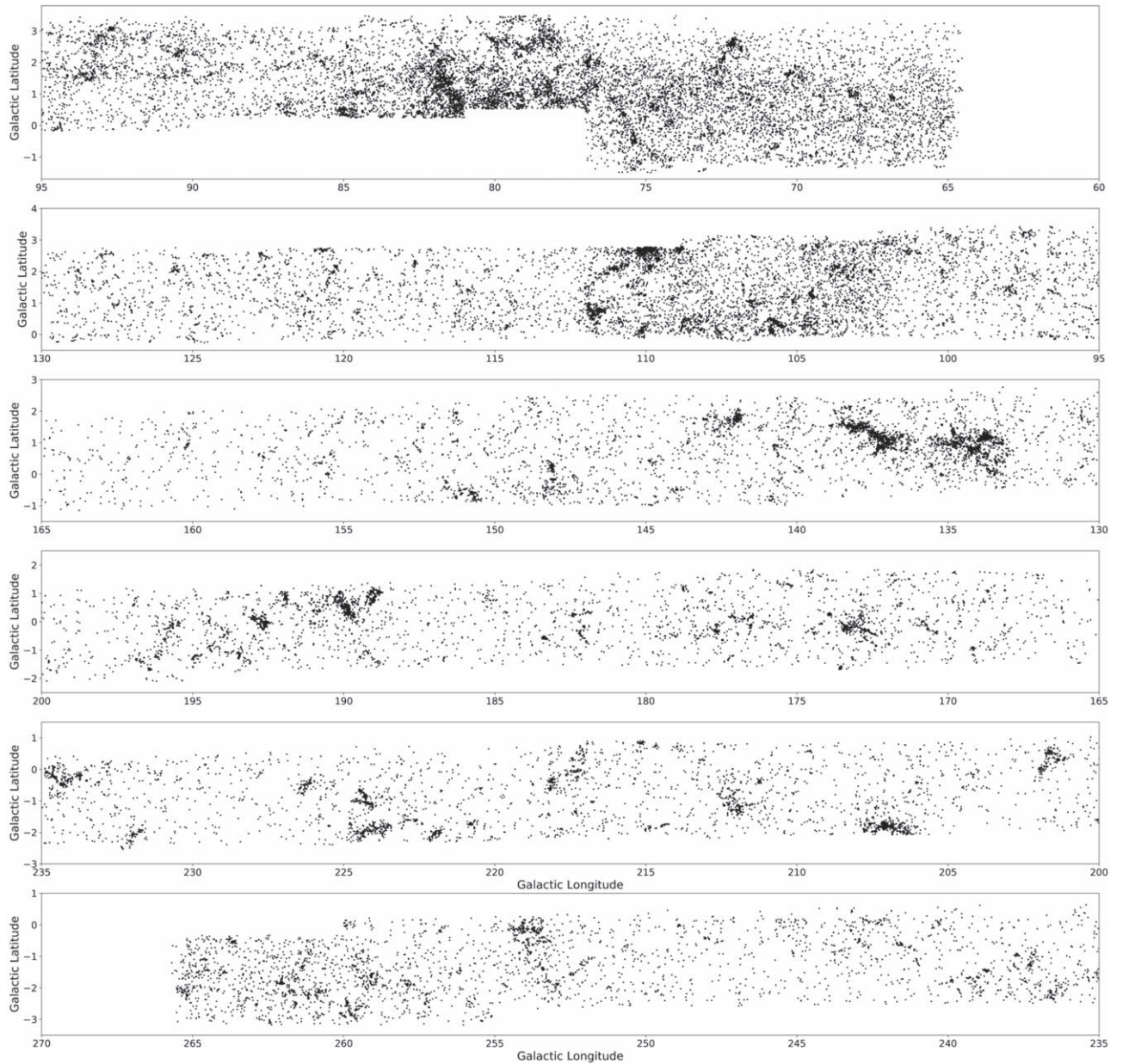
or star-forming complex. They are often surrounded by nonclustered YSOs that are likely to be associated with the cluster but did not satisfy the DBSCAN criteria. We present these clusters as starting points for future studies.

The cluster identification of each YSO is listed in Table 2. Table 4 lists the properties of the largest clusters (number of YSOs  $> 100$ ) found in the SFOG catalog, sorted by decreasing number of YSOs in the cluster. The table gives the cluster number and name, the number of YSOs, the coordinates of the cluster central point, the circular radius based on the separation of the most distant YSOs, and the association with previously identified star-forming regions. The complete version of the table is available in the electronic version of this paper. The electronic version also includes a list of WISE H II counterparts and a full listing of the SIMBAD objects located within the convex hull of each cluster. These sources are not assumed to be physically associated with the cluster, and no attempt has been made to filter the lists or match them to the YSOs. They are provided as a reference for more in-depth studies.

Images of each of the individual clusters were constructed using mosaics with a radius of five times the estimated convex hull radius of each cluster. The IRAC mosaics were downloaded from the IPAC servers, and the WISE Coadd<sup>5</sup> was used to generate mosaics of the fields in the WISE bands. For

<sup>5</sup> <https://irsa.ipac.caltech.edu/applications/ICORE>





**Figure 4.** Spatial distribution of all identified YSOs in the SFOG catalog from the combined 2MASS-IRAC, IRAC+WISE, and WISE photometric selection criteria and also incorporating the SMOG YSOs.

cases where a cluster was near the edge of the IPAC mosaics of the IRAC data, adjoining mosaics were combined to form images that could be used to display the full cluster. These mosaics were then used to create a selection of three-color images and a WISE  $12\ \mu\text{m}$  band gray-scale image overlaid with the cluster convex hull and the locations of the YSOs by evolutionary class. The full set of mosaics for every cluster is available on the Harvard Dataverse SFOG page (Winston et al. 2020).

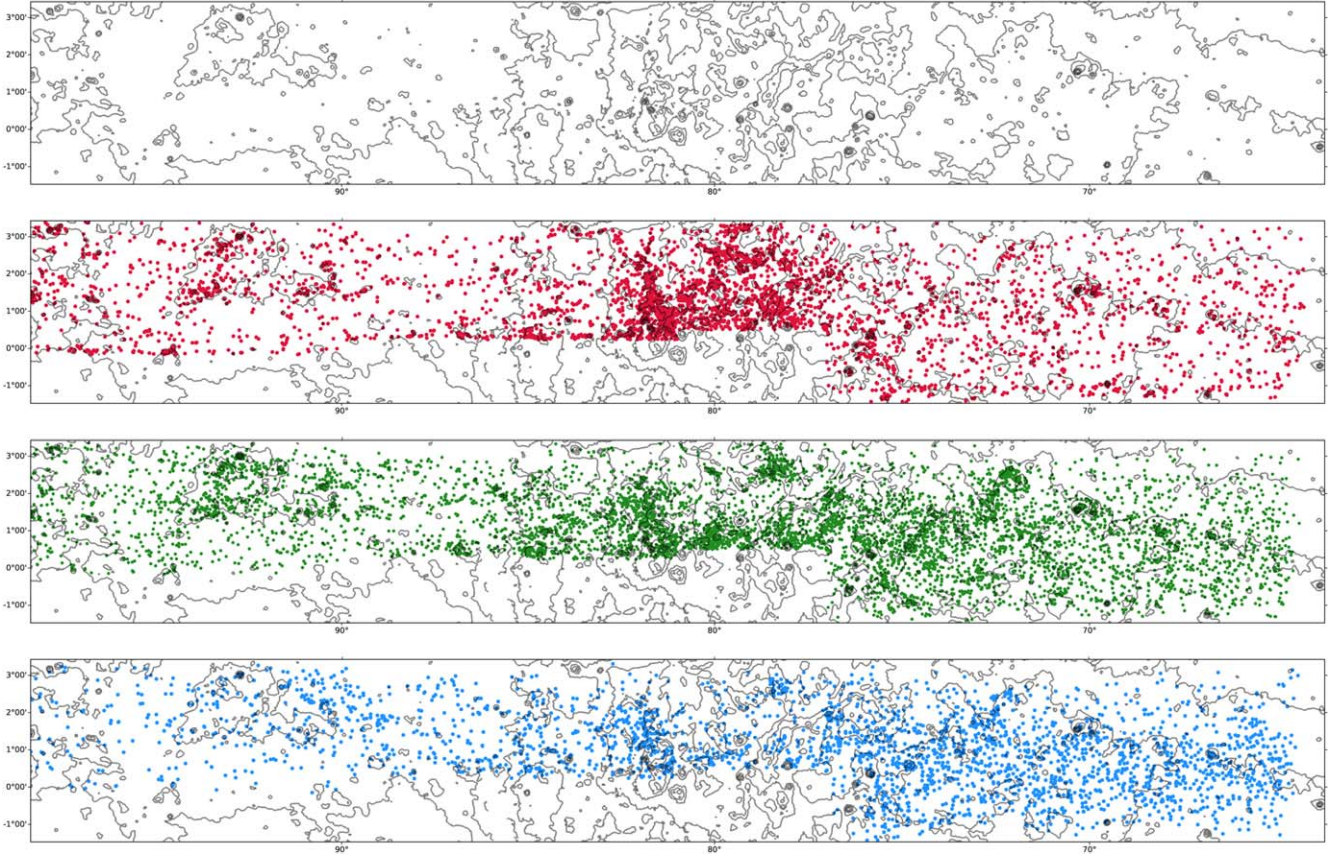
As can be seen from Table 4, all of the larger clusters are associated with previously known star-forming regions identified, for example, through surveys of HII regions, CO emission, radio continuum emission, detection of dark clouds, or the IR emission from compact sources or the surrounding nebula. However, we also find that six of the smaller clusters do not have any SIMBAD sources within their convex hulls and may be newly identified.

### 4.3. Catalog Contamination

The true level of contaminants remaining in the catalog used to select YSOs is difficult to estimate precisely given the broad range in spatial coverage and distances covered by the SFOG field. Spectral typing of the YSOs would provide confirmation of their nature, and a program to obtain spectra and X-ray observations of a selection of the clusters is currently underway. We discuss in the following sections the possible sources of contamination in the YSO catalog and their effects on the clustering analysis.

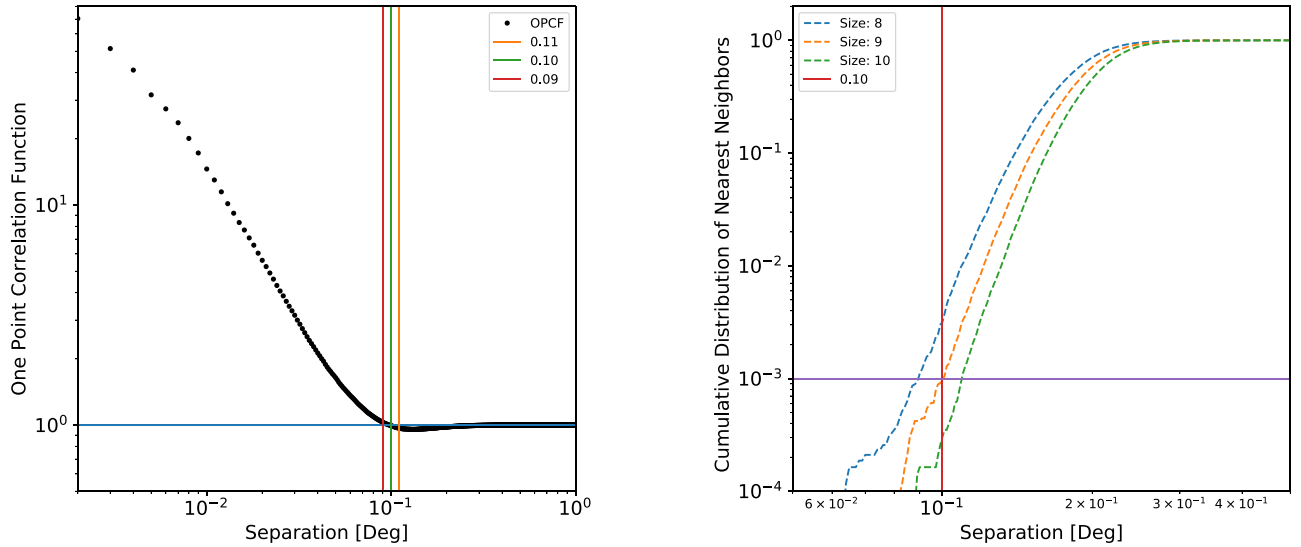
#### 4.3.1. IRAC Sample Contamination

To examine the effectiveness of the contamination cuts used in the GLIMPSE360 analysis, a comparison was made to the SMOG field of Paper I. By applying the same contamination cuts and selection criteria used in this paper, 1512 YSOs were



**Figure 5.** Spatial distribution by evolutionary classification of the identified YSOs in the SFOG catalog. Plotted in each panel are contours of the  $100\ \mu\text{m}$  IRAS IRIS image (Miville-Deschênes & Lagache 2005) as an indication of the dust distribution along the outer Galactic plane. Class I objects are plotted in red in the second panel, class II in green in the third panel, and class IIa/III in cyan in the fourth panel. This figure shows only one-sixth of the entire field; the online version contains an electronic figure set containing all six panels covering the full SFOG field.

(The complete figure set (6 images) is available.)

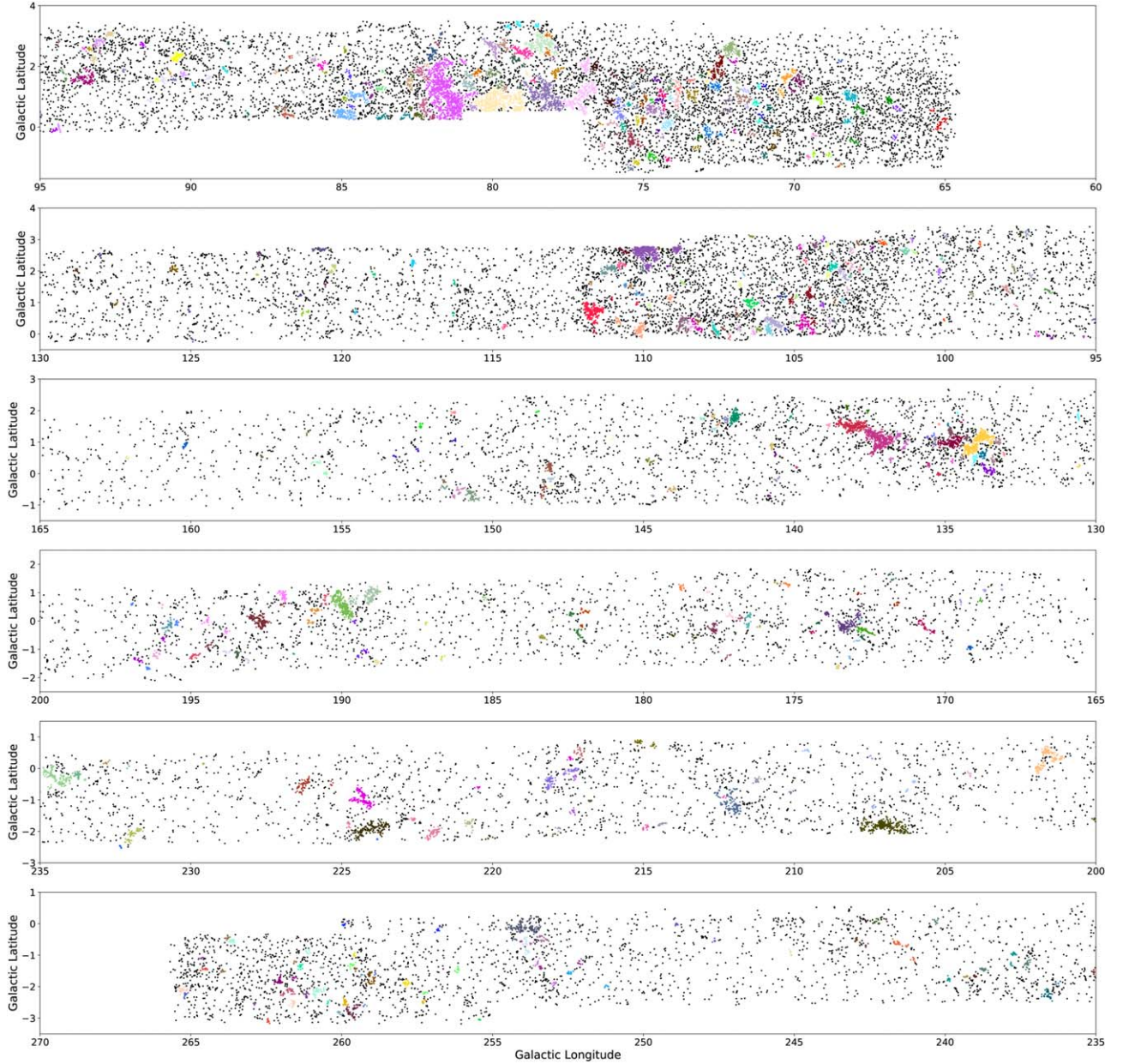


**Figure 6.** Selection of the criteria for the DBSCAN clustering algorithm. Left: one point correlation function showing the ratio of the YSO and random cumulative distributions, with a crossing point at  $0.1$  separation. Right: cumulative distribution of three nearest-neighbor distributions at the 8th, 9th, and 10th nearest neighbors. The  $10^{-3}$  probability occurs at  $0.1$  for a cluster density of nine members.

selected in the SMOG field. Of these, 1102 objects were matched at  $1''$  between the two catalogs. A majority of the nonmatching sources were included in the SFOG catalog due to

the differences between the selection criteria required because of the lack of IRAC  $5.8$  and  $8\ \mu\text{m}$  and MIPS  $24\ \mu\text{m}$  coverage in the GLIMPSE360 field. Plotting these sources on an IRAC





**Figure 7.** Spatial distribution of the identified YSOs showing the clusters identified by the DBSCAN method. The sources in each cluster are color-coded. The black dots represent those YSOs not identified as belonging to a cluster.

four-color CCD, 22 objects exhibited colors similar to those of galaxies. This would imply that 22/1512 or  $<2\%$  of the candidate YSOs are likely to be extragalactic contaminants.

By applying this percentage to the 28,837 objects in the GLIMPSE360 component of our catalog, we would estimate that 420 of these may possibly be contaminants. These would likely be found predominantly in the nonclustered population, since both foreground and background contaminants tend to be randomly distributed across the field.

#### 4.3.2. Contamination in the WISE Sample

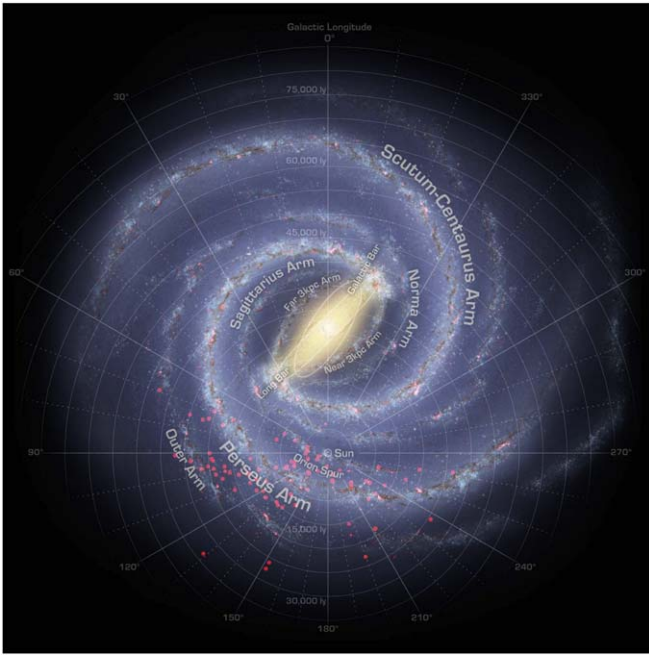
In order to assess the validity of the clustering algorithm, each of the regions was visually checked to look for interesting or spurious clusters. From this assessment, a total of three of the identified clusters, Nos. 515, 516, and 564, were found to be invalid due to spurious WISE sources. The WISE

photometry contaminant removal process removes confirmed diffraction spike objects, but those objects with tentative identification as photometric diffraction spikes are not removed. In each of these cases, a visual inspection of the images containing the “cluster” found them to contain a number of diffraction spike objects surrounding a saturated bright foreground star.

In removing these three spurious clusters, the number of clusters is reduced to 618 in total. The number of YSOs is reduced from 47,405 to 47,338 YSOs. The breakdown by evolutionary class becomes 10,461 class I, 29,552 class II, and 7325 class IIa/III YSOs.

#### 4.3.3. AGB Contamination

The Besançon Galactic population synthesis models were used to estimate the AGB population at four points along the



**Figure 8.** Schematic representation of the Milky Way (Hurt 2008) with red points overlaid showing the locations of the identified clusters with known distance measurements.

Galactic plane (Robin et al. 2003, 2014). The points chosen were at  $l = 70^\circ, 105^\circ, 180^\circ,$  and  $250^\circ$  and  $b = 1^\circ, 1^\circ, 0^\circ,$  and  $-1^\circ$ , respectively, with a  $1 \text{ deg}^2$  field at each location. The models predicted that these fields contained either one or two AGB stars each. By scaling these numbers up to the approximate  $600 \text{ deg}^2$  size of the SFOG field, we estimate that there may be between 600 and 1200 AGB stars in the SFOG field. This would imply that  $\sim 1200/47,338$  or  $\sim 2.5\%$  of the candidate YSOs may be AGB stars. The AGB stars would tend to be randomly distributed over the field and thus less likely to be included in a cluster, so these contaminants are not likely to affect our clustering analysis significantly.

## 5. YSO and Cluster Physical Characteristics

### 5.1. SED Model Fitting

The Python SEDFitter<sup>6</sup> package of Robitaille et al. (2007) was used to provide an estimate of the mass, age, disk, and accretion properties of the YSOs in the SFOG field. The code uses a sample grid of YSO model SEDs with varying age, mass, inclination, etc. to compare to the input photometry, with the scale factor  $S$  (dependent on source distance and luminosity) and extinction  $A_V$  as free parameters. The code returns a sample of best-fit models and their associated parameters. From our previous SMOG study, it was determined that within each cluster, the range in distances for the best-fit models covered the entire provided distance range, and that therefore the results could not be used in any meaningful way when the distance to the YSO was unknown.

Therefore, only those clusters with reliable distance estimates from WISE H II regions (Anderson et al. 2014) located within the cluster’s convex hull were used in the SEDFitter modeling. Reliable distances were found for 100 of the 618 clusters identified in the SFOG field. The H II regions

associated with each cluster and their distances are listed in the full online version of Table 4. The galactocentric distances for these 100 clusters ranged from 7.8 to 18.1 kpc.

The SEDFitter routine was run using a fixed distance range based on the distance to the cluster, allowing the  $A_V$  to vary from zero to 40 for each cluster separately. The YSOs lacking photometry across a sufficient number of bands were not fit successfully. Further, of those YSOs fit, not all had fits with low  $\chi^2$  values. Of the 100 clusters run, 96 contained some YSOs with “good” fits ( $\chi^2 < 3.0$ ). In general, four photometric points were required for a “good” fit to the SED model. In total, 6234 YSOs had reliable model fits across the 96 clusters. Figure 11 shows three examples of SEDFitter model fits to YSOs in the SFOG catalog.

Table 5 lists a sample of the results of the SEDFitter routine for the YSOs in the 96 clusters for which reliable fits were obtained. A number of parameters are presented for each model fit, including the best fit to the object mass, disk mass, age,  $A_V$ , central temperature, disk accretion rate, etc. and the  $\chi^2$  value. The weighted average values of all parameters for fits with  $\chi^2 < 3$  were calculated and are presented for each source. The upper and lower limit model fit parameters are also supplied in each case. The full data table, including all columns, is available online in electronic format.

Given the uncertainties in the distances and the sparse photometry for each source, we do not consider the individual ages and masses derived from the model fits to be entirely reliable and will not discuss them further here. However, cumulatively for all clusters, they can provide an insight into the relative ages and masses of the YSOs in different regions of the outer Galaxy. Figure 12 shows the age and relative mass of the YSOs in three clusters by their spatial distribution. The size of the circles indicates the mass of the YSO relative to the most massive object that we identified in that cluster. The color indicates the age, and the range in ages in Myr is shown by the color bar for each plot. There were no strong trends in the distribution of age or mass across the 618 clusters.

### 5.2. IMF

The IMF of the clusters in the outer Galaxy is of great importance to determine whether the environment of the outer Galaxy has had an effect on the star formation efficiency and rate. Because we have the SED-derived masses for only a small sample of YSOs for each cluster, we examined the IMF of all 96 clusters combined, as shown in Figure 13 (left). We then split the clusters into two groups based on the mass of the most massive identified member, with the cut at  $10 M_\odot$ , as shown in Figure 13 (right). There were 58 low-mass clusters (2929 YSOs) and 38 high-mass clusters (3305 YSOs). The clusters were then split into two groups based on their galactocentric radius, with the division between inner and outer Galaxy placed at 11.5 kpc, based on Huang et al. (2015), as shown in Figure 14. There were 46 clusters with 4533 YSOs with  $R_{\text{Gal}} < 11.5 \text{ kpc}$  and 50 clusters with 1701 YSOs with  $R_{\text{Gal}} > 11.5 \text{ kpc}$ .

In the figures, the red line plots the power-law slope with  $\Gamma = 1.35$ , roughly the value of the Salpeter slope (Salpeter 1955). The blue line shows the Miller & Scalo (1979) broken power-law fit with slopes of  $\Gamma_1 = 1.7$  and  $\Gamma_2 = 2.3$ . The Kroupa (2001) IMF with three distinct values of  $\Gamma$  for the low-, solar-, and

<sup>6</sup> <https://sedfitter.readthedocs.io>

**Table 4**  
SFOG YSO Cluster Descriptions

Cluster Number	Cluster Name	No. of YSOs	Central R.A. (J2000)	Central Decl. (J2000)	Circular Radius (deg)	Associated Region
104	G081.55+1.11	1177	309.049	42.535	1.200	Cygnus-X
237	G133.95+1.00	730	36.725	61.818	0.808	W3/W4
189	G109.96+2.63	704	343.771	62.535	0.826	S155
257	G137.20+1.07	632	42.959	60.566	1.035	W5
357	G189.93+0.50	555	92.373	20.589	0.597	S252
527	G079.72+0.83	519	307.866	40.931	0.989	Cygnus-X
389	G207.07−1.82	497	98.547	4.381	0.817	Rosette; S275, NGC 22
199	G111.67+0.73	444	348.712	61.478	0.688	S158, NGC 75
261	G138.08+1.53	422	44.976	60.561	0.987	W5
523	G078.26+1.05	400	306.513	39.876	0.780	Cygnus-X
375	G192.73−0.00	395	93.322	17.903	0.422	S254/258
360	G188.99+0.93	330	92.277	21.614	0.355	S247
572	G105.64+0.34	304	338.208	58.480	0.505	S138
417	G224.02−1.92	302	106.237	−10.750	0.649	Canis Majori
413	G224.31−0.93	288	107.243	−10.548	0.429	Canis Majori
429	G234.46−0.27	245	112.825	−19.208	0.532	NGC 2343
93	G077.04+0.98	237	305.682	38.838	0.773	Cygnus-X
34	G072.08+2.57	217	300.588	35.577	0.334	IRAS 20003+3524
96	G078.04+2.73	210	304.813	40.961	0.562	Cygnus-X
585	G104.64+0.31	207	336.613	57.932	0.481	IRAS 22246+5750
548	G108.73+0.32	205	343.558	59.915	0.483	CO <sup>a</sup>
316	G173.35−0.16	201	82.043	34.458	0.391	S234
453	G253.95−0.13	193	124.362	−35.807	0.616	GN 08.16.0
395	G212.06−1.14	192	101.444	0.260	0.419	S284
244	G134.79+0.99	185	38.357	61.500	0.781	W4
194	G110.11+0.15	178	346.206	60.336	0.396	IC 1470
123	G084.90+0.42	175	312.649	44.770	0.416	Pelican
190	G108.89+2.68	170	341.671	62.100	0.322	S155
593	G104.52+1.27	169	335.447	58.689	0.303	S135
269	G141.97+1.77	165	51.812	58.761	0.281	AFGL 490
408	G218.12−0.44	163	104.835	−4.822	0.228	S287
94	G076.91+2.05	156	304.449	39.335	0.333	Cygnus-X
566	G106.49+1.00	151	338.995	59.471	0.429	IRAS 22344+5909
286	G148.10+0.20	134	58.939	53.815	0.254	IRAS 03523+5343
122	G084.51+1.03	130	311.629	44.851	0.383	Pelican
66	G074.68+0.58	126	304.414	36.673	0.363	S104
191	G109.86+2.13	125	344.046	62.036	0.292	S155
387	G201.49+0.44	125	97.977	10.374	0.398	IC 446
422	G226.28−0.54	123	108.550	−12.103	0.338	Canis Majori
77	G075.35−0.44	121	305.929	36.634	0.340	DOBASHI 2314
420	G221.95−2.04	117	105.168	−8.962	0.316	DOBASHI 5043
536	G080.01+2.63	110	306.140	42.220	0.435	Cygnus-X
163	G093.44+1.59	108	319.954	51.908	0.406	NRAO 655
577	G103.70+2.15	107	333.188	58.967	0.315	S134
525	G079.00+2.47	105	305.550	41.299	0.405	Cygnus-X
301	G150.68−0.70	103	61.189	51.446	0.421	S206
147	G090.47+2.30	101	315.908	50.231	0.312	L988

**Note.** The clusters are shown in order of decreasing number of YSO members. Only the clusters with more than 100 YSOs are shown here. Table 4 is published in the electronic version in its entirety in a machine-readable format.

<sup>a</sup> Molecular cloud with CO emission identified by Ungerechts et al. (2000).

high-mass regions is shown in yellow, as presented by Weisz et al. (2015), who used a slope of  $\Gamma = 1.45$  above  $2 M_{\odot}$ .

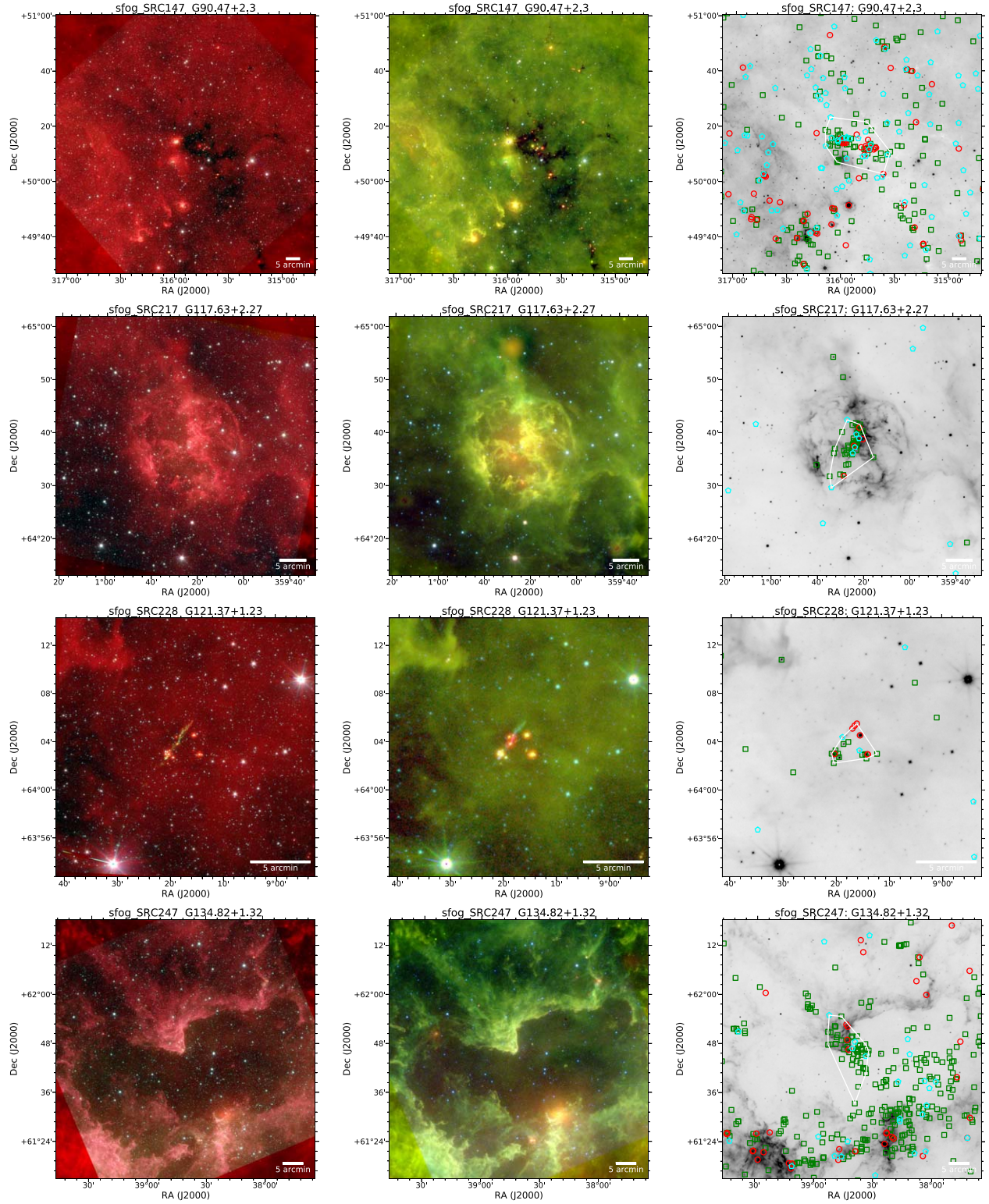
Linear regression fits were made to the high-mass end of the resultant histograms in the range  $3 M_{\odot} < M_{*} < 10 M_{\odot}$ . The slopes of the fits are listed in Table 6.

The slopes for the complete sample of YSOs and the nearby and low-mass clusters are broadly consistent with the values of  $\Gamma$  presented in the literature by Parravano et al. (2018) with  $\Gamma = 1.7 - 2.1$ , as discussed in detail in Paper I. The high-mass

and more distant clusters have shallower slopes than the others, though still consistent with the Salpeter value. The  $1\sigma$  uncertainties overlap for the high-/low-mass cluster slopes, though they do not quite overlap for the near/far cluster slopes. The near/far distance appears to be due to the difference in the modeled populations of the high- and low-mass clusters and not necessarily to a difference in the environment of the outer Galaxy.

To further investigate the comparison of the SEDFitter-generated IMFs of the inner and outer galaxies, we undertook a



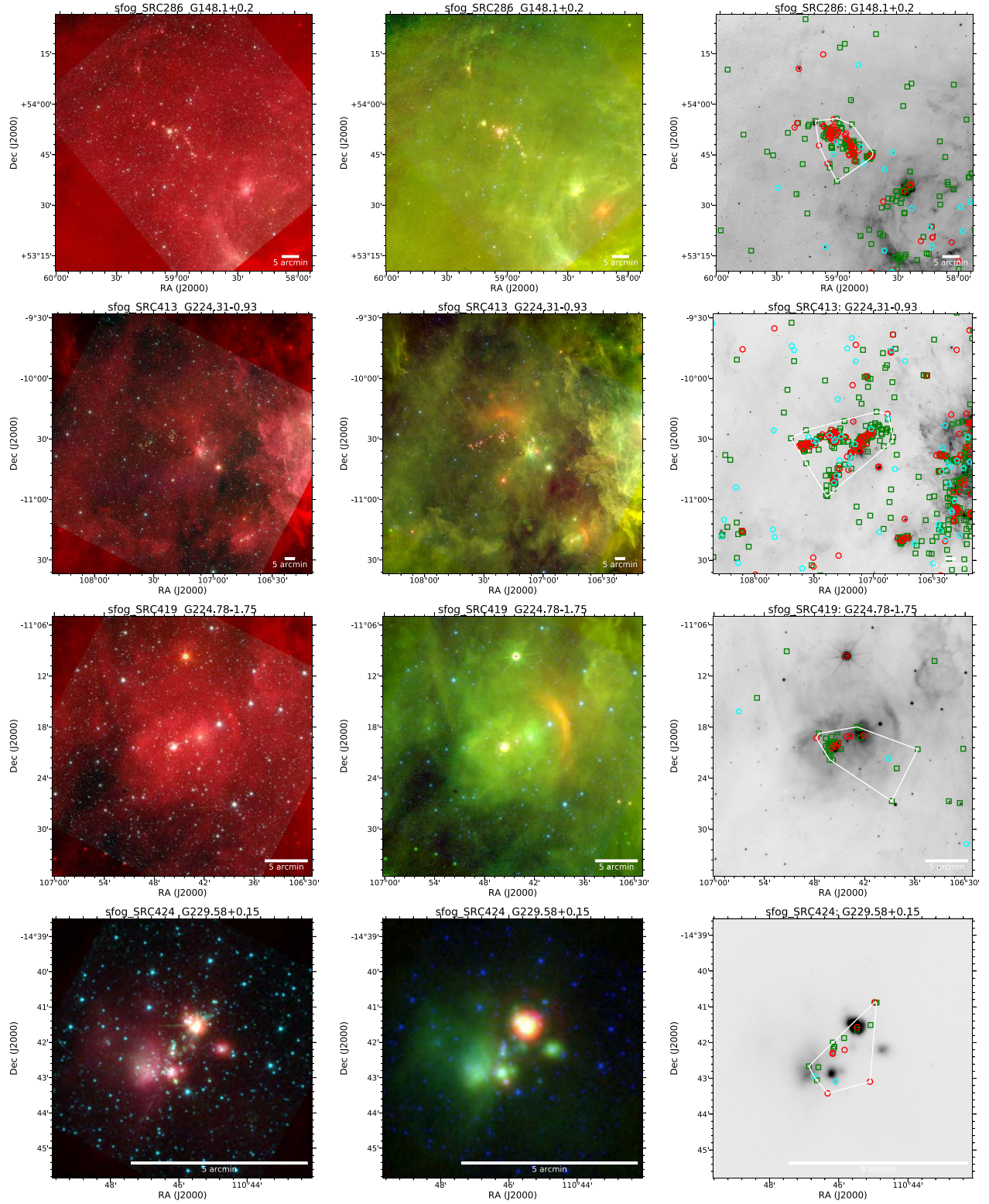


**Figure 9.** Four examples of clusters identified in the SFOG field: Nos. 147, 217, 228, and 247 (top to bottom). In the left column are three-color images in WISE  $12\ \mu\text{m}$  (red), IRAC  $4.5\ \mu\text{m}$  (green), and IRAC  $3.6\ \mu\text{m}$  (blue); the middle column contains three-color images with WISE  $22\ \mu\text{m}$  (red), WISE  $12\ \mu\text{m}$  (green), and IRAC  $4.5\ \mu\text{m}$  (blue); and the right column shows the WISE  $12\ \mu\text{m}$  in reverse gray scale with the identified YSOs and the calculated convex hulls for each cluster overlaid. The symbols show the positions of the class I (red circles), class II (green squares), and class IIa/III (cyan pentagons) YSOs. All of the cluster images and associated FITS files are available from Winston et al. (2020).

comparison of the Cygnus-X Legacy Survey catalog of IRAC and MIPS identified YSOs (Winston et al. 2020). This catalog contains a total of 30,646 YSOs—2029 class I, 27,672 class II, and 945 class IIa/III young stars—across the Cygnus-X north

and south fields. These YSOs were run through the SEDFitter routine in the same way as the SFOG data, with a fixed distance of 1.4 kpc, and the IMF was constructed. The linear regression fit to the data in the range  $3\ M_{\odot} < M_{*} < 10\ M_{\odot}$  was

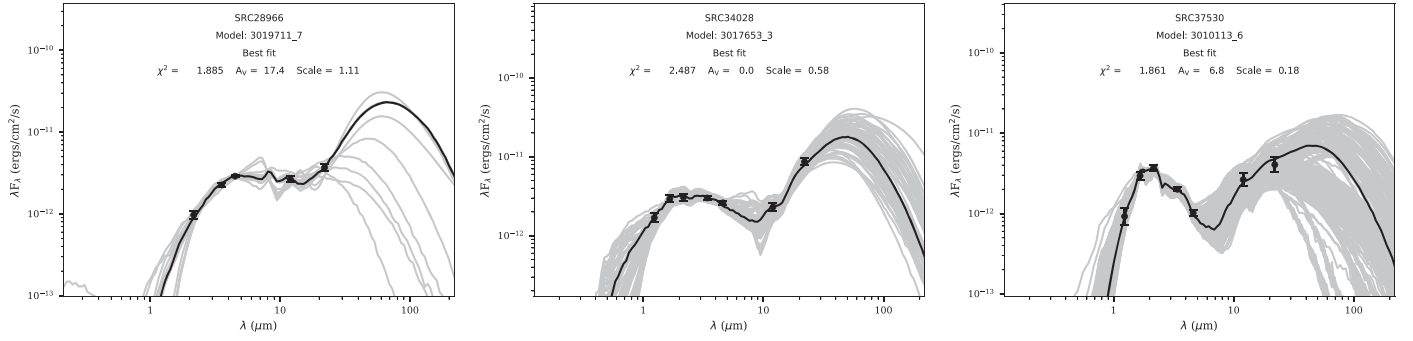




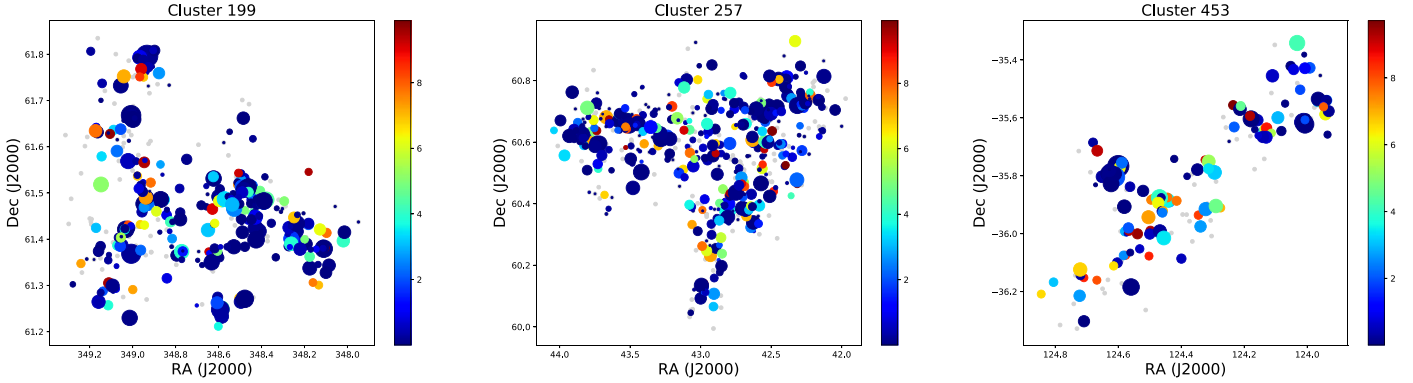
**Figure 10.** Four more examples of clusters identified in the SFOG field: Nos. 286, 413, 419, and 424 (top to bottom). In the left column are three-color images in WISE 12  $\mu\text{m}$  (red), IRAC 4.5  $\mu\text{m}$  (green), and IRAC 3.6  $\mu\text{m}$  (blue); the middle column contains three-color images with WISE 22  $\mu\text{m}$  (red), WISE 12  $\mu\text{m}$  (green), and IRAC 4.5  $\mu\text{m}$  (blue); and the right column shows the WISE 12  $\mu\text{m}$  in reverse gray scale with the identified YSOs and the calculated convex hulls for each cluster overlaid. The symbols show the positions of the class I (red circles), class II (green squares), and class IIa/III (cyan pentagons) YSOs. All of the cluster images and associated FITS files are available from Winston et al. (2020).

$\Gamma = 2.315 \pm 0.298$ . This value is consistent with the literature values and the slope measured for all SFOG clusters and those with  $\text{Max } M_{\odot} < 10$  or  $R_{\text{Gal}} < 11.5$  kpc.

The IMF is generally presumed to be universal; however, the precise origins of the IMF, its relation to the core mass function, and the effects of the galactic and local star-forming



**Figure 11.** Examples of three SEDFitter model fits to YSOs in the SFOG catalog with detections in 2MASS, IRAC, and WISE: SRC28966, SRC34028, and SRC37530 in clusters 15, 85, and 106, respectively.



**Figure 12.** Examples of the SEDFitter results for spatial distributions of ages and masses within each cluster. The relative sizes of each symbol indicate how massive the YSO is with respect to the most massive object in that cluster. The color bar shows how the color scaling relates to the calculated age of the YSO (in Myr), with the bluer objects being younger and the redder objects being older.

**Table 5**  
SFOG Field YSOs: SEDFitter Results

SF ID	Cl. No.	Dist. (kpc)	$R_{\text{Gal}}$ (kpc)	$N_{\text{data}}$	$\chi^2$	$A_V$ (mag)	$M_C$ ( $M_{\odot}$ )	Age (yr)	$M_{\text{disk}}$ ( $M_{\odot}$ )	$M_{\text{env}}$ ( $M_{\odot}$ )	$\dot{M}$ ( $M_{\odot}$ )	$T_*$ (K)	$L_*$ ( $L_{\odot}$ )
SRC588	5	3.4	7.8	5	2.475	0.213	2.492	3.900e+06	1.008e-05	5.367e-04	...	7.985e+03	4.251e+01
SRC591	5	3.4	7.8	5	1.856	0.000	1.068	3.815e+05	7.670e-05	3.610e-03	7.183e-07	4.264e+03	5.582e+00
SRC597	5	3.4	7.8	5	0.322	5.213	1.981	8.221e+06	3.102e-05	2.993e-06	...	7.854e+03	1.186e+01
SRC600	5	3.4	7.8	5	0.210	5.039	1.470	1.192e+06	1.645e-02	1.321e-04	...	4.537e+03	3.165e+00
SRC601	5	3.4	7.8	5	0.829	3.181	1.638	7.610e+06	8.970e-03	4.193e-08	...	5.322e+03	3.185e+00
SRC608	5	3.4	7.8	4	0.007	11.30	0.540	3.042e+05	9.975e-04	3.223e-03	2.440e-07	3.822e+03	2.954e+00
SRC613	5	3.4	7.8	5	0.776	0.345	1.915	6.945e+06	3.301e-05	2.837e-08	...	6.449e+03	1.303e+01
SRC616	5	3.4	7.8	4	0.393	8.739	2.009	6.060e+04	3.965e-03	1.839e-01	9.353e-06	4.316e+03	3.828e+01
SRC617	5	3.4	7.8	4	0.491	7.324	8.147	6.542e+03	4.394e-03	3.207e+01	1.131e-03	4.381e+03	9.556e+02
SRC622	5	3.4	7.8	5	0.971	0.157	3.736	1.364e+04	1.649e-02	6.810e+00	1.486e-04	4.327e+03	1.446e+02

**Note.** This table is available in its entirety in FITS format.

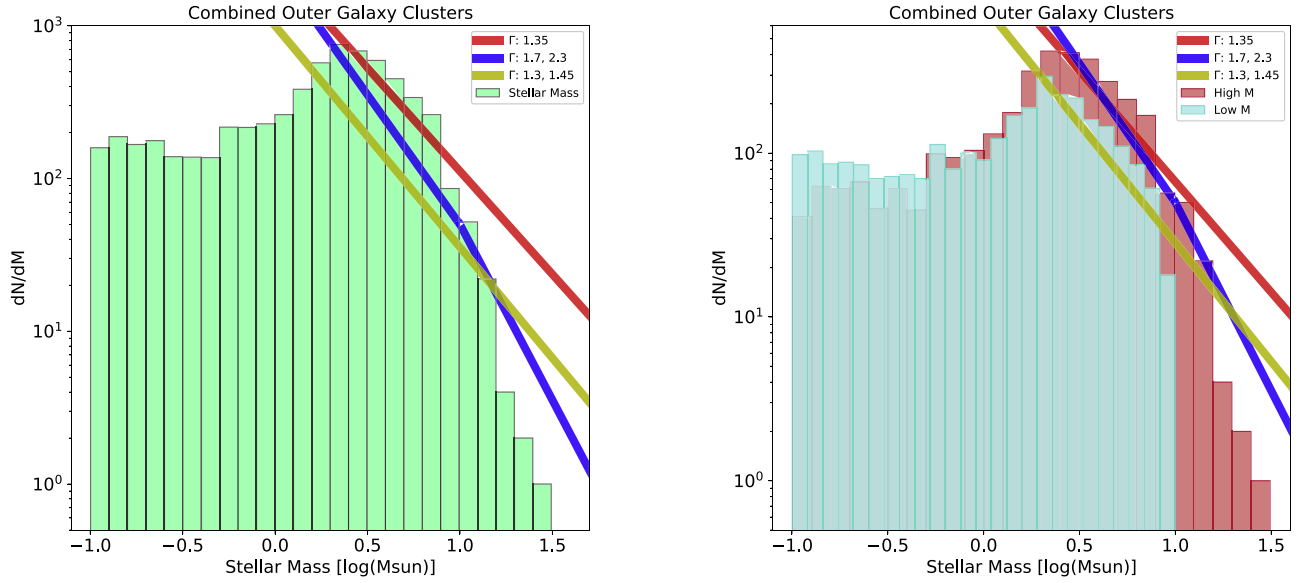
environment on its evolution are not fully understood. Paladini et al. (2019) provided a useful overview of our current understanding in their recent white paper. An accurate and in-depth study of the IMF in the outer Galaxy is beyond the scope of this paper. The results presented here should be taken as an indication that the IMF in the outer Galaxy is similar to that of the inner Galaxy, and that this may indicate that metallicity does not greatly impact the IMF. However, there are a number of caveats, both astrophysical and analytical, that affect the SEDFitter-derived IMF that we will now outline.

Variations in the IMF of an astrophysical origin could be due to, e.g., metallicity (e.g., Kroupa 2019), feedback from massive stars in the cluster either triggering or impeding star formation

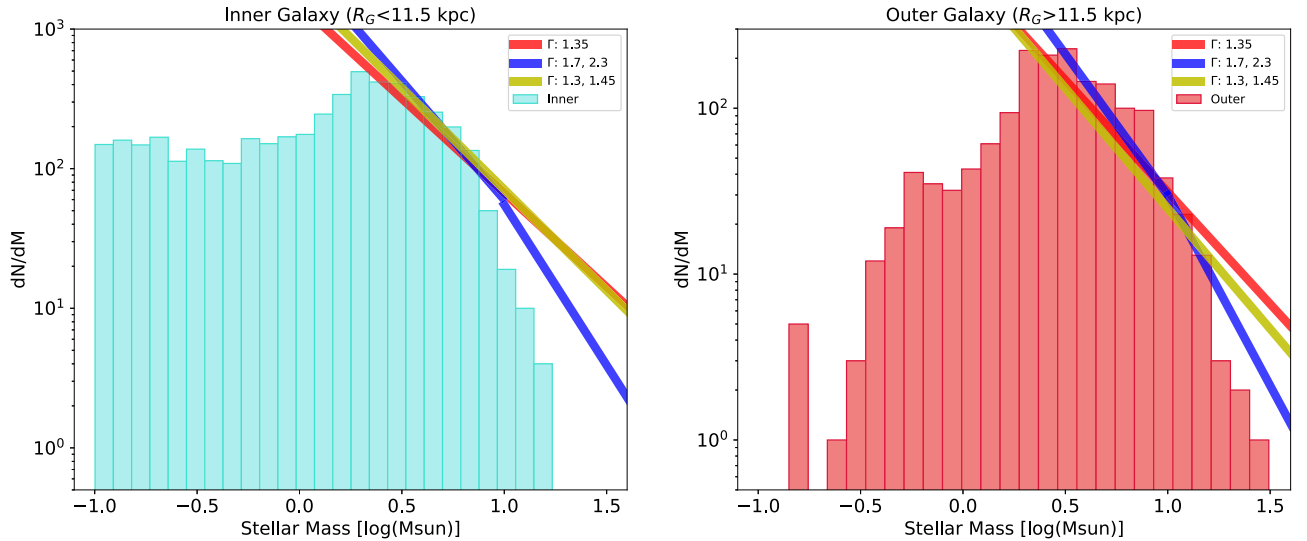
(e.g., Walch et al. 2013; Krumholz et al. 2016), jets and outflows (e.g., Cunningham et al. 2018), disk evolution (e.g., Povich et al. 2016), and clustered versus distributed star formation (e.g., Bonnell et al. 2011). Recent theory suggests that massive stars may form a few Myr after initial low-mass star formation begins, which would also lead to a variation in IMF with cluster age (Vázquez-Semadeni et al. 2019). Faster disk evolution in intermediate and massive stars has also been reported by Povich et al. (2016) in the M17 SWex star-forming region that would also bias the slope of the IMF in this mass regime.

The limitations of the data can impact the accuracy of the SEDFitter results. The distance estimates to the clusters are





**Figure 13.** Examination of the IMF for the outer Galactic regions identified in the SFOG field. Left: IMF as determined by combining the SEDFitter calculated masses for the 96 clusters with distance estimates and with members with  $\chi^2 < 3$  fits. The green histogram shows the best-fit model masses for all YSOs. The data show a reasonable correlation to both the Salpeter slope of the IMF ( $m \sim -1.35$ ) and more recent estimates ( $m \sim -2.7, -2.3$ ) for a broken power-law fit. Right: IMF split into the 58 clusters with a highest-mass member lower than  $10 M_{\odot}$  and 38 clusters with a highest-mass member greater than  $10 M_{\odot}$ .



**Figure 14.** Examination of the IMF for the outer Galactic regions identified in the SFOG field. Left: IMF for the 46 clusters within a galactocentric radius of 11.5 kpc. Right: IMF for the 50 clusters at a galactocentric radius greater than 11.5 kpc.

**Table 6**  
IMF Slopes and Uncertainties

Clusters	$\Gamma$	$\epsilon$
All clusters	1.918	0.419
Max $M_{\odot} < 10$	1.929	0.355
Max $M_{\odot} > 10$	1.275	0.345
$R_{\text{Gal}} < 11.5$ kpc	1.873	0.308
$R_{\text{Gal}} > 11.5$ kpc	1.146	0.241

derived from kinematic distances, which can have uncertainties of at least 10%–20% and are subject to systematic uncertainties (Anderson et al. 2014). Cantat-Gaudin & Anders (2020) found in their Gaia DR2 study of open clusters that not all previously identified members are physically associated with the regions;

this would also affect the SEDFitter results, which assume a single distance to each cluster.

Further, each cluster also has a spatial “depth,” so a YSO at the near edge of the region may lie a few hundred pc nearer than one at the far edge. The clusters are all of different ages, and within each cluster, the subclusters may also have a range in age. Thus, the disk fraction, and hence percentage of the population detected with Spitzer, will vary; this should be accounted for when stacking the clusters in the IMF but cannot be done here.

There are also the issues of varying levels of diffuse mid-IR emission and source crowding in both Spitzer and, to a greater extent, the WISE data. Diffuse emission reduces sensitivity in the IRAC bands, meaning fainter sources will not be detected. Source crowding is primarily a spatial resolution issue and most prevalent in the central cores of clusters. This leads to an

underestimation of the disk-bearing population, particularly of the lower-mass YSOs.

The mid-IR data also do not provide a measure of the diskless population of the clusters. These populations are generally identified using spectroscopy or X-ray observations. The coverage of X-ray studies is highly constrained in comparison to IR surveys and is generally focused on the centers of known star-forming regions (Winston et al. 2009, 2011; Broos et al. 2013). These studies show large populations of YSOs without disks and trends in disk fractions with proximity to OB stars.

Our selection method is not very sensitive to massive stars and may not detect all of the known massive stars in a region. This would lead to a steeper IMF slope at higher masses. A comparison with Skiff (2014) shows that of the 3302 objects in their catalog with spectral types O or B that were located within the convex hulls of the clusters, 591 were matched to a YSO in the SFOG catalog within  $2''$ , or 18% of the known massive population.

The accuracy of the SEDFitter analysis is limited by the small number of near- and mid-IR bands available for model fitting and the reddening across those bands. Povich et al. (2019) found a degeneracy between stellar effective temperature and extinction that can lead to inaccurately modeled masses. The presence of disk material adds to this issue, especially as disk/envelope masses are poorly constrained without far-IR and longer photometric data.

## 6. Comparison to Other Catalogs

The SFOG catalog of YSOs and identified clusters was compared to a number of other published surveys and databases covering the outer Milky Way.

### 6.1. Other Databases of YSOs and Clusters

The individual YSO candidates were matched to the online SIMBAD database to within a  $2''$  radius for all 47,338 YSOs, with 7039 matches. Of these, 4428 had been previously identified as a type of YSO, leaving 42,910 possibly new YSO candidates (some previous studies may not be listed on the SIMBAD database). Table 3 provides a sample listing of the YSOs with matches within  $2''$  of a SIMBAD source, giving the source identification and the object type. The full table is available online in electronic format and includes further selected information pertaining to the SIMBAD objects.

In their recent paper, Armentrout et al. (2020) listed 166 H II regions previously thought to be radio-quiet that they confirmed to be weak radio sources. Of the 166, 37 were found to lie within the convex hull of one of the 618 clusters in the SFOG field. Twenty of these matches were found between  $90^\circ$  and  $115^\circ$  Galactic longitude. There was no significant difference in the median number of YSOs or the median effective hull radius between the clusters in the radio-quiet Armentrout sample or the WISE H II sample. Further, when matching the H II regions to the catalog of YSOs, it was found that 58 had a YSO located within  $10''$  and 107 had a YSO located within  $1'$ .

The positions of the clusters were also compared to those of Avedisova (2002), who reported 66,887 clusters over the whole Galactic plane. Of these, 24,101 were within the boundaries of the SFOG field. Of the 618 clusters we found in the SFOG

field, 260 were matched to one or more of the Avedisova (2002) clusters, while 358 were new to this survey.

### 6.2. GAIA DR2

The Gaia DR2 was also searched for matches to the SFOG catalog. Of the 47,338 YSOs identified by SFOG, 25,919 had a Gaia counterpart within  $0''.5$ . These matched sources were then compared to the Bailer-Jones et al. (2018) catalog of calculated parallax distances for Gaia DR2 to determine the individual distances to the SFOG YSOs. Of the 25,919 matches, 23,424 had a distance listed in Bailer-Jones et al. (2018). A comparison was then made to the 100 identified clusters with known distances to determine if the distances of the YSOs in those clusters corresponded to the WISE H II catalog distances. The clustered YSOs showed both a large range in individual uncertainty of a few kiloparsecs and a spread within the cluster YSOs of a few kiloparsecs. While some individual YSOs had distances corresponding to that from WISE, the median and average values for the cluster did not. It can be noted that the Gaia distance uncertainties increase greatly beyond  $\sim 0.5$  kpc, and this may account for the lack of consistency in cluster distance estimates. Given the lack of consistency in the cluster distances, it was decided not to use these in the SEDFitter routine or to report them in the catalog.

Cantat-Gaudin & Anders (2020) calculated Gaia DR2 distances to previously identified open clusters using the mode of distance likelihoods of known members. Of these, 27 matched to an SFOG cluster and 12 had distances in the WISE H II catalog. The differences in the measured distances varied widely (from 250 to 7300 pc). Given that young clusters are embedded and Gaia is more likely to detect foreground stars or the front cluster population, we decided to use the kinematic distances in this paper.

### 6.3. Other IR Surveys

The first comparison was made to the SMOG field of Paper I to assess the effect of the loss of the 5.8 and  $8\ \mu\text{m}$  IRAC bands on the detection rate of YSOs. The SMOG field 2MASS +IRAC selection technique identified 3835 YSOs. The same GLIMPSE360 criteria used in this paper identified 1512 YSOs. Of these, 1102 objects matched between the two catalogs. This suggests that we are detecting 29% of the original catalog with the GLIMPSE360 criteria. Much of this difference in the number of YSOs detected is attributable to the more stringent  $3.6\ \mu\text{m}$  cut around 14 mag in the SFOG selection criteria; many of the undetected SMOG YSOs are fainter and redder than those selected by the SFOG criteria. This would imply that the total YSO population may be at least as high as 73,247 across the field to the depth of the SMOG observations. Of the remaining 410 unmatched YSOs that were identified with the new method and not in the original catalog, about 70 fall on the edges of the field where there was no overlap between the two IRAC fields of view and thus were not in the original “cleaned” source catalog after contaminant removal. A further  $\sim 30$  SFOG objects fall in the galactic contaminant color space of the IRAC four-band CCD. The rest show a weak  $[3.6-4.5]$  color excess that was not deemed sufficient for selection in the original SMOG criteria. The spatial distribution of the unmatched sources shows that the majority ( $>75\%$ ) are associated with matched YSOs and/or identified clusters.

The catalog of Tóth et al. (2014) AKARI YSOs was compared to the SFOG YSOs. The AKARI catalog was based on the Far-Infrared Surveyor All-Sky Survey catalog composed of photometry at four IR-wavelength bands centered at 65, 90, 140, and 160  $\mu\text{m}$ . Of the 44,001 AKARI YSOs, 14,986 were located in the SFOG field. Of these, 49 matched to the SFOG catalog within  $2''$  and 261 to within  $5''$ . However, given the AKARI spatial resolution of  $1'-1.5'$  at these wavelengths, the AKARI sources contain many individual YSOs and likely significant emission from the clouds surrounding these clusters, as well as deeply embedded YSOs, and are therefore not likely true or unique matches to the IRAC-identified YSOs in the SFOG survey.

We compared the WISE-identified YSOs in SFOG to the Marton et al. (2016) WISE single vector machine-selected YSO candidates. They identified 133,980 class I/II YSOs across the whole sky, with about 16,945 candidate YSOs within the SFOG field. Within a radius of  $1''.5$ , we match 5537 of 20,892 of our WISE-detected YSOs, or 26%, to the Marton et al. (2016) sample. Including the full SFOG catalog, we match 6425 of 47,338, or 14%, of our YSOs to the Marton sample. A comparison of the CCDs of the two samples of YSOs shows that ours has a more conservative color cut, with the Marton sample selecting YSOs with  $\text{WISE1} - \text{WISE2} < 0$ , which accounts for most of the difference in the catalogs.

We compared our detection of YSOs in the W5 star-forming region to that reported in Koenig et al. (2008), who published a Spitzer IRAC and MIPS survey of the region conducted during the cryogenic mission phase. They identified three deeply embedded objects, 171 class I, 1809 class II, 79 transition disks (for a total of 2062 IR-excess YSOs), and 15,709 class IIa/III stars. Of these, we match 1091 objects within a  $1''$  radius: 88 class I (52%), 947 class II (52%), 18 transition disks (23%), and 38 class IIa/III (0.2%); none were deeply embedded objects. The Koenig et al. (2008) class IIa/III objects have weak excess and are positionally associated with the region, hence the very low detection rate in the SFOG sample. The 50% detection rate for the IR-excess YSOs roughly corresponds to the brighter YSOs in the two shorter IRAC bands. The spatial distributions of the YSOs and the identified clusters match closely to those found in this paper.

The SFOG catalog was also compared to the Rivera-Ingraham et al. (2011) four-band IRAC and MIPS 24  $\mu\text{m}$  survey of the W3 star formation region. They identified 1566 YSOs, of which we match 446 SFOG YSOs to within  $1''$  radius. The Rivera-Ingraham et al. (2011) evolutionary classes are 184 class 0/I, 560 deeply embedded class 0/I, 549 class II, and 273 embedded class II. We match 65 (35%), 35 (2.2%), 343 (63%), and 3 (1.1%) of these classes, respectively. The SFOG catalog does not identify the deeply embedded YSOs; some of these were selected using Spitzer MIPS photometry and so may not show excess emission in the shorter wavelengths. The remainder were below 14th magnitude in the IRAC 3.6  $\mu\text{m}$  channel and so were not selected in SFOG. However, the spatial distribution of the W3 YSOs in the SFOG catalog traces a similar cluster distribution to that of Rivera-Ingraham et al. (2011).

We made a further comparison of the SFOG catalog to the Rebull et al. (2011) Spitzer IRAC and MIPS survey of the North American and Pelican nebula star-forming regions. They reported a total of 2196 YSOs in the field, with 262 that lie in the overlap region with the SFOG catalog field. Of these,

132 are matched to within a  $1''$  radius of an SFOG YSO, being again the brighter 3.6 and 4.5  $\mu\text{m}$  sources.

We thus draw two conclusions from the comparisons to these well-studied regions and other IR-based YSO catalogs. First, we are finding roughly half of the previously identified YSOs: those that are less embedded and brighter at 3.6  $\mu\text{m}$ . We also used more conservative color cuts and brighter magnitude limits to minimize the number of spurious YSO identifications. Second, from the YSOs we do detect, we find that we are reliably identifying the overall spatial structure and main clusters in these regions that were found in the other surveys that used deeper integrations or had coverage in the IRAC 5.8 and 8  $\mu\text{m}$  and MIPS 24  $\mu\text{m}$  bands.

## 7. Summary

We have undertaken a study of the 600  $\text{deg}^2$  SFOG field comprising the GLIMPSE360 and SMOG survey regions. We combined the Spitzer data with 2MASS near-IR photometry and used the WISE catalog of the field to identify more embedded YSOs.

1. We identify 42,757 YSOs with IR-excess emission in the GLIMPSE360 and WISE data. When combined with the SMOG field and after removing the spurious WISE sources, we find a total of 47,338 YSOs.
2. The evolutionary class of the YSOs was determined from the SED slope: 10,461 class I, 29,552 class II, and 7325 class IIa/III.
3. We identify 618 reliable clusters in the SFOG field. The ratio of YSOs identified as members of clusters was 25,528/47,338, or 54%. The smallest cluster has five members, and the largest has 1177 members, with a median size of 17 YSOs. Of the 618 clusters, 47 have more than 100 members, and 22 have more than 200 members.
4. One hundred clusters had a distance estimate from H II regions within their convex hulls. The SEDs of the YSOs in these clusters were fitted using the SEDFitter routine; of these, 96 had reliable fits.
5. From the modeled masses, the IMF was constructed for the clusters across the SFOG field. The slope of the combined IMF was found to be  $\Gamma = 2.38 \pm 0.20$  above  $3 M_{\odot}$ . Dividing the clusters by galactocentric distances, the slopes were  $\Gamma = 1.87 \pm 0.31$  above  $3 M_{\odot}$  for  $R_{\text{Gal}} < 11.5 \text{ kpc}$  and  $\Gamma = 1.15 \pm 0.24$  above  $3 M_{\odot}$  for  $R_{\text{Gal}} > 11.5 \text{ kpc}$ . These values are consistent with each other within the uncertainties and with those obtained in the inner Galaxy high-mass star formation regions. The slopes are likely also consistent with a universal Salpeter IMF.

This work is based on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract 1407. We gratefully acknowledge funding support for this work from NASA ADAP grant NNX16AF37G. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. Support for the IRAC instrument



was provided by NASA through contract 960541 issued by JPL. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. This research made use of Montage. It is funded by the National Science Foundation under grant No. ACI-1440620 and was previously funded by the National Aeronautics and Space Administration's Earth Science Technology Office, Computation Technologies Project, under cooperative agreement No. NCC5-626 between NASA and the California Institute of Technology. This research has made use of NASA's Astrophysics Data System.

*Facility:* Spitzer (IRAC).

*Software:* Astropy (Astropy Collaboration et al. 2013), SAOimageDS9 (Joye & Mandel 2003), Montage (Berriman et al. 2008), SEDFitter (Robitaille et al. 2007).

### Appendix A IRAC and 2MASS Source Selection

The removal of contaminating sources and the selection of YSOs in the GLIMPSE360 field were undertaken in a different manner to those implemented for the SMOG field, with IRAC four-band coverage that was based on the methods of Gutermuth et al. (2008, 2009). The locations of the SMOG field contaminants in color space were used as the basis for identifying the locations of the contaminants in the two-band IRAC data of GLIMPSE360.

Background galaxy contaminants, including candidate AGNs and PAH galaxies, and saturated sources were removed using a cut in color–magnitude space:

$$\begin{aligned} [3.6] &< 6.0 \text{ and} \\ [4.5] &< 5.5 \text{ and} \\ [3.6] &> 16.0 \text{ and} \\ [3.6] &> 14.0 \text{ and } [3.6 - 4.5] < 0.5 \text{ and} \\ [3.6] &> 2([3.6 - 4.5] + 0.5) + 14. \end{aligned} \quad (\text{A1})$$

The full GLIMPSE360 archive contained 49,378,049 sources. The catalog contained 7,527,352 objects after contaminants were removed. The SFOG fields lie in the direction of the outer Galaxy, where the level of shielding from the Galactic center from both the stellar population and the dust component is reduced and thus the extragalactic background is expected to be higher.

Three combinations of 2MASS and IRAC bands were used to select for less extincted objects. The photometry was first dereddened, and only sources with good values of extinction were included for selection:

$$(A_H - A_{4.5}/A_J - A_H)([J - H] - 0.6 + \sigma_{JH}) + 1.0 + \sigma_{H4.5} < [H - 4.5] \text{ and } [J - H] > 0, \quad (\text{A2})$$

$$\begin{aligned} (A_K - A_{4.5}/A_H - A_K)([H - K] + \sigma_{HK}) + 0.4 + \sigma_{K4.5} &< [K - 4.5] \text{ and } [H - K] > 0 \text{ and} \\ [K - 4.5] &> 0.2 + \sigma_{K4.5}, \end{aligned} \quad (\text{A3})$$

$$\begin{aligned} [3.6 - 4.5] - \sigma_{3.6,4.5} &> 0 \text{ and} \\ [K - 3.6] - \sigma_{K3.6} &> 0.2 * [3.6 - 4.5] + 0.3 \text{ and} \\ [K - 3.6] - \sigma_{K3.6} &> -1.0([3.6 - 4.5] - \sigma_{12}) + 0.8. \end{aligned} \quad (\text{A4})$$

For these three selection criteria, 20,339, 24,560, and 16,941 candidate YSOs were selected, respectively, for a combined total of 28,837 YSOs identified using the GLIMPSE360 photometry.

### Appendix B WISE Source Selection

Following the process laid out by Fischer et al. (2016), spurious detections were cleaned from the catalog. The first step was to remove those sources with uppercase flags in bands W1, W2, and W3. Upper limits in bands W1, W2, and W3 were then removed, and a saturation cutoff of  $W1 > 5$  was applied to the data. The initial catalog contained 14,483,596 sources. The remaining catalog contained 543,457 sources, or  $\approx 4\%$  of the original catalog. The contaminating background galaxies and source-selection criteria for the WISE data were taken from Koenig & Leisawitz (2014) and Fischer et al. (2016) and adapted to the requirements of the GLIMPSE360 field. We slightly adjusted the criteria for removal of AGNs and star-forming galaxy contaminants from those of Koenig & Leisawitz (2014) as follows:

$$\begin{aligned} \text{SFG} &= [W2 - W3] > 2.3 \text{ and} \\ [W1 - W2] &< 1.0 \text{ and} \\ [W1 - W2] &< 0.46([W2 - W3] - 0.78) \text{ and} \\ [W1] &> 14, \end{aligned} \quad (\text{B1})$$

$$\begin{aligned} \text{AGN} &= [W1] > 1.8([W1 - W3] + 4.1) \text{ and} \\ [W1] &> 14 \text{ or} \\ [W1] &> [W1 - W3] + 10.0. \end{aligned} \quad (\text{B2})$$

Of those 14,483,596 sources, 13,940,139 were identified as contaminants, leaving a cleaned catalog of 543,457 sources.

The cleaned catalog was then searched for YSOs according to criteria taken from the Fischer et al. (2016) and Koenig & Leisawitz (2014) papers. From these, there were 2245 transition disk candidates and 7094 YSOs identified from the four-band WISE diagram and 7244 class I and 10,447 class II sources identified from the WISE three-band diagram. There was a total of 20,892 candidate YSOs identified with WISE.

### Appendix C IRAC+WISE Source Selection

The IRAC+WISE YSOs were selected following the same selection cutoffs as were applied to the WISE sources. The cleaned WISE catalog was matched to the contaminant-removed IRAC catalog, and sources with a  $1''.5$  or closer spatial coincidence were considered to be the same object, giving 318,588 objects in the joined catalog. The YSOs were then selected by replacing WISE bands 1 and 2 with IRAC bands 1 and 2 and replacing this photometry in the WISE CCD YSO selections. The 2MASS selection criteria were not applied here, as they replicate the IRAC+2MASS selection from the GLIMPSE360 data. A total of 11,196 candidate YSOs were selected using this method.

### ORCID iDs

Elaine Winston  <https://orcid.org/0000-0001-9065-6633>  
Joseph L. Hora  <https://orcid.org/0000-0002-5599-4650>  
Volker Tolls  <https://orcid.org/0000-0003-1841-2241>

## References

- Allen, L., Bourke, T., Brooke, T., et al. 2006, Gould's Belt: Star Formation in the Solar Neighborhood, Spitzer Proposal, [30574](#)
- Allen, L., Calvet, N., D'Alessio, P., et al. 2004, [ApJS](#), **154**, 363
- Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, [ApJS](#), **212**, 1A
- Armentrout, W., Anderson, L., Wegner, T., Balser, D., & Bania, T. 2020, [ApJ](#), submitted
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, [A&A](#), **558**, A33
- Avedisova, V. S. 2002, [ARep](#), **46**, 193
- Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., et al. 2018, [AJ](#), **156**, 58
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, [ARA&A](#), **48**, 339
- Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, [PASP](#), **115**, 953
- Berriman, G. B., Good, J. C., Laity, A. C., et al. 2008, in ASP Conf. Ser. 394, Astronomical Data Analysis Software and Systems XVII, ed. R. W. Argyle, P. S. Bunclark, & J. R. Lewis (San Francisco, CA: ASP), 83
- Bloemen, J. B. G. M., Bennett, K., Bignami, G. F., et al. 1984, [A&A](#), **135**, 12
- Bonnell, I. A., Smith, R. J., Clark, P. C., et al. 2011, [MNRAS](#), **410**, 2339
- Bressert, E., Ginsburg, A., Bally, J., et al. 2012, [ApJL](#), **758**, L28
- Broos, P. S., Getman, K. V., Povich, M. S., et al. 2013, [ApJS](#), **209**, 32
- Cantat-Gaudin, T., & Anders, F. 2020, [A&A](#), **633**, A99
- Carey, S., Ali, B., Berriman, B., et al. 2008, Spitzer Mapping of the Outer Galaxy (SMOG), Spitzer Proposal, [50398](#)
- Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, [PASP](#), **121**, 76
- Casali, M., Adamson, A., Alves de Oliveira, C., et al. 2007, [A&A](#), **467**, 777
- Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, [PASP](#), **121**, 213
- Cooper, H. D. B., Lumsden, S. L., oudmaijer, R. D., et al. 2013, [MNRAS](#), **430**, 1125
- Cunningham, A. J., Krumholz, M. R., McKee, C. F., et al. 2018, [MNRAS](#), **476**, 771
- Dawson, S. A., & Schroder, K.-P. 2010, [MNRAS](#), **404**, 917
- Dunham, M. M., Allen, L. E., Evans, N. J., II, et al. 2015, [ApJS](#), **220**, 11
- Ester, M., Kriegl, H.-P., Sander, J., et al. 1996, in KDD'96: Proc. Second Int. Conf. on Knowledge Discovery and Data Mining, ed. E. Simoudis, J. Han, & U. Fayyad (Palo Alto, CA: AAAI Press), 226
- Evans, N. J. 1999, [ARA&A](#), **37**, 311
- Evans, N. J., Allen, L. E., Blake, G. A., et al. 2003, [PASP](#), **115**, 965
- Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, [ApJS](#), **154**, 10
- Fischer, W. J., Padgett, D. L., Stapelfeldt, K. L., et al. 2016, [ApJ](#), **827**, 96
- Ginsburg, A., Bressert, E., Bally, J., et al. 2012, [ApJL](#), **758**, L29
- Gutermuth, R., Megeath, S. T., Muzerolle, J., et al. 2004, [ApJS](#), **154**, 374
- Gutermuth, R., Megeath, S. T., Myers, P. C., et al. 2009, [ApJS](#), **184**, 18
- Gutermuth, R., Myers, P. C., Megeath, S. T., et al. 2008, [ApJ](#), **674**, 336
- Hambly, N. C., Collins, R. S., Cross, N. J. G., et al. 2008, [MNRAS](#), **384**, 637
- Hewett, P. C., Warren, S. J., Leggett, S. K., et al. 2006, [MNRAS](#), **367**, 454
- Hodgkin, S. T., Irwin, M. J., Hewett, P. C., et al. 2009, [MNRAS](#), **394**, 675
- Hora, J. L., Bontemps, S., Megeath, S. T., et al. 2009, AAS Meeting Abstracts, **213**, 356.01
- Huang, Y., Liu, X. W., Zhang, H. W., et al. 2015, [RAA](#), **15**, 1240
- Hurt, R. 2008, Annotated Roadmap to the Milky Way, <http://www.spitzer.caltech.edu/images/1925-ssc2008-10b-A-Roadmap-to-the-Milky-Way-Annotated>
- Joncour, I., Duchêne, G., Moraux, E., et al. 2018, [A&A](#), **620**, A27
- Joye, W. A., & Mandel, E. 2003, in ASP Conf. Ser., Astronomical Data Analysis Software and Systems XII, ed. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook (San Francisco, CA: ASP), 489
- Koenig, X. P., Allen, L. E., Gutermuth, R. A., et al. 2008, [ApJ](#), **688**, 1142
- Koenig, X. P., & Leisawitz, D. T. 2014, [ApJ](#), **791**, 131
- Kroupa, P. 2001, [MNRAS](#), **322**, 231
- Kroupa, P. 2019, The Systematically Varying IMF and Some Consequences, Zenodo, doi:[10.5281/zenodo.3756450](#)
- Krumholz, M. R., Myers, A. T., Klein, R. I., et al. 2016, [MNRAS](#), **460**, 3272
- Lada, C. J. 1987, in IAU Symp. 115, Star-forming Regions, ed. M. Peimbert & J. Jugaku (Dordrecht, NL: D. Reidel), 1
- Lada, C. J., & Wilking, B. A. 1984, [ApJ](#), **287**, 610
- Lada, C. J., Muench, A. A., Luhman, K. L., et al. 2006, [AJ](#), **131**, 1574
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, [MNRAS](#), **379**, 1599
- Mainzer, A., Bauer, J., Grav, T., et al. 2011, [ApJ](#), **731**, 53
- Majewski, S., Babler, B., Churchwell, E., et al. 2007, Galactic Structure and Star Formation in Vela-Carina, Spitzer Proposal, [40791](#)
- Marton, G., Toth, L. V., Paladini, R., et al. 2016, [MNRAS](#), **458**, 3479
- Mead, K. N., & Kutner, M. L. 1988, [ApJ](#), **330**, 399
- Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2012, [AJ](#), **144**, 192
- Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, [AJ](#), **151**, 5
- Meyer, M. R., Backman, D., Beckwith, S. V. W., et al. 2004, The Formation and Evolution of Planetary Systems: Placing Our Solar System in Context, Spitzer Proposal, [148](#)
- Miller, G. E., & Scalo, J. M. 1979, [ApJS](#), **41**, 513
- Miville-Deschênes, M.-A., & Lagache, G. 2005, [ApJS](#), **157**, 302
- Paladini, R., Povich, M., Armus, L., et al. 2019, [BAAS](#), **51**, 335
- Parravano, A., Hollenbach, D., & McKee, C. F. 2018, [MNRAS](#), **480**, 2449
- Povich, M. S., Maldonado, J. T., Haze Nuñez, E., et al. 2019, [ApJ](#), **881**, 37
- Povich, M. S., Townsley, L. K., Robitaille, T. P., et al. 2016, [ApJ](#), **825**, 125
- Rebull, L. M., Guieu, S., Stauffer, J. R., et al. 2011, [ApJS](#), **193**, 25
- Rivera-Ingraham, A., Martin, P. G., Polychroni, D., et al. 2011, [ApJ](#), **743**, 39
- Robin, A. C., Reyle, C., Derriere, S., et al. 2003, [A&A](#), **409**, 523
- Robin, A. C., Reyle, C., Fliri, J., et al. 2014, [A&A](#), **569**, A13
- Robitaille, T. P., Whitney, B. A., Indebetouw, R., et al. 2007, [ApJS](#), **169**, 328
- Rudolph, A. L., Simpson, J. P., Haas, M. R., et al. 1997, [ApJ](#), **489**, 94
- Salpeter, E. E. 1955, [ApJ](#), **121**, 161
- Saral, G., Hora, J. L., Audard, M., et al. 2017, [ApJ](#), **839**, 108
- Skiff, B. A. 2014, [yCat](#), **2023**, 0
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, [AJ](#), **131**, 1163
- Tóth, L. V., Marton, G., Zahorecz, S., et al. 2014, [PASJ](#), **66**, 17
- Ungerechts, H., Umbanhowar, P., & Thaddeus, P. 2000, [ApJ](#), **537**, 221
- Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., et al. 2019, [MNRAS](#), **490**, 3061
- Walch, S., Whitworth, A. P., Bisbas, T. G., et al. 2013, [MNRAS](#), **435**, 917
- Weisz, D. R., Clifton Johnson, L., Foreman-Mackey, D., et al. 2015, [ApJ](#), **806**, 198
- Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, [ApJS](#), **153**, 1
- Whitney, B., Arendt, R., Babler, B., et al. 2008, GLIMPSE360: Completing the Spitzer Galactic Plane Survey, Spitzer Proposal, [60020](#)
- Whitney, B. & GLIMPSE360 Team 2009, AAS Meeting Abstracts, **214**, 210.01
- Winston, E., Hora, J., Gutermuth, R., et al. 2019, [ApJ](#), **880**, 9
- Winston, E., Hora, J., & Tolls, V. 2020, "SFOG Cluster Fits Images," V1, Harvard Dataverse, doi:[10.7910/DVN/1TCGIM](#)
- Winston, E., Megeath, S. T., Wolk, S. J., et al. 2007, [ApJ](#), **669**, 493
- Winston, E., Megeath, S. T., Wolk, S. J., et al. 2009, [AJ](#), **137**, 4777
- Winston, E., Wolk, S. J., Gutermuth, R., et al. 2018, [AJ](#), **155**, 241
- Winston, E., Wolk, S. J., Megeath, S. T., et al. 2011, [ApJ](#), **741**, 166
- Wouterloot, J. G. A., Brand, J., Burton, W. B., et al. 1990, [A&A](#), **230**, 21
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, [AJ](#), **140**, 868
- Xu, Y., Reid, M. J., Zheng, X. W., et al. 2006, [Sci](#), **311**, 54