

## SURVEYING THE AGENTS OF GALAXY EVOLUTION IN THE TIDALLY STRIPPED, LOW METALLICITY SMALL MAGELLANIC CLOUD (SAGE-SMC). III. YOUNG STELLAR OBJECTS

M. SEWIŁO<sup>1,2</sup>, L. R. CARLSON<sup>3</sup>, J. P. SEALE<sup>4</sup>, R. INDEBETOUW<sup>5</sup>, M. MEIXNER<sup>4</sup>, B. A. WHITNEY<sup>2,6</sup>,  
T. P. ROBITAILLE<sup>7</sup>, J. M. OLIVEIRA<sup>8</sup>, K. GORDON<sup>4,9</sup>, M. R. MEADE<sup>6</sup>, B. L. BABLER<sup>6</sup>, J. L. HORA<sup>10</sup>,  
M. BLOCK<sup>11</sup>, K. MISSELT<sup>11</sup>, J. TH. VAN LOON<sup>8</sup>, C.-H. R. CHEN<sup>12</sup>, E. CHURCHWELL<sup>6</sup>, AND B. SHIAO<sup>4</sup>

<sup>1</sup> The Johns Hopkins University, Department of Physics and Astronomy, 366 Bloomberg Center,  
3400 N. Charles Street, Baltimore, MD 21218, USA; [mmsewilo@pha.jhu.edu](mailto:mmsewilo@pha.jhu.edu)

<sup>2</sup> Space Science Institute, 4750 Walnut St. Suite 205, Boulder, CO 80301, USA; [bwhitney@spacescience.org](mailto:bwhitney@spacescience.org)

<sup>3</sup> Sterrewacht Leiden, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands, [carlson@strw.leidenuniv.nl](mailto:carlson@strw.leidenuniv.nl)

<sup>4</sup> Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21218, USA;

[seale@stsci.edu](mailto:seale@stsci.edu), [meixner@stsci.edu](mailto:meixner@stsci.edu), [kgordon@stsci.edu](mailto:kgordon@stsci.edu), [shiao@stsci.edu](mailto:shiao@stsci.edu)

<sup>5</sup> Department of Astronomy, University of Virginia, P.O. Box 3818, Charlottesville, VA 22903, USA; [remy@virginia.edu](mailto:remy@virginia.edu)

<sup>6</sup> Department of Astronomy, 475 North Charter St., University of Wisconsin, Madison, WI 53706, USA;

[meade@sal.wisc.edu](mailto:meade@sal.wisc.edu), [brian@sal.wisc.edu](mailto:brian@sal.wisc.edu), [churchwell@astro.wisc.edu](mailto:churchwell@astro.wisc.edu)

<sup>7</sup> Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany; [robitaille@mpia-hd.mpg.de](mailto:robitaille@mpia-hd.mpg.de)

<sup>8</sup> School of Physical & Geographical Sciences, Lennard-Jones Laboratories, Keele University, Staffordshire ST5 5BG, UK;

[j.oliveira@keele.ac.uk](mailto:j.oliveira@keele.ac.uk), [j.t.van.loon@keele.ac.uk](mailto:j.t.van.loon@keele.ac.uk)

<sup>9</sup> Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium

<sup>10</sup> Center for Astrophysics, 60 Garden St., MS 67, Harvard University, Cambridge, MA 02138, USA; [jhora@cfa.harvard.edu](mailto:jhora@cfa.harvard.edu)

<sup>11</sup> Steward Observatory, University of Arizona, 933 North Cherry Ave., Tucson, AZ 85719, USA; [kmissett@as.arizona.edu](mailto:kmissett@as.arizona.edu)

<sup>12</sup> Max Planck Institute for Radio Astronomy, Auf dem Hügel 69, D-53121 Bonn, Germany; [rchen@mpifr-bonn.mpg.de](mailto:rchen@mpifr-bonn.mpg.de)

Received 2012 December 21; accepted 2013 September 11; published 2013 October 29

### ABSTRACT

The *Spitzer Space Telescope* Legacy Program SAGE-SMC allows global studies of resolved stellar populations in the SMC in a different environment than our Galaxy. Using the SAGE-SMC IRAC (3.6–8.0  $\mu\text{m}$ ) and MIPS (24 and 70  $\mu\text{m}$ ) catalogs and images combined with near-infrared ( $JHK_s$ ) and optical ( $UBVI$ ) data, we identified a population of  $\sim 1000$  intermediate- to high-mass young stellar objects (YSOs) in the SMC (three times more than previously known). Our method of identifying YSO candidates builds on the method developed for the Large Magellanic Cloud by Whitney et al. with improvements based on what we learned from our subsequent studies and techniques described in the literature. We perform (1) color–magnitude cuts based on five color–magnitude diagrams (CMDs), (2) visual inspection of multi-wavelength images, and (3) spectral energy distribution (SED) fitting with YSO models. For each YSO candidate, we use its photometry to calculate a measure of our confidence that the source is not a non-YSO contaminant, but rather a true YSO, based on the source’s location in the color–magnitude space with respect to non-YSOs. We use this CMD score and the SED fitting results to define two classes of sources: high-reliability YSO candidates and possible YSO candidates. We found that, due to polycyclic aromatic hydrocarbon emission, about half of our sources have [3.6]–[4.5] and [4.5]–[5.8] colors not predicted by previous YSO models. The YSO candidates are spatially correlated with gas tracers.

**Key words:** circumstellar matter – galaxies: dwarf – infrared: stars – Magellanic Clouds – stars: formation – stars: pre-main sequence

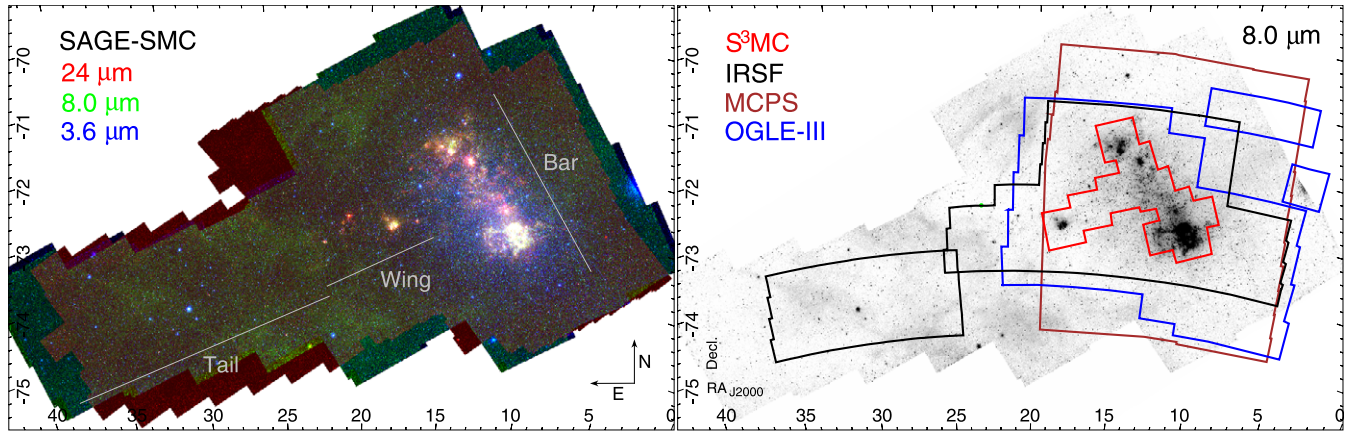
**Online-only material:** figure sets, machine-readable tables

### 1. INTRODUCTION

The Small Magellanic Cloud (SMC) is one of the closest galaxy companions to the Milky Way. Located at a distance of  $\sim 60$  kpc (e.g., Hilditch et al. 2005; Haschke et al. 2012), this gas-rich irregular dwarf galaxy is characterized by metallicity much lower than that in our Galaxy ( $\sim 0.1$ – $0.2 Z_{\odot}$ ; Russell & Dopita 1992; Rolleston et al. 1999; Rolleston et al. 2003; Lee et al. 2005). The formation of stars in the high-redshift universe occurred in a metal-poor environment; thus, the SMC can be used as a template for detailed studies of star formation under metal-poor conditions. The proximity of the SMC makes it ideal to study star formation on galactic scales and within individual star formation regions without the confusion and extinction of the Galactic plane. We can study resolved (proto)stellar populations throughout the entire SMC and how they relate to the galaxy’s extensively studied structure and gas and dust distribution. Such global studies are very challenging in the Galaxy due to our location within it and distance determinations

within the Galaxy are highly uncertain. The distance to the SMC is 10 times its line-of-sight depth; therefore, treating all sources as though they are at a single known distance is a reasonable assumption.

There are two main components of the SMC, namely the bar and the wing (see Figure 1). The gas-rich bar of the SMC extending from the northeast to southwest over roughly  $3^{\circ}$  is the place of the most vigorous star formation. It does, however, contain old stellar populations as opposed to the wing which mainly hosts young stars. The wing originates from the bar and extends  $\sim 2^{\circ}$  toward the southeast. Extended farther to the east is the tail of the SMC. The tail is the densest portion of the Magellanic Bridge (the H I filament connecting SMC with the Large Magellanic Cloud, LMC; e.g., McGee & Newton 1986; Muller et al. 2003). The formation of the Bridge was most likely a result of the tidal interaction between the SMC and LMC. Both young, embedded star formation (Gordon et al. 2009 and references therein; C.-H. R. Chen, in preparation) and old (Bagheri et al. 2013) stellar populations have been detected



**Figure 1.** Left: the three-color composite image of the SMC combining the MIPS  $24\ \mu\text{m}$  (red), IRAC  $8.0\ \mu\text{m}$  (green), and IRAC  $3.6\ \mu\text{m}$  (blue) images from the SAGE-SMC survey. The three main components of the SMC are indicated: the bar, wing, and tail (high-density portion of the Magellanic Bridge). Right: the comparison between the coverage of the SAGE-SMC survey represented by the  $8.0\ \mu\text{m}$  image, and the coverage of other surveys (color-coded as indicated in the legend) whose data are used in this paper: S<sup>3</sup>MC (Bolatto et al. 2007), IRSF (Kato et al. 2007), MCPS (Zaritsky et al. 2002), and OGLE-III (Udalski et al. 2008). The SAGE-SMC survey area is fully covered by the 2MASS survey (Skrutskie et al. 2006) and by 6X2MASS survey for R.A.  $\lesssim 28^{\circ}2$  (Cutri & 2MASS Team 2004). The data from the 2MASS and 6X2MASS surveys are the integral part of the SAGE-SMC IRAC catalogs. See Section 2.2 and Appendix B for details on all the ancillary surveys.

in the Bridge, indicating that both the gas and stellar content of the SMC were pulled out during its formation.

The *Spitzer Space Telescope* (*Spitzer*; Werner et al. 2004) was the first instrument that enabled studies of the resolved (proto)stellar populations in the SMC with the unprecedented sensitivity at the range of infrared (IR) wavelengths. The sensitivity and resolution of *Spitzer*, combined with the proximity of the SMC, allowed observations of the protostars and small clusters at sub-parsec resolution ( $\sim 2''$  or  $\sim 0.6$  pc) at 3.6, 4.5, 5.8, and  $8.0\ \mu\text{m}$  with the Infrared Array Camera (IRAC; Fazio et al. 2004) and at resolutions of  $6''$  ( $\sim 1.7$  pc),  $18''$  ( $\sim 5.2$  pc), and  $40''$  ( $\sim 11.6$  pc) at 24, 70, and  $160\ \mu\text{m}$ , respectively, with the Multiband Imaging Photometer for *Spitzer* (MIPS; Rieke et al. 2004).

Before the advent of *Spitzer*, only one protostar had been identified in the SMC. The protostar, located in the N 76B H II region, was identified by Gatley et al. (1982) based on its near-IR colors, correlation with ionized gas, and spectral characteristics. Identification of a significant number of embedded young stellar object (YSO) candidates in the SMC was enabled for the first time by the The *Spitzer* Survey of the Small Magellanic Cloud (S<sup>3</sup>MC; Bolatto et al. 2007). S<sup>3</sup>MC imaged a  $\sim 2^{\circ}5 \times 1^{\circ}$  region along the bar and a  $\sim 1^{\circ}5 \times 1^{\circ}$  region along the wing in all the IRAC and MIPS bands. Based on the S<sup>3</sup>MC photometric catalog containing  $\sim 400,000$  mid- and far-IR point sources, Bolatto et al. (2007) compiled a list of 282 YSO candidates. The selection criteria were based on the sources'  $5.8\ \mu\text{m}$  magnitudes and  $[5.8]-[8.0]$  colors, chosen to minimize contamination by background galaxies. Subsequent, more detailed studies on YSOs in the SMC concentrated on individual star-forming regions: NGC 346 (Simon et al. 2007) and NGC 602 (Carlson et al. 2007, 2011; Gouliermis et al. 2007), identifying  $\sim 150$  additional YSO candidates (see Section 3.1).

The *Spitzer* Legacy Project Surveying the Agents of Galaxy Evolution in the Tidally Stripped, Low Metallicity Small Magellanic Cloud (SAGE-SMC; Gordon et al. 2011b) imaged a 30 square degree region of the SMC in the IRAC (3.6, 4.5, 5.8, and  $8.0\ \mu\text{m}$ ) and MIPS (24, 70, and  $160\ \mu\text{m}$ ) bands. The SAGE-SMC survey covered the full SMC including the bar, wing, and the Magellanic Bridge (tail). SAGE-SMC allows us to construct a census of YSOs in the whole SMC. The goal of this work

is to construct a list of high-reliability, bright YSO candidates and characterize them as a population to learn about the environments in which intermediate and high-mass stars form (e.g., correlation with gas tracers, clustering, triggering).

Our approach combines the successful methods used by Whitney et al. (2008) and Gruendl & Chu (2009) to identify YSOs in the LMC based on the data from the galaxy-wide *Spitzer* Legacy Project Surveying the Agents of Galaxy Evolution (SAGE; Meixner et al. 2006). We apply color-magnitude selection criteria based on multiple color-magnitude diagrams (CMDs) and perform spectral energy distribution (SED) fitting with YSO models following the method developed by Whitney et al. (2008). Gruendl & Chu (2009) use a color-magnitude cut based on a single CMD and perform a visual inspection of images and SED shapes. While Whitney et al. (2008) inspected all the IRAC and MIPS  $24\ \mu\text{m}$  images, Gruendl & Chu (2009) showed that a careful examination of the images over a broader wavelength range (from optical to mid-IR, as well as CO and H $\alpha$ ) provides valuable information about the environment of the sources, making the classification process more informed. Also, the Gruendl & Chu (2009) analysis was based on aperture photometry, allowing them to identify more massive YSO candidates (slightly extended in IRAC bands) that were missed by Whitney et al. (2008), who used point source catalogs. Our analysis is based primarily on the Whitney et al. (2008) method; however, we improve it by incorporating the inspection of all available images and supplementing point source catalog data with aperture photometry and streamline it by using 5 CMDs rather than 13.

Throughout this paper, we use the YSO classification scheme introduced by Robitaille et al. (2006). They defined stages of YSO evolution which depend on the source's physical parameters rather than the observational characteristics of its SED as in the more traditional class scheme defined for low-mass YSOs (Lada 1987; André et al. 1993). These two classification schemes are roughly equivalent in terms of main characteristics of each phase of the YSO evolution; however, stages are defined by the circumstellar dust distribution, and thus can also be applied (with some limitations; Robitaille 2008) to high-mass objects. Stage 0 protostars are in the earliest stage of evolution. The protostar and possibly a disk are deeply

embedded in the massive accreting envelope of gas and dust and cannot be observed directly. The cold envelopes can only be detected at far-IR and sub-millimeter wavelengths. In this main accretion phase, the envelope mass is significantly larger than the mass of the central protostellar object and collimated bipolar molecular outflows are present. In the more evolved Stage I, the envelope’s density decreases as the accretion from the envelope onto the disk continues. The forming star that has already accreted much of its mass drives bipolar outflows. Stage I YSOs emit the bulk of their radiation at the mid-IR to far-IR wavelengths. Stage II is characterized by optically thick disk, bipolar outflow, and possibly the remains of a tenuous infalling envelope. Stage II YSOs emit mostly in the near-IR to mid-IR regime. Stage III YSOs have optically thin disks or no disks, and may still have optical bipolar jets. These objects are mostly detected at optical and near-IR wavelengths. The *Spitzer* mid-IR data reveal primarily Stage I and Stage II YSOs.

Photometric YSO selection and characterization is a complex procedure because it is not possible to completely and cleanly separate YSOs from all other kinds of sources in multidimensional color–magnitude space. We use several complementary metrics to select YSO candidates and determine the likelihood that a given source is a YSO: binary cuts in color–magnitude space, fitting SEDs with YSO models, manual inspection of images and fit SEDs, and a CMD score that quantifies the degree to which non-YSOs overlap with the source’s location in color–magnitude space. Photometric selection requires as complete and reliable a photometric catalog as possible, described in Section 2. Both the binary CMD cuts and continuous CMD score require prior knowledge of the location of YSOs and non-YSOs in color–magnitude space, described in Section 3. In Section 4, we describe the construction of the initial list of YSO candidates, and in Sections 5 and 6, we describe how we refine it using the various tools. The results are presented in Section 7. We provide a summary of this work and final remarks in Section 8.

## 2. PHOTOMETRIC DATA

This work is based primarily on *Spitzer* SAGE-SMC data products, including IRAC and MIPS photometric lists and images. A detailed description of the SAGE-SMC data products can be found in the documentation for the SAGE-SMC survey available through the *Spitzer* Science Center (Gordon et al. 2011a) and in Gordon et al. (2011b). We complemented these with ancillary point-source photometry from previously conducted optical and near-IR surveys. We also performed new aperture photometry to produce a more complete data set.

### 2.1. Primary Photometry: SAGE-SMC

The basis on which we built our study are photometry lists produced by the SAGE-SMC team from IRAC and MIPS 24  $\mu\text{m}$  images. Sources fall into two categories: IRAC catalog sources (with or without MIPS 24  $\mu\text{m}$  catalog matches) and MIPS 24  $\mu\text{m}$  sources with no IRAC catalog counterparts. Sources from the IRAC catalog fulfill a set of stringent criteria during point-spread function (PSF) photometry, and thus we consider them to be highly reliable. However, this stringency also leads to marginally extended sources and sources in crowded areas being neglected. We therefore also considered a second category, MIPS 24  $\mu\text{m}$  sources with no matches in the IRAC catalog within 1". This category is less reliable because the MIPS 24  $\mu\text{m}$

catalogs we use contain all the sources extracted from images; we expect contamination from spurious sources. MIPS 24  $\mu\text{m}$  observations were carried out in two epochs; many of the sources extracted are only measured in one of the two and are likely to be false detections.

The IRAC Single Frame + Mosaic Photometry (SMP) Archive contains  $\sim 2.19$  million sources with faint limits of 18.5, 18.1, 16.2, and 15.4 mag, for 3.6, 4.5, 5.8, and 8.0  $\mu\text{m}$ , respectively. The faint limit is defined as the threshold magnitude above which 99% of sources in a given band are detected. The IRAC SMP Archive includes sources that fulfill a set of criteria developed to ensure that each source is a legitimate astronomical source and that the fluxes reported for the IRAC bands are of high quality (Gordon et al. 2011a). We use the IRAC SMP Archive rather than the even more reliable IRAC SMP Catalog for completeness in both the number of sources and flux measurements at each wavelength. We provide a short summary on the construction of the IRAC SMP Archive in Appendix A. We refer to the IRAC SMP Archive as the IRAC catalog for the remainder of the paper.

The MIPS 24  $\mu\text{m}$  Epoch 1 and Epoch 2 full lists contain  $\sim 66,000$  and  $\sim 69,000$  sources, respectively; the faint limit is  $\sim 12$  mag for both epochs (Gordon et al. 2011a, 2011b). These source lists consist of all the sources extracted from the MIPS 24  $\mu\text{m}$  mosaics and thus are expected to include spurious sources. The more reliable SAGE-SMC MIPS 24  $\mu\text{m}$  catalogs consist of sources that fulfill very stringent criteria; however, as a consequence, the number of sources is greatly reduced in comparison to the full lists. We use the full lists in this paper for completeness. Spurious sources can be identified via data inspection (see Section 4.2). We refer to this 24  $\mu\text{m}$  full list as the MIPS 24  $\mu\text{m}$  catalog.

We matched the SAGE-SMC IRAC and MIPS 24  $\mu\text{m}$  catalogs in the SAGE-SMC database. We performed the matching between the IRAC sources and two epochs of MIPS separately due to the nature of the MIPS catalog and the positional uncertainty of the IRAC and MIPS sources.

At the faint end, a MIPS 24  $\mu\text{m}$  source may be present in one epoch but not in the other. There are also rare cases of incorrect source extractions where one source is split into a few sources; it may occur for Epoch 1 source extraction, but not Epoch 2 or vice versa. The positional uncertainty for IRAC and MIPS sources is 0".3 and  $<0".1$ , respectively, and thus it is possible for the distance between an IRAC source and a MIPS 24  $\mu\text{m}$  source in one epoch to be slightly larger than 1", but slightly smaller than 1" in the other. As a consequence, our selection criteria would miss a 24  $\mu\text{m}$  match to the IRAC source if only one epoch of the MIPS 24  $\mu\text{m}$  catalog were used. To decrease the number of unreliable or spurious sources that contaminate MIPS 24  $\mu\text{m}$  catalogs, we only included 24  $\mu\text{m}$  sources with signal-to-noise ratio  $>5$ . Out of  $\sim 2.19$  million IRAC sources, 41,179 have MIPS 24  $\mu\text{m}$  matches within a conservative match radius of 1" in one or both MIPS epochs with 27,255 matches from Epoch 1 and 29,500 from Epoch 2.

We used MIPS 24  $\mu\text{m}$  Epoch 1 and Epoch 2 catalogs to create a list of MIPS 24  $\mu\text{m}$  sources with no IRAC counterparts in the IRAC catalog within 1". The SAGE-SMC IRAC point source catalog is the foundation of our initial source selection. Some of the SMC’s YSOs, however, may not be included in the IRAC catalog. Massive YSOs illuminate larger volumes than lower-mass YSOs, thus they are often marginally extended in IRAC bands. Moreover, the stringent definition of a point source adopted by the SAGE-SMC IRAC pipeline team using PSF

photometry excluded some of the sources from the final source lists. Chen et al. (2009) and Gruendl & Chu (2009) discuss this effect for the LMC SAGE survey that used the same data processing as SAGE-SMC. To make our list of YSO candidates more complete, we searched the MIPS  $24\ \mu\text{m}$  catalogs for sources not included in the IRAC catalog. We expect most of the YSOs extended in IRAC bands to be point-like at MIPS  $24\ \mu\text{m}$  band since the MIPS resolution at  $24\ \mu\text{m}$  is three times poorer than IRAC resolution ( $6''$  versus  $\sim 2''$ ;  $\sim 1.7\ \text{pc}$  versus  $\sim 0.6\ \text{pc}$  at 60 kpc). This selection added 68,631 potential YSOs to our list, and we must have supplemented their  $24\ \mu\text{m}$  measurements with photometry at shorter wavelengths (see Section 2.3).

## 2.2. Ancillary Data: Point Source Catalogs

To provide photometry for the SAGE-SMC IRAC catalog sources (already matched to MIPS  $24\ \mu\text{m}$  catalog) over a broader wavelength range (from optical to  $24\ \mu\text{m}$ ), the IRAC and MIPS catalogs were cross-matched to optical and near-IR point source catalogs. These are from the Magellanic Clouds Photometric Survey (MCPS; Zaritsky et al. 2002) and OGLE-III (Udalski et al. 2008) in the optical ( $U$ ,  $B$ ,  $V$ , and  $I$  and  $V$  and  $I$  bands, respectively). For the near-IR  $J$ ,  $H$ , and  $K_s$ , we used the catalog from the Infrared Survey Facility (IRSF) Magellanic Clouds Point Source Survey (Kato et al. 2007), the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006), and the 2MASS 6X Deep Point Source Catalog (6X2MASS; Cutri & 2MASS Team 2004). We also used IRSF and 2MASS images for some aperture photometry. The IRSF survey has higher resolution and is significantly deeper than 2MASS, but it covers a smaller area (see Figure 1). Details of these catalogs are given in Appendix B.

We considered sources from any two catalogs (IRAC-MIPS24, IRAC-MCPS, MIPS24-MCPS, IRAC-IRSF, MIPS24-IRSF, IRAC-OGLEIII, MIPS24-OGLEIII) a match when a distance between their positions is  $\leq 1''$ . If a source from one catalog has two matching sources within  $1''$  in another catalog, we only considered the closest match.

## 2.3. Aperture Photometry

For the  $\sim 69,000$   $24\ \mu\text{m}$  sources with no IRAC catalog match, no fluxes at wavelengths shorter than  $24\ \mu\text{m}$  are available in the SAGE-SMC catalog. We performed aperture photometry at their positions to obtain fluxes for each source over a wavelength range from  $1.235\ \mu\text{m}$  (IRSF/2MASS  $J$  band) to  $24\ \mu\text{m}$ , whenever a reliable flux measurement is possible. Wavelength coverage depends on the availability of images. In the near-IR, we use  $JHK_s$  images from the IRSF survey where available and 2MASS images outside the IRSF survey coverage. In the mid-IR, we performed aperture photometry on SAGE-SMC images in IRAC 3.6, 4.5, 5.8, and  $8.0\ \mu\text{m}$  and MIPS  $24\ \mu\text{m}$ .

For IRAC, we used both short (0.6 s) and long exposure (12 s) images (Gordon et al. 2011a) to obtain good flux measurements for the brightest and faintest sources. Faint data eventually suffer Malmquist effects in the short exposures (sources appear to be brighter than they really are), while long exposures saturate consistently at some bright points. In the SAGE-SMC catalog, fluxes were adopted from the short exposure catalog for magnitudes brighter than 12, 11, 9, and 9 for the IRAC 3.6, 4.5, 5.8, and  $8.0\ \mu\text{m}$  bands, respectively, and long exposure fluxes were used for magnitudes fainter than the saturation limits of 9.5, 9.0, 6.5, and 6.5 in the corresponding bands. When both criteria

were met, the uncertainty-weighted average of the short and long exposure fluxes was used. The limits were chosen to maximize the overlap of these two conditions without compromising the results. We compared the short and long exposure aperture photometry fluxes and determined that the limiting values for the 3.6, 4.5, and  $8.0\ \mu\text{m}$  bands used by the SAGE-SMC IRAC pipeline (Gordon et al. 2011a) are appropriate for our analysis. For the  $5.8\ \mu\text{m}$  band, we adopt limiting magnitudes of 9.4 at the faint end and 8.2 at the bright end.

We used aperture sizes of  $5''$  for  $JHK_s$  and all IRAC bands, and  $8''$  for the MIPS  $24\ \mu\text{m}$ . Aperture corrections were 1.090, 1.103, 1.117, 1.129, and 1.56 for 3.6, 4.5, 5.8,  $8.0$ , and  $24\ \mu\text{m}$  bands, respectively. These corrections were determined by performing aperture photometry on the subsampled PSFs available from the *Spitzer* Science Center with the exact same aperture and background annulus (since the background subtracts a bit of the outer PSF wing) as we use for real photometry. Background emission was calculated as the 2 sigma-clipped mean of an annulus ranging from 1.75 to 2 times the source radius. Aperture photometry used conservative uncertainties from background annulus. Uncertainties were increased where large gradients existed across the background annulus and in regions of crowding and confusion. Because uncertainties were already conservative, we required a relatively low signal-to-noise cutoff of 2 for aperture photometry fluxes.

We also performed aperture photometry for selected sources from the IRAC catalog (see Section 4.1 on the initial source selection) to provide as complete photometry as possible for all the sources. For IRAC catalog sources, some point source flux measurements may be missing from the IRAC catalog in one or more bands. These could be sources that are marginally resolved in some IRAC bands, those that have low-quality point source flux measurements (and thus did not fulfill the criteria for the inclusion in the IRAC catalog; Gordon et al. 2011a), or other glitches in the automated source extraction. Aperture photometry fluxes were used to fill in the bands from  $J$  to MIPS  $24\ \mu\text{m}$  where SAGE-SMC catalog fluxes are unavailable.

## 3. LITERATURE RESOURCES AND ANALYTIC TOOLS USED FOR YSO VERIFICATION

In this section, we describe the main resources from the literature we incorporated to identify YSO candidates and verify their nature. We gathered information on previously identified YSOs and YSO candidates (Section 3.1), as well as catalogs of other populations of sources (Section 3.2). We use these catalogs and YSO models to understand the behavior of YSOs and non-YSOs in color-magnitude space, which guide our selection of YSO candidates. We assess the effectiveness of our methods and contamination of our YSO candidate list by comparison with previously known YSO and non-YSO lists. We use SED fitting (Section 3.3) to confirm the nature of YSO candidates, to see how well it works for the YSO selection, and to constrain physical properties of the sources. We use CO, H $\alpha$ , and H $\alpha$  images (Section 3.4) to study the association of YSO candidates with the gas tracers.

### 3.1. Previously Known YSOs and YSO Candidates

There are 33 SMC sources confirmed spectroscopically as bona fide YSOs (van Loon et al. 2008, 2010a; Oliveira et al. 2011, 2013). Oliveira et al. (2013) studied these sources from optical to far-IR wavelengths, including optical spectra, 3–5  $\mu\text{m}$  spectra and *Spitzer* spectra with the *Spitzer* InfraRed

Spectrograph (IRS) and the SED mode of MIPS. They detected H<sub>2</sub>O and/or CO<sub>2</sub> ice absorption features—a definite signature of the YSO at an early evolutionary stage—toward 14 objects. A further two sources exhibit silicates in absorption, another indicator of youth. Many sources also exhibit polycyclic aromatic hydrocarbon (PAH) and fine-structure emission typical of (young) compact H II regions.

Large-scale, high-resolution and high-sensitivity photometric studies of YSOs in the mid- and far-IR in the SMC were enabled for the first time by the *Spitzer* S<sup>3</sup>MC survey (see Section 1). Bolatto et al. (2007) identified 282 YSO candidates using S<sup>3</sup>MC in the survey area that covered the bar and part of the wing of the SMC (see Figure 1). These regions include sites of the most active star formation. The Bolatto et al. (2007) YSO candidate selection was based on the 5.8 μm absolute magnitudes ( $-10 < M_{5.8} < -6$ ) and [5.8]–[8.0] colors ( $[5.8] - [8.0] > 1.2$ ), and required high signal-to-noise ( $> 10$ ) detections at 5.8 μm, 8.0 μm, and one neighboring band. As a consequence, their list of YSO candidates includes bright and highly reddened sources; in general, their classification as YSOs has not been verified by other means. A detailed analysis of *Spitzer* sources in individual star formation regions allows more certain classification of sources as they can be investigated individually over a wide wavelength range. Based on the data from optical to mid-IR, two star formation regions in the SMC were studied extensively: NGC 346 (Simon et al. 2007) using S<sup>3</sup>MC data and NGC 602 (Carlson et al. 2007, 2011; Gouliermis et al. 2007), which is outside of the S<sup>3</sup>MC survey area but was part of a pilot study with IRAC (Program ID 125, P.I. Fazio).

Simon et al. (2007) conducted the first detailed study of an individual H II region in the SMC aimed to examine the region’s young stellar content. Based on the S<sup>3</sup>MC data, they identified 61 definite YSOs and 50 possible YSOs in the brightest SMC H II region N 66 (containing the young cluster NGC 346) using SED fitting with YSO models (Robitaille et al. 2006) and a set of 4 selection criteria based on the [3.6]–[4.5] and [4.5]–[8.0] colors. The SEDs of definite YSOs can only be fit by YSO models, while the possible YSOs are objects with no unique SED classification. Simon et al. (2007) demonstrated that it is not possible to obtain a complete and uncontaminated sample of YSO candidates based solely on a set of color diagnostics as in the work of Bolatto et al. (2007). Color selection combined with SED fitting is a more reliable method of the YSO identification, as also shown by Whitney et al. (2008) in their galaxy-wide LMC survey.

Carlson et al. (2011) further improved the method of YSO identification by complementing the *Spitzer* data with high resolution optical data obtained with the Advanced Camera for Surveys (ACS) on board the *Hubble Space Telescope* (*HST*). Carlson et al. (2011) identified 41 embedded YSO candidates in the star-forming cluster NGC 602 in the wing of the SMC. The sources were identified by *Spitzer*/IRAC colors and SED fitting with a grid of YSO models (Robitaille et al. 2006), but only high resolution ACS optical images allowed the morphological identification of background galaxies, which are often identical in color to low-mass YSOs. The higher resolution *HST* images also show that at least 70% of the *Spitzer* YSO candidates are multiples or protoclusters. The SED fitting, combining multi-wavelength data from the *HST* 0.55 μm to *Spitzer*/MIPS 24 μm, allowed the characterization of YSOs, with results including mass, accretion rate, luminosity, envelope mass, and more; from these, their evolutionary stages were estimated. Based on the same IRAC data set, Gouliermis et al. (2007) independently

identified 22 candidate YSOs in NGC 602; 19 of these sources are in the Carlson et al. (2011) sample.

Studies based primarily on the *Spitzer* photometric data identified a total of 456 YSO candidates in the SMC (433 individual sources; see Section 6.1.2).

### 3.2. Catalogs of Other Types of Sources

Several types of sources can be confused with YSOs in color–magnitude space. We use catalogs of some of these sources to identify regions of color–magnitude space where there is the least confusion between YSOs and other types of objects. This information will allow us to select the initial YSO candidates list that is relatively free of contaminants. We will also use these non-YSO catalogs to assess the residual contamination from non-YSOs on our final list of YSO candidates. We consider a catalog of evolved star (asymptotic giant branch or AGB stars) candidates identified in the SMC by Boyer et al. (2011), which are similar in color to some bright YSO models. We also compare the Bonanos et al. (2010) catalog of massive stars, which was made based primarily on the SAGE-SMC data. A list of planetary nebulae (PNe) are considered (G. Jacoby 2009, private communication) by matching to the SAGE-SMC catalog. Our list of H $\alpha$  emission line stars and small nebulae comes from Meyssonnier & Azzopardi (1993); dusty OB stars are taken from Sheets et al. (2013). Finally, we prepare a comparison list of expected background galaxy photometry using IRAC and MIPS 24 μm catalogs from The *Spitzer* Wide-area InfraRed Extragalactic Survey (SWIRE; Lonsdale et al. 2003). Note that all of these lists except massive stars and dusty OB stars were constructed photometrically. See Appendix C for further details.

### 3.3. YSO Models and the SED Fitter

Following the method of YSO identification developed by Whitney et al. (2008), we use the distribution of Robitaille et al. (2006) model YSOs in color–magnitude space to guide our selection of YSO candidates. Robitaille et al. (2006) computed a set of 20,000 radiation transfer models of axisymmetric YSOs. Each YSO model predicts an emergent SED at 10 viewing angles, resulting in a total of 200,000 model SEDs. In Section 5.2, we describe fitting our YSO candidates’ photometry to these model SEDs to derive physical parameters. The YSO models cover stellar masses from 0.1 to 50  $M_{\odot}$ . Evolutionary stages range from the very early stage of envelope infall to the late disk-only stage; the geometry consists of a pre-main sequence star, a disk, an infalling envelope and bipolar cavities (Whitney et al. 2003a, 2003b). Whitney et al. (2008) includes a detailed discussion on the distribution of the model YSOs in the CMDs and the limitations of the models when applied to the LMC/SMC.

One of the most important limitations is source confusion. The YSO models assume a single central source; however, at the distance to the SMC, the resolution of *Spitzer* corresponds to  $\sim 0.6$  pc (or  $\sim 120,000$  AU) at IRAC bands and  $\sim 1.7$  pc (or  $\sim 360,000$  AU) at MIPS 24 μm, thus multiple YSOs are likely to be within a single PSF, particularly in dense forming clusters. These multiple sources detected as a single *Spitzer* point source are likely to be (proto)cluster members, but in some cases they may be physically unrelated objects in chance superposition. The SEDs of a cluster will be dominated by its most massive member(s) due to the known steep mass–luminosity relation. Indeed, studies show that  $> 70\%$  of *Spitzer*

YSO candidates are protoclusters and in the majority of them a single object dominates the luminosity at all wavelengths (Chen et al. 2009; Carlson et al. 2011; Oliveira et al. 2013). Thus, treating the *Spitzer* YSOs as single massive objects is a reasonable approximation. Another caveat to note here is that the Robitaille et al. (2006) YSO model grid was developed for the Galaxy, and hence, it is based on the well-defined Galactic dust grain models and assumes a gas-to-dust ratio of 100. We account for the higher gas-to-dust ratio in the SMC ( $\sim 350$  in star formation regions; Leroy et al. 2007) by rescaling by 3.5 times the YSO model envelope infall rates, disk, and envelope masses. The physical parameters of YSO candidates are estimated by SED fitting with the Robitaille et al. (2006) YSO models using the linear regression SED fitting tool developed by Robitaille et al. (2007; see Section 5.2).

### 3.4. Ancillary Data: Images

We use H I, H $\alpha$ , and CO images to study the spatial correlation of the YSO candidates with the neutral, ionized, and molecular gas, respectively.

1. The H I image is a combination of the observations of Stanimirović et al. (1999) of the bar and the wing of the SMC and those of Muller et al. (2003) of the Magellanic Bridge. Both images were made by combining the high-resolution data from the Australia Telescope Compact Array (ATCA; half-power beam width  $\sim 98''$ ) and lower resolution data from the Parkes 64-m radio telescope (14'7).
2. The H $\alpha$  image was provided by F. Winkler, S. D. Points, R. C. Smith, the Magellanic Cloud Emission Line Survey (MCELS) Team, and NAO/AURA/NSF. The survey was conducted using the UM/CTIO Curtis Schmidt telescope at the Cerro Tololo Inter-American Observatory (CTIO) and imaged the central  $3^{\circ}5 \times 4^{\circ}5$  of the SMC with a spatial resolution of  $4''.6$ .
3. The  $^{12}\text{CO}$  ( $J = 1-0$ ) image comes from the NANTEN Submillimeter Observatory survey of the SMC (Y. Fukui 2012, private communication). The survey covered the northeast and southwest regions of the main bar and the H II regions N 84 and N 88 in the wing with a resolution of  $2''.6$  (Mizuno et al. 2001).

## 4. CONSTRUCTION OF THE YSO CANDIDATE LIST

The goal of this paper is to construct a reliable, although not necessarily complete, list of the YSO candidates across the SMC. Thus, the selection criteria described in the following sections are designed to select bright YSO candidates which are significantly less confused with non-YSO populations than fainter sources (background galaxies in particular) in color–magnitude space.

### 4.1. Initial Selection: Color–Magnitude Cuts

We follow the empirical method developed by Whitney et al. (2008) for the LMC (see Section 1) with some modifications to select our initial sample of YSO candidates. We reduce the number of CMDs used for determining color–magnitude criteria. The Whitney et al. (2008) selections are based on 13 CMDs; by using a large number of CMDs, their initial list of YSOs contained a large number of sources, ensuring that the final sample is as complete as possible within the limitations of the method. However, at the same time, they increased contamination from other populations of sources.

Although an effort was made to remove contaminating sources from the final list of the YSO candidates, Whitney et al. (2008) still expect a significant contamination, particularly from background galaxies.

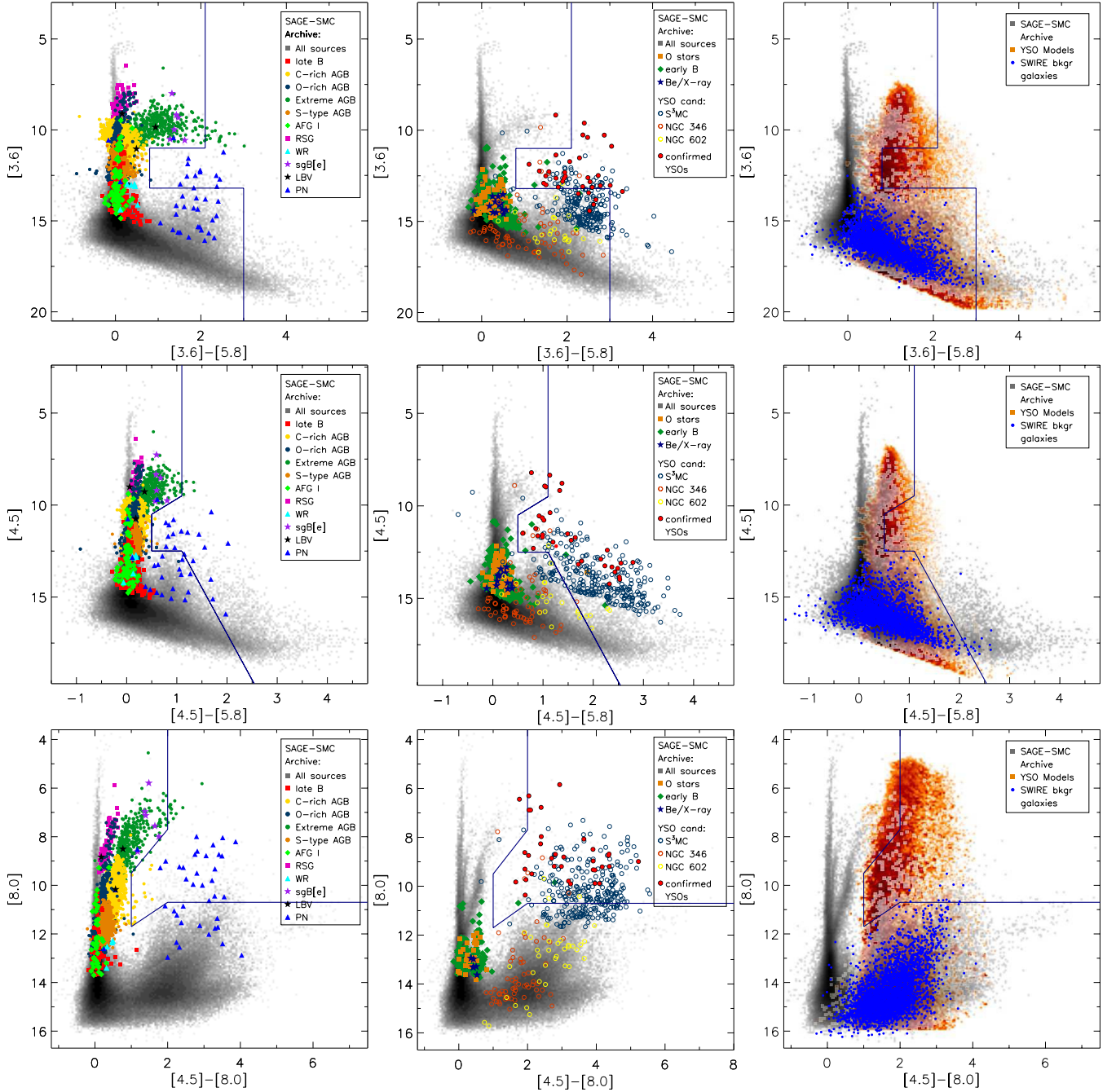
We tested a number of CMDs using different combinations of IRAC 3.6–8.0  $\mu\text{m}$ , and MIPS 24  $\mu\text{m}$  bands to select a set of CMDs for YSO analysis. Using multiple CMDs for source selection allows us to select more YSO candidates than using only one CMD. For example, if color–magnitude criteria are based only on the [8.0] versus [4.5]–[8.0] CMD, all sources without a flux measurement in either the 8.0 or 4.5  $\mu\text{m}$  band are excluded. To make sure we do not lose sources due to the lack of photometry in some bands (non-detections or measurements not included in the catalog due to low reliability/bad quality), we select CMDs that represent combinations of all IRAC and MIPS 24  $\mu\text{m}$  bands. Each CMD-cut may introduce more contaminants into our initial list (a non-YSO may lie close to the boundary in one CMD and just barely fulfill one color–magnitude criterion but not all the others); thus the selection criteria should be based on a small number of CMDs that show the greatest separation between YSOs and non-YSOs. Guided by our experience from previous work on YSOs, we select five CMDs sampling all five *Spitzer* bands, selecting more sources (than if only one CMD were used), while introducing only a moderate contamination from other populations of sources. The CMDs chosen are shown in Figure 2: [3.6] versus [3.6]–[5.8], [4.5] versus [4.5]–[5.8], [8.0] versus [4.5]–[8.0], [4.5] versus [4.5]–[24], and [8.0] versus [8.0]–[24].

The three plots in each row of Figure 2 show the same CMD, highlighting the locations of evolved stars in the left plot, young stars in the middle, and YSO models and background galaxies in the right panels. The catalogs and models are described in Section 3 and Appendix C. These populations of sources shown in different colors (as indicated in the legends) are overlaid on the SAGE-SMC IRAC and MIPS 24  $\mu\text{m}$  catalog sources shown in the gray Hess diagram. The plots in the right panels of Figure 2 show YSO two-dimensional radiation transfer models from Robitaille et al. (2006) displayed in the orange scale Hess diagram overlaid on the catalog sources (gray). Following the method developed by Whitney et al. (2008), we use the distribution of the model YSOs in color–magnitude space to guide our selection of YSO candidates.

Our selection criteria are listed in Equations (1)–(5) and are shown graphically in Figure 2 as navy lines. A source is included in the initial list of YSO candidates if it is located to the right (redward) of the navy line in any of the five selection CMDs. We avoid regions populated by background galaxies and those overlapping with evolved stars, selecting regions where intermediate to massive YSOs are expected to be located. Low-mass YSOs are confused with the background galaxies in the CMDs, and thus they are largely excluded by our color–magnitude criteria (see Figure 3). Our sample is also biased toward early evolutionary stages (see Figure 1 in Whitney et al. 2008); more evolved YSOs (mostly Stage III) overlap with evolved stars in color–magnitude space.

$$\left\{ \begin{array}{l} \text{or} \\ \text{or} \end{array} \right. \left\{ \begin{array}{l} [3.6] < 13.2 \quad \text{and} \quad ([3.6] - [5.8]) > 2.1 \\ 11 < [3.6] < 13.2 \quad \text{and} \quad 0.8 < ([3.6] - [5.8]) \leq 2.1 \\ [3.6] \geq 13.2 \quad \text{and} \quad ([3.6] - [5.8]) > 3.0 \end{array} \right. \quad (1)$$

$$\left\{ \begin{array}{l} \text{or} \\ \text{or} \end{array} \right. \left\{ \begin{array}{l} [4.5] \leq 12.5 \quad \text{and} \quad ([4.5] - [5.8]) > 1.1 \\ [4.5] < 12.5 \quad \text{and} \quad 0.5 < ([4.5] - [5.8]) \leq 1.1 \\ \text{and} \quad [4.5] > (11.33 - 1.67 \times ([4.5] - [5.8])) \\ \text{or} \quad [4.5] > 12.5 \quad \text{and} \quad [4.5] < (7.0 + 5.0 \times ([4.5] - [5.8])) \end{array} \right. \quad (2)$$



**Figure 2.** Distribution of evolved stars (left), young stars (middle), YSO models, and background galaxies (right) with respect to the entire population of sources from the SAGE-SMC IRAC SMP Archive (plotted in grey scale) shown for each CMD used for the source selection. Our YSO color–magnitude selection criteria are represented by the navy lines in each CMD. We selected sources to the right of these lines—regions that are expected to be occupied primarily by the YSOs. Our analysis was based on this initial sample. The CMDs show that we are biased toward intermediate- to high-mass stars and young evolutionary stages (see also Figure 1 in Whitney et al. 2008).

$$\left\{ \begin{array}{l} \text{or } \left[ \begin{array}{l} [8.0] < 10.7 \text{ and } ([4.5] - [8.0]) > 2.0 \\ 1.0 < ([4.5] - [8.0]) \leq 2.0 \text{ and } [8.0] < (12.7 - ([4.5] - [8.0])) \\ \text{and } [8.0] > (11.3 - 1.8 \times ([4.5] - [8.0])) \end{array} \right] \end{array} \right\} \quad (3)$$

$$\left\{ \begin{array}{l} \text{or } \left[ \begin{array}{l} [8.0] < 10 \text{ and } ([8.0] - [24]) > 2.7 \\ [8.0] < 10 \text{ and } ([8.0] - [24]) \leq 2.7 \\ \text{and } [8.0] > (12.5 - 1.67 \times ([8.0] - [24])) \end{array} \right] \\ \text{or } [8.0] \geq 10 \text{ and } [8.0] < (5.72 + 1.09 \times ([8.0] - [24])) \end{array} \right\} \quad (5)$$

$$\left\{ \begin{array}{l} [4.5] \leq 12.2 \text{ and } ([4.5] - [24]) > 4.0 \\ \text{or } [4.5] < 12.2 \text{ and } 3.2 < ([4.5] - [24]) \leq 4.0 \\ \text{and } [4.5] > (15 - 1.25 \times ([4.5] - [24])) \\ \text{or } [4.5] > 12.2 \text{ and } [4.5] < (6.43 + 1.03 \times ([4.5] - [24])) \end{array} \right\} \quad (4)$$

We apply these selections to the entire IRAC catalog. For those MIPS 24  $\mu\text{m}$  sources with no IRAC catalog match within 1", we use the aperture photometry described in Section 2.3. The resulting color-selected lists contain 4927 and 536 sources with IRAC catalog photometry and with aperture photometry, respectively. To ensure that sources from our sample are reliable,

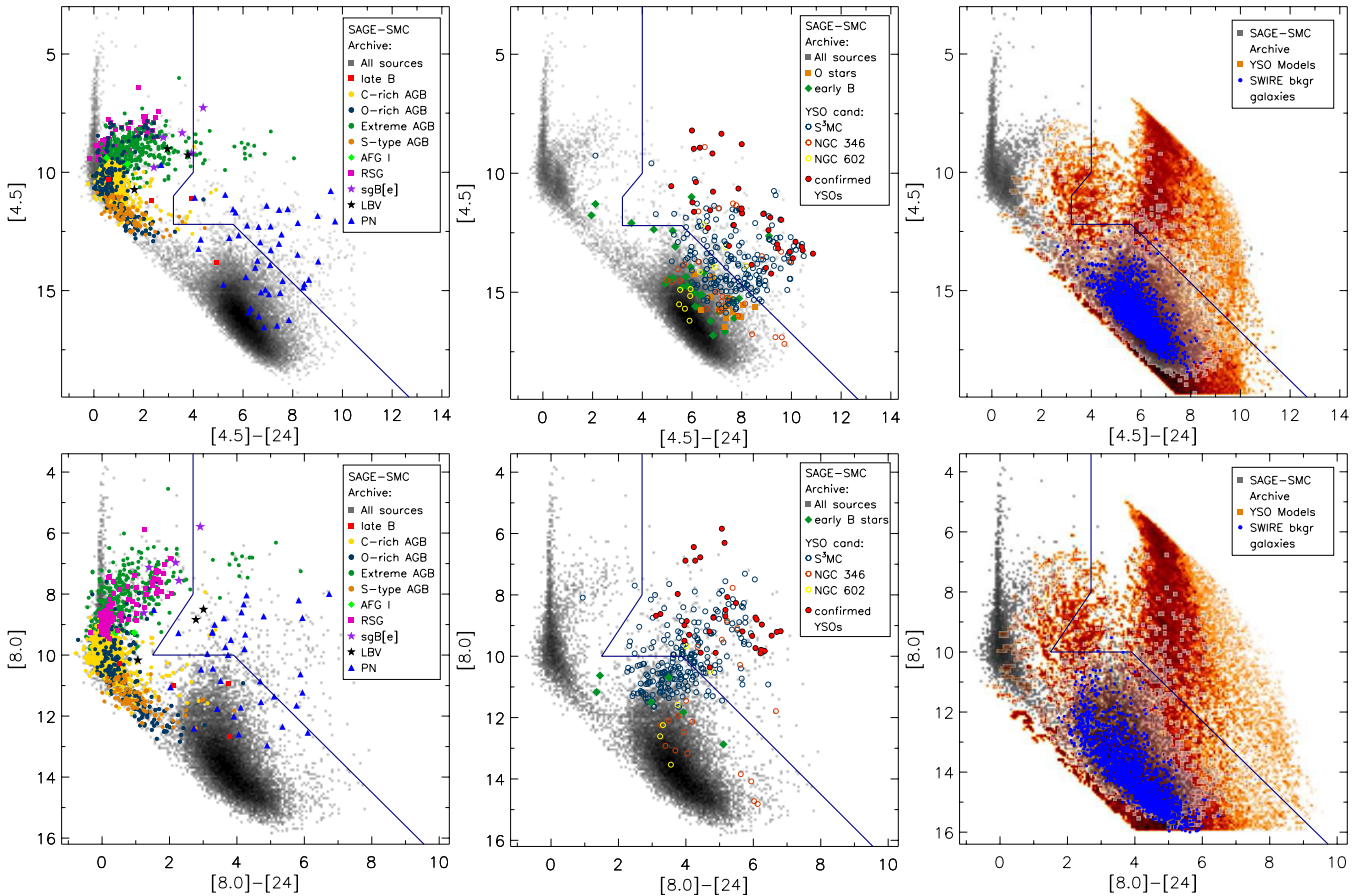


Figure 2. (Continued)

we require that each source has at least 3 *Spitzer* detections, i.e., valid flux measurements from the catalog or aperture photometry among all IRAC and MIPS 24  $\mu\text{m}$  bands. This requirement reduces the number of IRAC catalog sources to 3187 and MIPS 24  $\mu\text{m}$  sources without IRAC catalog matches to 388, giving a total of 3575 initial YSO candidates.

#### 4.2. Detailed Source Inspection

To further refine our YSO candidate list, we perform a detailed inspection of all color-selected sources. We examine all available images to assess morphology and environment, and SEDs with preliminary SED YSO model fitting results. We develop a computer program that automatically performs the aperture photometry and SED fitting with the YSO models (Robitaille et al. 2007) for thousands of sources. The SED fitting at this stage of our analysis is preliminary and serves as a guidance in the process of YSO identification; a thorough SED fitting is done for the sources we identify as YSO candidates following inspection (see Section 5.2).

For each source, the program generates an output that compiles image cutouts at several wavelengths, SEDs, the fitted YSO models, and physical parameters from the best-fit models. The image cutouts include IRSF *JHK<sub>s</sub>* (or 2MASS in regions not covered by IRSF), IRAC 3.6, 4.5, 5.8, and 8.0  $\mu\text{m}$ , MIPS 24 and 70  $\mu\text{m}$ ,  $\text{H}\alpha$ , and CO. Although we use MIPS 70  $\mu\text{m}$  image during the visual inspection, it is not used in the SED fitting due to its low resolution (FWHM of 18'') compared to the IRAC (2'') and MIPS 24  $\mu\text{m}$  data (6''). However, the presence or absence of the 70  $\mu\text{m}$  emission can support source identification.

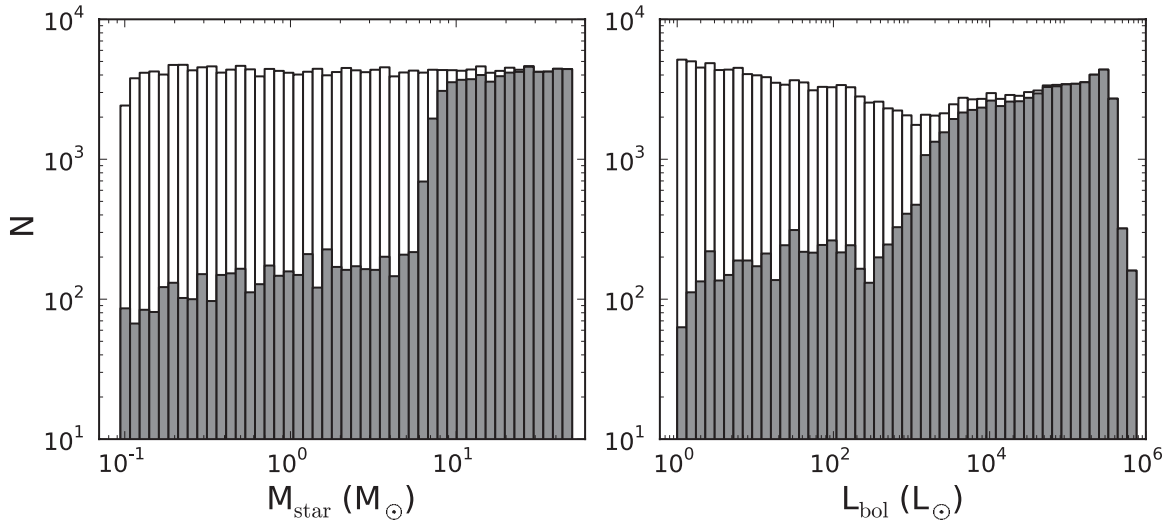
The images allow us to examine individual sources and their surrounding environment in detail at a wide wavelength range.

The photometric aperture and background annulus are marked in the cutout images, allowing for the evaluation of the aperture: adequate aperture size, multiplicity of the sources, etc. The plots showing the SEDs include the fluxes from the SAGE-SMC catalog (when available), providing an additional check of the aperture photometry. Additionally, for each source we prepare a set of the five CMDs used for color–magnitude cuts with the position of the source and the lines showing the selection criteria marked. We inspect these plots together with the image cutouts and the SED fits. We remove spurious sources—those that are not seen in the images (products of the automatic data extraction, mostly sources originating from the MIPS 24  $\mu\text{m}$  catalog with no IRAC catalog counterpart), objects with unreliable photometry, and those morphologically identified as background galaxies.

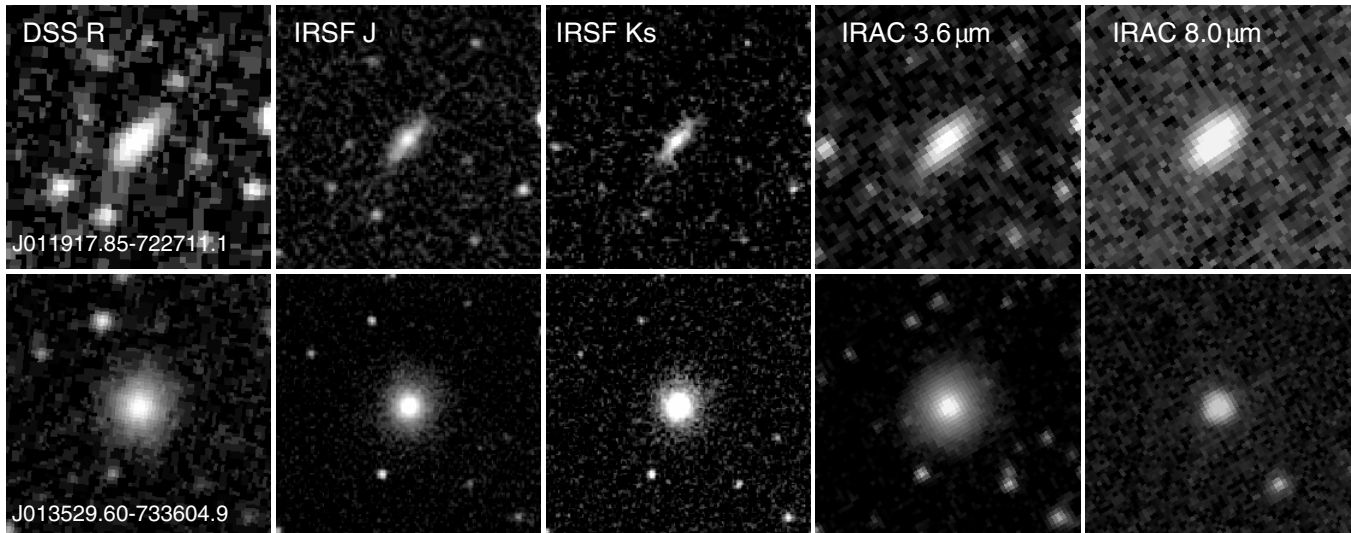
Figure 4 shows two examples of probable galaxies based on their morphological shapes. Source SSTM1SAGE2 J011917.85-722711.1 at the top of Figure 4 is elongated in all bands, suggesting that it is a roughly edge-on disk galaxy. Source SSTISAGEMA J013529.60-733604.9, on the other hand, resembles a face-on galaxy with a bright center surrounded by a smooth, regularly shaped extended emission. We gained experience in classifying mid-IR sources as galaxies from our work on NGC 602 (Carlson et al. 2011) where we could verify our classification based on high-resolution *HST* images.

Each source is inspected by two examiners independently. Sources for which there is disagreement are re-examined by a third person. We select a total of 1175 sources as YSO





**Figure 3.** Histograms of stellar mass (left) and luminosity (right) comparing the distribution of all YSO models (outline; Robitaille et al. 2006) to those models that fulfill the color–magnitude criteria (shaded). In general, our criteria select from intermediate- to high-mass stars; however, about 7.2% of models that fulfill the color–magnitude criteria have stellar mass less than  $5 M_{\odot}$ . This is the result of using logical ORs between different sets of criteria (Equations (1)–(5)), i.e., when the source does not fulfill the criteria in one or more CMDs, it may be located in the region occupied by lower luminosity YSOs in these CMDs.



**Figure 4.** DSS *R*, IRSF *J* and *K<sub>s</sub>*, and IRAC 3.6 and 8.0  $\mu\text{m}$  image cutouts for two sources suspected of being galaxies based on their morphology. Sources with similar morphologies were removed from the list of the YSO candidates during the source inspection. Top: MIPS 24  $\mu\text{m}$  Epoch 2-only source SSTM1SAGE2 J011917.85-722711.1; Bottom: IRAC source SSTISAGEMA J013529.60-733604.9, also in the MIPS 24  $\mu\text{m}$  catalog.

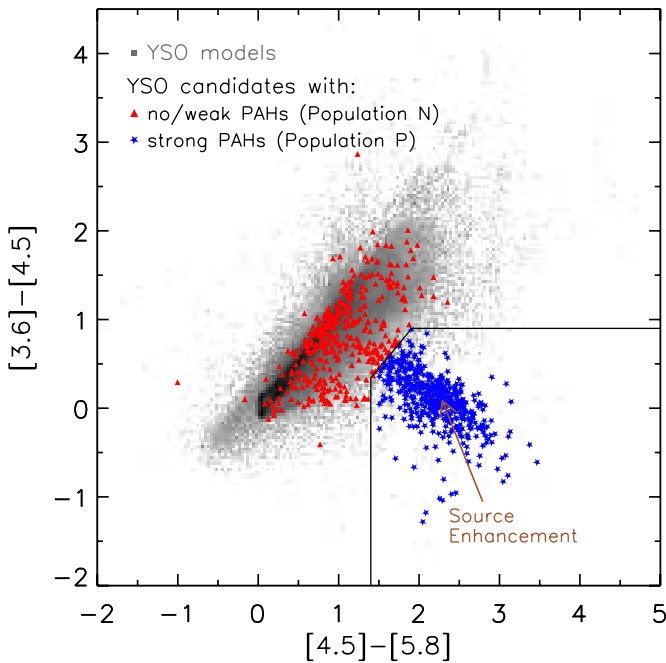
candidates. The spatial distribution of these sources reveal that a fraction of them lie at the outskirts of the SMC. Their positions indicate that they may be background galaxies rather than YSOs. We inspect  $\sim 85$  sources in Digitized Sky Survey (DSS) *R* images to search for any indication that they are galaxies; we identify and remove 51 sources whose morphological shapes across a range of wavelengths support their classification as galaxies and rather than YSOs. The resulting list contains 1124 YSO candidates, 1014 of which are IRAC catalog point sources.

Finally, we compare our full list to previously spectroscopically confirmed sources within  $1''$  of our candidates using SIMBAD and literature searches. We remove a total of 73 sources spectroscopically confirmed as non-YSOs. These include the following SIMBAD matches: 26 PNe, 7 carbon stars, 7 other evolved stars, 7 eclipsing binaries, 10 variables of other types, 1 active galactic nucleus (AGN), 3 quasars (QSOs), and 4 other galaxies. We also remove two sources that were found to be in globular clusters. A comparison to the catalog of 193 spectroscopically confirmed quasars behind the SMC (Kozłowski et al.

2011) return another 6 matches. This results in a list of 1051 potential YSOs. A detailed comparison between our finalized list to previous YSO studies is presented in Section 6.1.

#### 4.3. Two Populations of YSO Candidates: PAH Emission

Visual inspection of the sources' SEDs reveals a significant number of sources with SEDs ( $\lambda F_{\lambda}$  versus  $\lambda$ ) showing a characteristic dip at 4.5  $\mu\text{m}$ . This SED feature can be explained by the presence of PAHs toward YSO candidates. Diffuse emission in the 3.6, 5.8, and 8.0  $\mu\text{m}$  IRAC bands is dominated by the PAH features at 3.3, 6.2, 7.7, and 8.6  $\mu\text{m}$  when they are present in abundance, but the 4.5  $\mu\text{m}$  band remains relatively unaffected. In the presence of the strong PAH emission, the flux in the 3.6, 5.8, and 8.0  $\mu\text{m}$  bands increases, making the flux difference between these bands and the 4.5  $\mu\text{m}$  band larger. Consequently, the SEDs of sources with PAHs show a dip at 4.5  $\mu\text{m}$ . The depth of this feature depends on the strength of the PAH emission; it gets deeper with increasing PAH strength as



**Figure 5.**  $[3.6]-[4.5]$  vs.  $[4.5]-[5.8]$  CCD showing the distribution of the YSO candidates from our sample with respect to the Robitaille et al. (2006) YSO models. The solid line shows a boundary between Population N (red) and Population P (blue) sources:  $[4.5]-[5.8] = 1.4$  and  $[3.6]-[4.5] = 0.9$  and  $[3.6]-[4.5] = 1.1 \times ([4.5]-[5.8]) - 1.2$ . The position of the Source Enhancement is indicated with an arrow.

3.6, 5.8, and 8.0  $\mu\text{m}$  fluxes are amplified. To find out how many sources from our sample are affected by PAH emission, we use the  $[3.6]-[4.5]$  versus  $[4.5]-[5.8]$  color-color diagram (CCD), allowing us to assess the depth of the 4.5  $\mu\text{m}$  dip (and thus PAH strength) in the SEDs.

The  $[3.6]-[4.5]$  versus  $[4.5]-[5.8]$  CCD reveals two distinct populations of sources (see Figure 5). The first one stretching from the lower left to the upper right coincides well with the Robitaille et al. (2006) YSO models shown as a Hess diagram. We will refer to this population as Population N. The second population (called Population P) of the YSO candidates starts at  $([4.5]-[5.8], [3.6]-[4.5]) \sim (1.5, 0.5)$  and extends toward the lower right. Another feature in the CCD is a concentration of sources at about  $(2.3, 0)$ —labeled as Source Enhancement in Figure 5. A detailed analysis of Population P and the Source Enhancement is beyond the scope of this paper and will be discussed in a separate paper. We found that these two features consist of sources affected by PAHs with varying PAH strength.

The first indication that YSOs with PAHs occupy a distinct area in the color-color space was reported by Carlson et al. (2011). There we identified 41 YSO candidates in the star-forming cluster NGC 602 in the wing of the SMC. Of the 41 sources, 26 have 3.6, 4.5, and 5.8  $\mu\text{m}$  data; 23 of those show PAH emission. Based on this small sample of sources, Carlson et al. (2011) defined two regions in the  $[3.6]-[4.5]$  versus  $[4.5]-[5.8]$  CCD, one occupied by YSO candidates with weak PAHs, and the other with those showing strong PAHs. All the sources with strong PAHs from Carlson et al. (2011) lie in the Population P area. In Carlson et al. (2012), we used the same CCD selection-criteria to identify PAH-enhanced YSO candidates in nine active star-formation regions in the LMC. As noted there, this enhancement would shift sources to redder colors in the  $[4.5]$  versus  $[4.5]-[8.0]$  CMD and both brighter and redder in the  $[8.0]$  versus  $[4.5]-[8.0]$  and  $[3.6]$  versus  $[3.6]-[5.8]$

CMDs. There should be little effect in  $[4.5]$  versus  $[4.5]-[24]$ , which sources may shift to the blue and brighter in  $[8.0]$  versus  $[8.0]-[24]$ . Thus, sources with PAH enhancement are likely to be moved into our color-selected CMD space.

## 5. REFINING THE YSO CANDIDATE LIST

Our subsequent analysis focuses on better characterizing our manually constructed list, understanding its reliability, and assessing the effectiveness of automated methods.

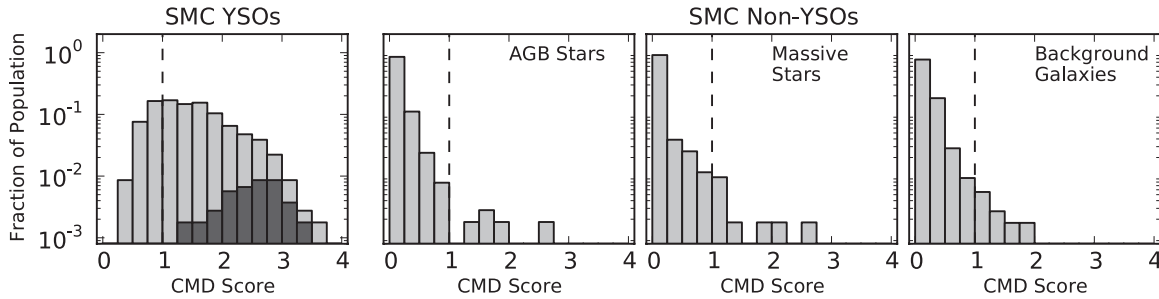
### 5.1. CMD Scores

We have developed a “CMD score” as a quantitative measurement of a YSO candidate’s position in color-magnitude space with respect to non-YSO sources. While visually inspecting every candidate YSO can be an effective means of separating true YSOs from spurious sources or non-YSO contaminants such as evolved stars or background galaxies, a visual examination is a highly labor-intensive procedure. Moreover, the process is highly subjective: if the same dataset is inspected by a number of investigators, a source could be classified differently by each investigator. Our CMD score calculation is a means of quantifying the varying certainties in YSO classification without the use of multiple inspections by a number of people.

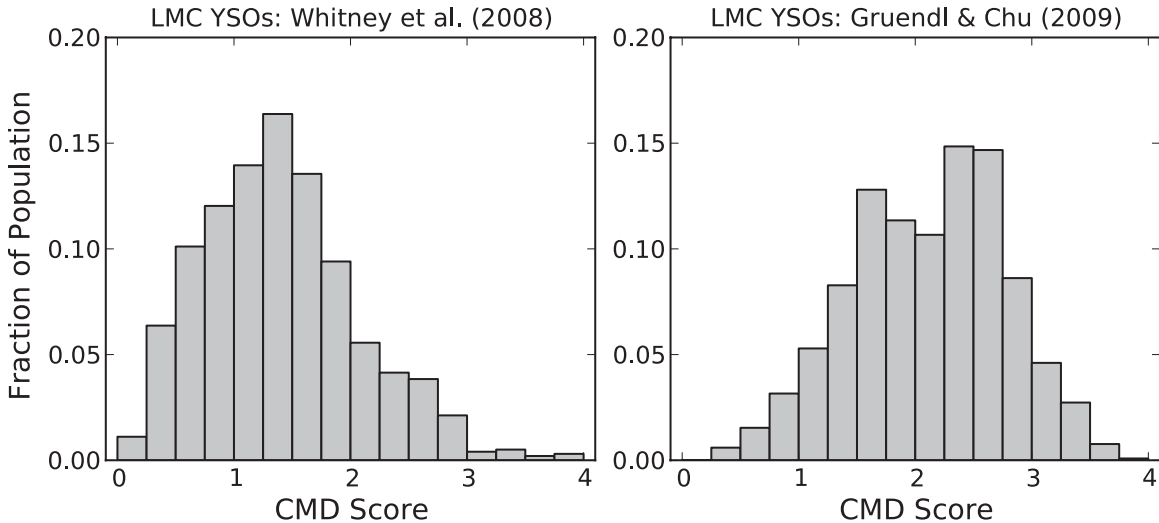
For each YSO we compare the color-magnitude position to those of sources photometrically confused with YSOs. These include evolved stars, background galaxies, and massive stars. This CMD score represents a measure of our confidence that the source is *not* a member of one of these other categories, making the basic assumption that a source is more likely to be a certain type of object (e.g., evolved star or background galaxy) if it is similar in color and magnitude to other objects of that type. The most securely classified YSOs are those that are most dissimilar in color and magnitude from non-YSOs. The farther a YSO candidate is from non-YSOs in color-magnitude space, the lower the probability that it is a non-YSO and the higher its CMD score. Thus, a higher CMD score indicates a higher confidence in a candidate’s classification as a YSO.

The exact method of CMD score calculation is presented in Appendix D. Our procedure uses all five of our selection CMDs and the catalogs of AGB stars and first-ascent red giant branch (RGB) stars (Boyer et al. 2011), galaxies (SWIRE; Lonsdale et al. 2003), and massive stars (Bonanos et al. 2010), presented in Appendix C. The method, however, is general and could be modified for use with other types of objects and other CMD combinations.

We calculate the CMD score for each YSO and scale by the total number of CMDs into which it falls, i.e., the number of valid flux measurements. The score is sensitive to the luminosity of the source in two ways: (1) brighter sources tend to have more reliable flux values and thus more CMDs, and (2) brighter sources are farther away in color-magnitude space from background galaxies. Therefore, our score calculation method preferentially assigns higher scores to the most luminous sources. This point underlines the fact that the score is *not* a measure of whether or not a source has YSO-like photometric properties. Indeed, bona fide low-luminosity YSOs have mid-IR photometry indistinguishable from background galaxies, and will thus receive low CMD scores. The approximate separation between YSO candidates and non-YSOs occurs at the CMD score of  $\sim 1$ . Note, however, that some sources previously identified as non-YSOs lie in the outskirts of CMD regions



**Figure 6.** Comparison between the distribution of CMD scores for the SMC YSO candidates (leftmost plot) and scores of non-YSO populations (from left to right): evolved stars, massive stars, and background galaxies. The spectroscopically confirmed YSOs are shown in dark gray in the first plot on the left. As expected, the score distributions of non-YSO populations peak at lower scores than the distribution of scores of the YSO candidates and spectroscopically confirmed YSOs. The approximate separation between YSO candidates and non-YSOs occurs at the CMD score of 1 indicated by a vertical dashed line.



**Figure 7.** Distribution of CMD scores for the LMC YSO candidates from Whitney et al. (2008, left) and Gruendl & Chu (2009, right). Similar to the SMC candidate YSOs, the score distributions for the LMC YSO candidates peak at high scores.

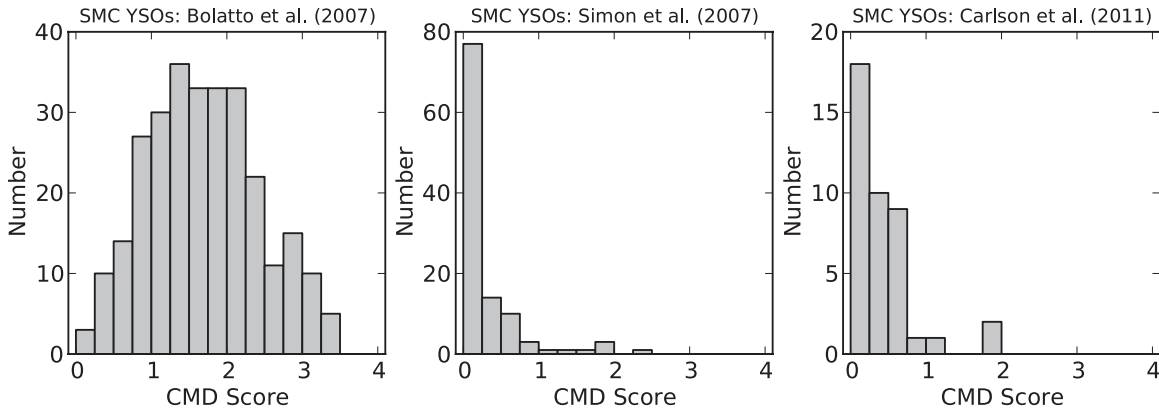
occupied by the population as a whole, giving them higher CMD scores. Moreover, some candidate YSOs are located in CMD regions potentially contaminated by non-YSO populations (see below). Together, the implication is that while the CMD scores are a good measure of the reliability of the YSO identification, because YSOs and non-YSOs can occupy similar regions of color–magnitude space, some contamination in the candidate YSO list by non-YSOs is expected.

We examine score distributions in Figure 6. The leftmost plot in Figure 6 shows the distribution of CMD scores for all the sources we choose as YSO candidates during the inspection. The distribution of CMD scores for spectroscopically confirmed YSOs is also included for comparison. The other three panels show the score distributions for non-YSO catalogs (AGBs, massive stars, and galaxies), sources that have previously been identified as non-YSOs in other studies have low CMD scores (distributions peak at score of 0–0.5), while those we identified as candidate YSOs (including all spectroscopically confirmed YSOs) have higher CMD scores. This result suggests that our selection process is effective in identifying sources separated in color–magnitude space from non-YSOs. For comparison, we show the scores of candidate YSOs in both the Whitney et al. (2008) and Gruendl & Chu (2009) LMC galaxy-wide studies in Figure 7. Like those identified in this study, the candidate YSOs in the LMC peak at relatively high scores:  $\sim 1.5$  and  $\sim 2$  for the Whitney et al. (2008) and Gruendl & Chu (2009) YSO lists, respectively.

In Figure 8, we show the distribution of scores for the previously identified SMC YSO candidates from Bolatto et al. (2007, left), Simon et al. (2007, middle), and Carlson et al. (2011, right). While the distribution of the CMD scores for the Bolatto et al. (2007) sample is similar to the distribution of YSO candidates from our list and those in the LMC, the YSO candidates identified by Simon et al. (2007) and Carlson et al. (2011) on a smaller scale—in NGC 346 and NGC 602 star formation regions, respectively—have low scores. This is a consequence of the fact that these two studies of individual H II regions selected relatively low-luminosity sources that lie close to or overlap with the CMD region occupied by background galaxies in the color–magnitude space. This result reiterates that the low CMD score does not indicate that the source is definitely not a YSO, but rather that it is more likely to be a contaminant. Low-score sources require a more careful analysis. As mentioned before, Carlson et al. (2011) showed that *HST* images are a powerful tool in distinguishing between YSOs and background galaxies. They are, however, only available for a couple of H II regions in the SMC.

### 5.2. SED Fitting with the YSO Models

We fit the SEDs of the YSO candidates with the Robitaille et al. (2006) YSO models using the Robitaille et al. (2007) SED fitting tool (see Section 3.3). The SED YSO model fitting provides more evidence supporting the YSO classification for sources that are well-fit. We constrain the physical parameters



**Figure 8.** Distribution of CMD scores for the previously identified YSO candidates from Bolatto et al. (2007, left), Simon et al. (2007, middle), and Carlson et al. (2011, right). The CMD scores for YSO candidates from the S<sup>3</sup>MC survey have a similar distribution to that for YSO candidates identified in this work (see Figure 6). However, studies that concentrated on individual star formation regions (NGC 346 and NGC 602; Simon et al. 2007 and Carlson et al. 2011, respectively) selected lower-luminosity YSO candidates that lie close to or overlap with the region occupied by background galaxies in the color–magnitude space (see Figure 2). As a result, these YSO candidates have low CMD scores.

of these sources. Stellar masses estimated from the fitting can be used to estimate a lower limit for the star formation rate (SFR) in the SMC (see Section 7.2).

#### 5.2.1. Preparations for the Fitting: Photometry and Photometric Uncertainties

As discussed in Section 2.3, for sources from the IRAC catalog, we combine the near-IR and mid-IR catalog and aperture photometry to constrain the SEDs as fit by the models. When available, we use photometry from the SAGE-SMC catalog and fill in photometry at missing wavelengths with aperture photometry. Because the aperture photometry generally overestimates the flux relative to the catalog (see above), one must be careful when combining the catalog and aperture photometry. A balance must be struck between the use of upper limits and the necessity of data points to constrain the fit.

We set the aperture photometry as an upper limit for sources with at least six catalog fluxes of the possible eight ( $JHK_s$ , all IRAC, and MIPS 24  $\mu\text{m}$ ;  $\sim 71\%$  sources) or with all five *Spitzer* bands in the catalog ( $\sim 12\%$ ). For sources with only three catalog fluxes ( $\sim 3\%$ ), we adopt aperture photometry only. The SEDs of all other sources (those with four or five catalog fluxes and at least one *Spitzer* band missing from the catalog) are inspected on an individual basis to determine if the aperture photometry values should be set as upper limits (e.g., if the source is not clearly visible or in an area of high confusion). For sources from the MIPS 24  $\mu\text{m}$  catalog without IRAC catalog matches, we use the near-IR and mid-IR aperture photometry fluxes for the SED fitting.

To account for systematic effects and calibration uncertainties (we combine multiple surveys), we reset the near- and mid-IR flux uncertainties to at least 10% in all bands. We further increase flux uncertainties for sources affected by PAHs (Population P sources; see Section 4.3). The current version of the fitter does not incorporate PAHs, thus the sources with PAHs are not well-fit with the YSO models even though they may be bona fide YSOs. To account for the fact that we fit the models without PAHs to the SEDs of sources showing the PAH emission, we increased the uncertainties of the measured fluxes for Population P sources to 20%, 30%, and 40% in the bands affected by the PAH emission: 3.6, 5.8, and 8.0  $\mu\text{m}$ , respectively (see also Carlson et al. 2011).

Optical data are available only for IRAC catalog sources (see Section 2.2). We use the MCPS  $UB$  and OGLE-III  $VI$  data

when available, supplemented by the MCPS  $VI$  data when the OGLE-III  $V$  and/or  $I$  fluxes are missing. We set optical fluxes in all bands as upper limits based on the results of studies showing that as many as 70% of single *Spitzer* sources at the distance of the Magellanic Clouds are in fact (proto)clusters when observed at high-resolution with the *HST* (Chen et al. 2009; Carlson et al. 2011; Oliveira et al. 2013; see also Section 4.1). It is uncertain whether an optical source is the same source as that detected by *Spitzer* or is a different object. Some of SEDs for sources from our sample are double-peaked, suggesting that different sources contribute to the SED at different wavelengths (see Section 6.2).

The 0.3–24  $\mu\text{m}$  combined photometry (both fluxes and magnitudes with uncertainties:  $F_U(\sigma_U), \dots, F_{24\mu\text{m}}(\sigma_{24\mu\text{m}})$ ;  $mag_U(\sigma_U), \dots, mag_{24\mu\text{m}}(\sigma_{24\mu\text{m}})$ ) used for the SED fitting is provided in Tables 1 and 2 for YSO candidates we classify in the following sections as high-reliability and possible YSO candidates, respectively. The sources are ordered by right ascension with names starting with a letter Y followed by a running number (Source Name). The IRAC and MIPS designations are included for the IRAC sources and MIPS 24  $\mu\text{m}$  sources with no IRAC counterparts, respectively (IRAC or MIPS Designation), as well as the J2000 equatorial coordinates in degrees. We include a three-digit flag for each band (columns  $Flag_U, \dots, Flag_{24\mu\text{m}}$ ). The first digit indicates what survey the flux in a given band comes from; the second digit tells whether it is from the catalog or aperture photometry; the third digit indicates whether the flux is used as an upper limit in SED fitting. A detailed description of the flags can be found in Table 1 footnotes. Tables 1 and 2 provide information on whether the source is affected by PAH emission (PAH; Section 4.3). Both tables also include CMD scores (CMDScore; Section 5.1) and the value of the normalized  $\chi^2$  per data point ( $\chi^2/pt$ ; see Section 5.2.2; Robitaille et al. 2007; Carlson et al. 2011) for the best fit YSO model ( $\chi_{\text{min}}^2/pt$ )—these two parameters are used for source classification.

Out of 1051 sources, 40 MIPS 24  $\mu\text{m}$  sources with no IRAC matches correspond to multiple IRAC sources and may be YSO candidate clusters. Due to their different nature, they are listed in a separate table. Table 3 provides the names (starting with a letter “C” to indicate cluster nature), MIPS designations, J2000 equatorial coordinates, aperture photometry fluxes and magnitudes with uncertainties, and flags. It also includes information on PAHs and CMD scores.

**Table 1**  
The 0.3–24  $\mu\text{m}$  Photometry of High-reliability YSO Candidates in the SMC

Source Name	IRAC or MIPS Designation <sup>a</sup>	R.A. (°, J2000)	Decl. (°, J2000)	$F_U(\sigma_U)^c$ (mJy)	...	$\text{mag}_U(\sigma_U)$	...	Flag $_U^d$	...	PAH <sup>e</sup>	Score	$\chi_{\text{min}}^2/pt^f$
Y1	SSTISAGEMA J001849.02-732129.8	4.70425	-73.35829	...	...	...	...	000	...	0	1.6	1.1
Y2	SSTISAGEMA J001956.85-732217.4	4.98688	-73.37152	...	...	...	...	000	...	0	1.5	4.3
Y3	SSTISAGEMA J002023.87-732021.3	5.09947	-73.33926	...	...	...	...	000	...	-1	1.7	13.6
Y5	SSTISAGEMA J002430.71-740138.2	6.12798	-74.02728	<0.101	...	>18.17	...	113	...	1	1.2	4.4
Y6	SSTISAGEMA J002432.91-740150.1	6.13715	-74.03059	<0.063	...	>18.69	...	113	...	1	1.1	6.5
Y7	SSTISAGEMA J002519.76-730824.5	6.33237	-73.14014	<0.020	...	>19.94	...	113	...	1	0.9	2.8
Y8	SSTISAGEMA J002602.56-724717.9	6.51069	-72.78833	<0.248	...	>17.20	...	113	...	0	1.6	7.8
Y9	SSTISAGEMA J002630.45-730048.2	6.62689	-73.01340	<0.179	...	>17.56	...	113	...	0	1.5	7.9
Y11	SSTISAGEMA J002726.22-723155.1	6.85925	-72.53199	...	...	...	...	000	...	-1	1.5	7.7
Y13	SSTISAGEMA J002806.82-731450.1	7.02843	-73.24727	...	...	...	...	000	...	0	0.6	0.9
Y33	SSTM1SAGE1 J003617.19-735958.8 <sup>b</sup>	9.07163	-73.99966	...	...	...	...	000	...	0	1.9	4.4

**Notes.**

<sup>a</sup> IRAC designations listed in the table are preceded by “SSTISAGEMA.” MIPS Epoch 1 and Epoch 2 designations start with “SSTM1SAGE1” and “SSTM1SAGE2,” respectively. For 24  $\mu\text{m}$  sources present in both epochs (indicated with a footnote “b”), the MIPS Epoch 1 designation is used.

<sup>b</sup> MIPS 24  $\mu\text{m}$  sources present in both Epoch 1 and Epoch 2 MIPS 24  $\mu\text{m}$  catalogs.

<sup>c</sup> The flux and magnitude uncertainties are listed in separate columns in the electronic table.

<sup>d</sup> Flags indicating what survey the flux in a given band comes from (first digit), whether the flux is from the catalog or was measured by aperture photometry (second digit), and whether it was used as an upper limit in the fitting (third digit). First digit: 0—no catalog data available or catalog photometry not adopted; 1—MCPS; 2—OGLE-III; 3—IRSF; 4—2MASS; 5—IRAC; 6—MIPS. Second digit: 0—no flux measurement available; 1—flux from the catalog; 2—flux from the aperture photometry. Third digit: 0—no data; 1—used as a valid data point in the fitting; 2—set as a lower limit, and 3—set as an upper limit for the fitting (see the text for details).

<sup>e</sup> For sources for which we could determine the PAH strength (i.e., with complete photometry between 3.6 and 5.8  $\mu\text{m}$ ): 0—sources without PAH emission; 1—sources associated with PAHs. For YSO candidates with incomplete photometry between 3.6 and 5.8  $\mu\text{m}$ , “PAH” is set to -1. See Section 4.3 for details.

<sup>f</sup>  $\chi_{\text{min}}^2/pt$  is a  $\chi^2$  per data point for the best-fit YSO model (see the text and Table 5).

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

**Table 2**  
The 0.3–24  $\mu\text{m}$  Photometry of Possible YSO Candidates in the SMC

Source Name	IRAC or MIPS Designation <sup>a</sup>	R.A. (°, J2000)	Decl. (°, J2000)	$F_U(\sigma_U)^c$ (mJy)	...	$\text{mag}_U(\sigma_U)$	...	Flag $_U^d$	...	PAH <sup>e</sup>	Score	$\chi_{\text{min}}^2/pt^f$
Y4	SSTISAGEMA J002048.76-732808.6	5.20318	-73.46907	...	...	...	...	000	...	0	0.9	4.3
Y10	SSTISAGEMA J002704.15-720353.8	6.76733	-72.06496	<0.278	...	>17.08	...	113	...	-1	0.8	12.8
Y12	SSTISAGEMA J002800.81-731444.3	7.00341	-73.24565	<0.047	...	>19.01	...	113	...	1	0.5	5.3
Y14	SSTISAGEMA J002810.43-725844.5	7.04348	-72.97905	...	...	...	...	000	...	0	3.1	7.2
Y17	SSTISAGEMA J003138.16-734748.0	7.90904	-73.79667	...	...	...	...	000	...	1	0.8	4.0
Y23	SSTISAGEMA J003331.66-714216.6	8.38194	-71.70464	<0.161	...	>17.68	...	113	...	0	0.7	3.6
Y26	SSTISAGEMA J003521.44-732658.6	8.83935	-73.44963	<0.100	...	>18.19	...	113	...	1	0.7	3.8
Y27	SSTISAGEMA J003525.14-732430.6	8.85479	-73.40851	...	...	...	...	000	...	1	0.6	3.9
Y30	SSTISAGEMA J003543.51-732110.7	8.93130	-73.35297	<0.127	...	>17.93	...	113	...	0	2.2	3.3
Y38	SSTISAGEMA J003654.19-735255.4	9.22582	-73.88208	<0.034	...	>19.38	...	113	...	1	0.7	3.9
Y124	SSTM1SAGE1 J004530.17-731241.1 <sup>b</sup>	11.37569	-73.21143	...	...	...	...	000	...	1	2.7	2.8

**Notes.**

<sup>a</sup> IRAC designations listed in the table are preceded by “SSTISAGEMA.” MIPS Epoch 1 and Epoch 2 designations start with “SSTM1SAGE1” and “SSTM1SAGE2,” respectively. For 24  $\mu\text{m}$  sources present in both epochs (indicated with a footnote “b”), the MIPS Epoch 1 designation is used.

<sup>b</sup> MIPS 24  $\mu\text{m}$  sources present in both Epoch 1 and Epoch 2 MIPS 24  $\mu\text{m}$  catalogs.

<sup>c</sup> The flux and magnitude uncertainties are listed in separate columns in the electronic table.

<sup>d</sup> Flags indicating what survey the flux in a given band comes from (first digit), whether the flux is from the catalog or was measured by aperture photometry (second digit), and whether it was used as an upper limit in the fitting (third digit). First digit: 0—no catalog data available or catalog photometry not adopted; 1—MCPS; 2—OGLE-III; 3—IRSF; 4—2MASS; 5—IRAC; 6—MIPS. Second digit: 0—no flux measurement available; 1—flux from the catalog; 2—flux from the aperture photometry. Third digit: 0—no data; 1—used as a valid data point in the fitting; 2—set as a lower limit, and 3—set as an upper limit for the fitting (see the text for details).

<sup>e</sup> For sources for which we could determine the PAH strength (i.e., with complete photometry between 3.6 and 5.8  $\mu\text{m}$ ): 0—sources without PAH emission; 1—sources associated with PAHs. For YSO candidates with incomplete photometry between 3.6 and 5.8  $\mu\text{m}$ , “PAH” is set to -1. See Section 4.3 for details.

<sup>f</sup>  $\chi_{\text{min}}^2/pt$  is a  $\chi^2$  per data point for the best-fit YSO model (see the text and Table 5).

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

### 5.2.2. The Fitting Results

Using the photometric data described above, we fit the Robitaille et al. (2006) YSO models to the SEDs of all color-selected sources with at least four valid data points (not

including upper limits) that are single sources in the *Spitzer* images. We exclude multiple sources listed in Table 3 from the fitting since the current version of the SED fitter does not account for multiplicity. Out of a total of 1011 single sources, 1007 fulfill the fitting requirements.

**Table 3**  
The 1.25–24  $\mu\text{m}$  Aperture Photometry of Possible YSO Candidate Clusters in the SMC

Source Name	MIPS Designation <sup>a</sup>	R.A. ( $^{\circ}$ , J2000)	Decl. ( $^{\circ}$ , J2000)	$F_J^{c,d}$ (mJy)	$\sigma_J$ (mJy)	...	$\text{mag}_J$	$\sigma_{\text{mag}_J}$	...	Flag <sub>J</sub> <sup>e</sup>	...	PAH <sup>f</sup>	Score
C1	SSTM1SAGE2 J004255.40-732410.9	10.73084	-73.40303	0.120	0.036	...	17.81	0.33	...	021	...	1	1.0
C2	SSTM1SAGE2 J004518.27-732244.3	11.32611	-73.37897	0.080	0.024	...	18.25	0.33	...	021	...	1	2.8
C3	SSTM1SAGE1 J004519.96-732214.5	11.33318	-73.37070	0.310	0.093	...	16.78	0.33	...	021	...	1	2.2
C4	SSTM1SAGE1 J004618.01-732335.5 <sup>b</sup>	11.57504	-73.39318	1.364	0.409	...	15.17	0.33	...	021	...	0	1.4
C5	SSTM1SAGE1 J004635.97-730537.4	11.64987	-73.09372	...	...	...	...	...	...	000	...	1	2.4
C6	SSTM1SAGE1 J004640.29-731528.4 <sup>b</sup>	11.66786	-73.25788	1.020	0.306	...	15.48	0.33	...	021	...	1	2.2
C7	SSTM1SAGE1 J004830.72-730629.5	12.12799	-73.10818	1.340	0.402	...	15.19	0.33	...	021	...	1	2.5
C8	SSTM1SAGE1 J004842.67-732523.2 <sup>b</sup>	12.17779	-73.42311	0.394	0.118	...	16.52	0.33	...	021	...	1	1.5
C9	SSTM1SAGE1 J004901.42-725354.0 <sup>b</sup>	12.25591	-72.89832	0.478	0.143	...	16.31	0.33	...	021	...	-1	0.3
C10	SSTM1SAGE2 J004925.13-732156.1	12.35471	-73.36559	0.534	0.160	...	16.19	0.33	...	021	...	1	1.8
C11	SSTM1SAGE1 J004926.52-732213.9	12.36052	-73.37051	0.733	0.220	...	15.84	0.33	...	021	...	0	1.7
C12	SSTM1SAGE1 J004933.74-732646.7 <sup>b</sup>	12.39059	-73.44630	0.548	0.164	...	16.16	0.33	...	021	...	1	1.9
C13	SSTM1SAGE2 J005006.73-725247.7	12.52803	-72.87993	0.248	0.074	...	17.02	0.33	...	021	...	1	1.2
C14	SSTM1SAGE1 J005047.63-724810.2 <sup>b</sup>	12.69847	-72.80283	0.623	0.187	...	16.02	0.33	...	021	...	1	1.7
C15	SSTM1SAGE1 J005140.29-731334.6 <sup>b</sup>	12.91788	-73.22627	1.308	0.392	...	15.21	0.33	...	021	...	0	2.4
C16	SSTM1SAGE1 J005256.91-731107.7	13.23712	-73.18548	...	...	...	...	...	...	000	...	1	1.4
C17	SSTM1SAGE2 J005647.24-722108.1	14.19684	-72.35226	0.438	0.131	...	16.40	0.33	...	021	...	1	2.3
C18	SSTM1SAGE2 J005705.51-733303.1	14.27295	-73.55086	0.191	0.057	...	17.30	0.33	...	021	...	1	0.9
C19	SSTM1SAGE1 J005845.12-721246.7 <sup>b</sup>	14.68802	-72.21297	0.190	0.057	...	17.31	0.33	...	021	...	1	1.0
C20	SSTM1SAGE2 J005855.86-720859.2	14.73276	-72.14979	0.256	0.077	...	16.98	0.33	...	021	...	1	2.0
C21	SSTM1SAGE2 J005912.25-722419.0	14.80106	-72.40527	0.617	0.185	...	16.03	0.33	...	021	...	1	1.7
C22	SSTM1SAGE2 J005923.30-720846.7	14.84707	-72.14630	0.157	0.047	...	17.51	0.33	...	021	...	1	1.1
C23	SSTM1SAGE2 J010006.60-721355.3	15.02750	-72.23202	9.450	2.835	...	13.07	0.33	...	021	...	0	1.2
C24	SSTM1SAGE1 J010042.35-720629.3 <sup>b</sup>	15.17644	-72.10815	0.145	0.043	...	17.60	0.33	...	021	...	1	1.2
C25	SSTM1SAGE2 J010058.27-713529.0	15.24280	-71.59138	2.124	0.637	...	14.69	0.33	...	021	...	1	2.8
C26	SSTM1SAGE1 J010131.45-715138.4	15.38105	-71.86066	0.185	0.056	...	17.34	0.33	...	021	...	1	2.0
C27	SSTM1SAGE2 J010134.15-715053.9	15.39230	-71.84830	0.267	0.080	...	16.94	0.33	...	021	...	1	2.2
C28	SSTM1SAGE1 J010314.80-715535.5 <sup>b</sup>	15.81166	-71.92653	<0.159	...	...	>17.51	...	...	023	...	1	1.1
C29	SSTM1SAGE2 J010340.87-720428.4	15.92028	-72.07454	0.271	0.081	...	16.92	0.33	...	021	...	1	1.1
C30	SSTM1SAGE2 J010348.82-720352.9	15.95342	-72.06469	0.896	0.269	...	15.63	0.33	...	021	...	1	2.5
C31	SSTM1SAGE1 J010504.48-715901.0	16.26866	-71.98363	0.583	0.175	...	16.09	0.33	...	021	...	0	2.0
C32	SSTM1SAGE2 J010540.77-720326.6	16.41987	-72.05740	0.277	0.083	...	16.90	0.33	...	021	...	0	1.6
C33	SSTM1SAGE2 J010540.94-720203.0	16.42060	-72.03416	0.133	0.040	...	17.70	0.33	...	021	...	1	1.0
C34	SSTM1SAGE2 J010617.64-742205.5	16.57349	-74.36819	1.890	0.567	...	14.81	0.33	...	021	...	0	1.4
C35	SSTM1SAGE1 J010906.86-731317.0	17.27858	-73.22140	0.049	0.017	...	18.78	0.38	...	021	...	1	1.3
C36	SSTM1SAGE2 J011349.04-731805.9	18.45432	-73.30165	1.032	0.310	...	15.47	0.33	...	021	...	0	2.3
C37	SSTM1SAGE1 J011359.28-731512.5 <sup>b</sup>	18.49698	-73.25346	<0.408	...	...	>16.48	1.00	...	023	...	1	2.6
C38	SSTM1SAGE2 J011433.51-731359.0	18.63964	-73.23305	0.026	0.013	...	19.47	0.52	...	021	...	1	1.0
C39	SSTM1SAGE1 J011445.98-731618.4 <sup>b</sup>	18.69158	-73.27179	<0.072	...	...	>18.36	...	...	023	...	1	1.2
C40	SSTM1SAGE1 J011446.58-731945.0 <sup>b</sup>	18.69410	-73.32918	1.027	0.308	...	15.48	0.33	...	021	...	1	2.6

**Notes.**

<sup>a</sup> MIPS Epoch 1 and Epoch 2 designations start with “SSTM1SAGE1” and “SSTM1SAGE2,” respectively. For 24  $\mu\text{m}$  sources present in both epochs (indicated with a footnote “b”), the MIPS Epoch 1 designation is used.

<sup>b</sup> MIPS 24  $\mu\text{m}$  sources present in both Epoch 1 and Epoch 2 MIPS 24  $\mu\text{m}$  catalogs.

<sup>c</sup> The flux and magnitude uncertainties are listed in separate columns in the electronic table.

<sup>d</sup> The  $JHK_s$  photometry was obtained using the IRSF images, except for source C34 for which 2MASS images were used (outside the IRSF survey area).

<sup>e</sup> Flags indicating what survey the flux in a given band comes from (first digit), whether the flux is from the catalog or was measured by aperture photometry (second digit), and whether it was used as an upper limit in the fitting (third digit). First digit: 0—no catalog data available or catalog photometry not adopted; 1—MCPS; 2—OGLE-III; 3—IRSF; 4—2MASS; 5—IRAC; 6—MIPS. Second digit: 0—no flux measurement available; 1—flux from the catalog; 2—flux from the aperture photometry. Third digit: 0—no data; 1—used as a valid data point in the fitting; 2—set as a lower limit, and 3—set as an upper limit for the fitting (see the text for details).

<sup>f</sup> For sources for which we could determine the PAH strength (i.e., with complete photometry between 3.6 and 5.8  $\mu\text{m}$ ): 0—sources without PAH emission; 1—sources associated with PAHs. For YSO candidates with incomplete photometry between 3.6 and 5.8  $\mu\text{m}$ , “PAH” is set to -1. See Section 4.3 for details.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

For the fitting we allow a distance range from 58.9 kpc to 66.4 kpc, which corresponds to the minimum and maximum distance to the SMC determined based on classical Cepheids ( $\sim 30$ – $300$  Myr; Haschke et al. 2012 and references therein). The latest distance measurement of  $63.1 \pm 3.0$  kpc by Haschke et al. (2012) is based on a large sample of  $\sim 2600$  Cepheids. The same work measured the depth of the SMC for the Cepheids to be

$\sim 5$ – $6$  kpc. The range of distances we assume accounts for both the distance uncertainties and the depth of the SMC. We consider models with interstellar extinction ( $A_V$ ) in a range from 0 to 10 mag (lower than in the Galaxy due to the lower dust-to-gas ratio; e.g., Leroy et al. 2007) in addition to the extinction caused by the YSO’s circumstellar environment which is intrinsic to the models.

**Table 4**  
Well-fit YSO Candidate Fitting Requirements

Score Range <sup>a</sup>	$\chi_{\min}^2/pt^b$	#YSOs <sup>c</sup> /Total
>2	$\leq 5$	126/171
1–2	$\leq 4$	274/564
0.5–1	$\leq 3$	104/242
0–0.5	$\leq 2$	3/7

**Notes.**

<sup>a</sup> The lower and upper endpoints are exclusive and inclusive, respectively.

<sup>b</sup>  $\chi_{\min}^2/pt$  is a  $\chi^2$  per data point for the best-fit YSO model.

<sup>c</sup> The ratio of the number of YSO candidates that fulfill the criteria listed in Column 2 to the total number of YSO candidates in the corresponding score range.

We examine SED fits by eye to define well-fit models. We quantify how well a source is fit by a given model SED by considering the value of normalized  $\chi^2/pt$ . The CMD scores are a measure of YSO identification reliability, and thus we allow higher  $\chi^2/pt$  for the high-score sources, while for lower-score sources we are more restrictive and require lower values of  $\chi^2/pt$ . Sources with high CMD scores are less likely to be contaminants than sources with low scores; however, their SED fits may not be well-constrained due to inadequacies of the models (e.g., PAH emission and multiplicity not incorporated in the models), intrinsic variability of the sources or bad data points. For low-score sources, we require additional evidence that they are YSOs; such evidence can be provided by good YSO model fits to their SEDs. We visually inspect SED fits for four CMD score ranges: 0–0.5, 0.5–1, 1–2, and >2 (see Figure 6) and determine the threshold  $\chi^2/pt$  for sources to be considered well-fit in each score range as listed in Table 4. The  $\chi^2/pt$  limit ranges from 5 for sources with scores >2 to 2 for sources with scores  $\leq 0.5$ .

We also perform the SED fitting with the stellar photosphere models to remove stars that could still be on our YSO candidate list (residual contamination by stars). Twenty seven sources (all with scores <0.5) were well-fit by the stellar photosphere models, and thus removed from the list. Out of 1011 single sources, 984 remain on the list.

Based on the scores and the SED fitting results, we define two categories of YSO candidates as follows.

1. High-reliability YSO candidates: all sources with CMD scores  $\geq 1$  (regardless of how well they are fit with YSO models) and sources with CMD scores <1 that are well-fit with YSO models. Sources with matches in the catalogs of evolved stars, PNe, or massive stars (see Section 6.3 and Appendix C) are excluded.
2. Possible YSO candidates: sources with CMD scores <1 that are not well-fit with the YSO models and those classified as non-YSO candidates in other photometric studies (see Table 6).

The high-reliability and possible YSO candidate lists contain 742 and 242 sources, respectively. We keep sources included in non-YSO photometric catalogs (see Section 6.3) on the possible YSO list since their non-YSO classification is unconfirmed, and thus the YSO classification is not ruled out. There is also a possibility of a mismatch between *Spitzer* sources and sources

identified at optical or near-IR wavelengths with higher angular resolution. Information about the alternative classification of sources is provided in Table 6, which also contains information on overlap between our catalog of candidate YSOs and previously known YSO candidates.

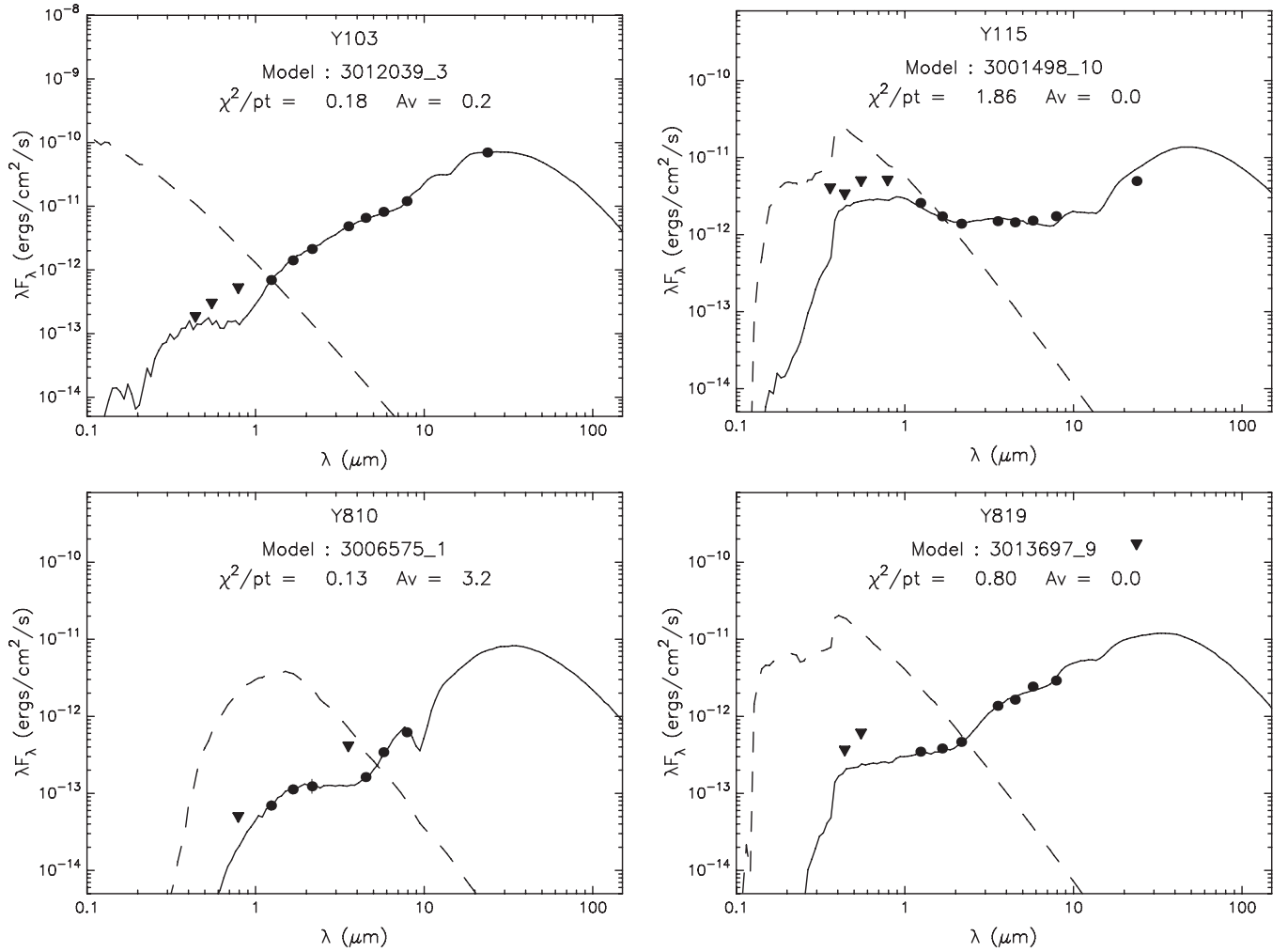
The SEDs and YSO fits for high-reliability and possible YSO candidates are shown in Appendix E. Example SEDs of different categories are presented in Figures 9 and 10.

Figure 9 shows four example SEDs of high-reliability YSO candidates with good YSO fits. Two YSO candidates whose SEDs are shown in the top panel represent high-score sources (2.6 and 1.7 for Y103 and Y115, respectively), while those in the bottom panel are examples of low-score sources (1.0 and 0.7 for Y810 and Y819, respectively). The low score of Y810 is mostly the result of the source having only three valid *Spitzer* fluxes, while for Y819 it is a combination of incomplete photometry and a close position to evolved stars in the color–magnitude space. Y103 and Y115 are spectroscopically confirmed YSOs (Oliveira et al. 2013). The SEDs of both sources are well-fit with Stage I YSO models of comparable age ( $<10^5$  yr) and masses of  $\sim 15 M_{\odot}$  and  $\sim 9 M_{\odot}$ , respectively. The difference in the shape of the model SEDs is the result of the models' different inclination angles: high (nearly edge-on) and low for models fitting the Y103 and Y115, respectively. The YSO with a flat SED can be confused with normal or active background galaxies (see Figure 14 for examples).

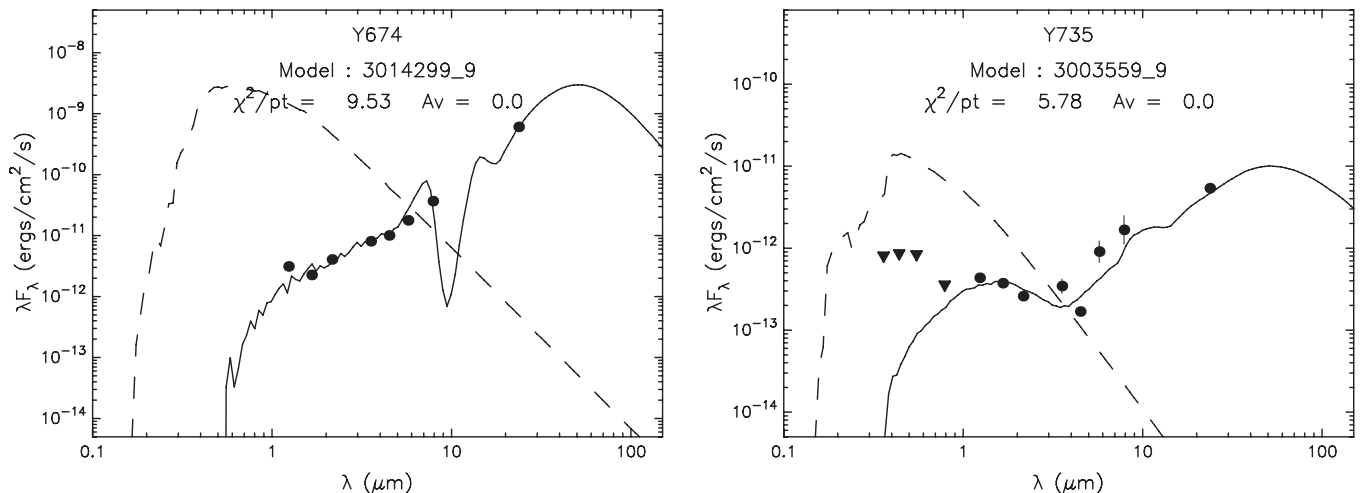
Figure 10 shows example SEDs of two high-reliability YSO candidates (Y674 and Y735) which are not well-fit with the YSO models. Both Y674 and Y735 are classified as high-reliability YSO candidates based on their high scores: 3.4 and 2.4, respectively. The poor YSO model fit to Y674's SED ( $\chi^2/pt \sim 10$ ) is most likely the result of the excess *J*-band emission. The 4.5  $\mu\text{m}$  dip in the Y735's SED indicates strong PAH emission, which is not accounted for in the YSO models, resulting in the poor YSO fit ( $\chi^2/pt \sim 6$ ) even with our adjustments to the 3.6, 5.8, and 8  $\mu\text{m}$  error bars.

We estimate physical parameters of well-fit sources by averaging parameters of all YSO models that fit the source's SED with  $\chi^2/pt$  in a range between the  $\chi^2/pt$  for the best-fitting model ( $\chi_{\min}^2/pt$ ) and  $\chi_{\min}^2/pt+2$ . The YSO model fits with  $\chi^2/pt$  within this arbitrary range of  $\chi^2/pt$  provide acceptable fits to the SEDs as determined by the visual examination of the fits. Robitaille et al. (2007) point out that it is not possible to derive formal confidence levels since the sampling of models in 14 dimensional space is too sparse to resolve the shape of the minima in the  $\chi^2$  surface. Robitaille et al. (2007) argue that this approach may result in overestimating the uncertainties on physical parameters, but these increased uncertainties would prevent overinterpretation of the results.

Table 5 provides estimates of stellar luminosity, stellar temperature, stellar mass, envelope mass, and disk mass, as well as derived evolutionary state for well-fit, high-reliability YSO candidates. The uncertainties of the average physical parameters represent standard deviations of the mean (see footnotes to Table 5). In addition to average quantities, Table 5 also includes physical parameters of the best-fit YSO model, which is plotted in the SED for each source (see Appendix E). The envelope mass and disk mass returned by the Robitaille et al. (2007) fitter are multiplied by 3.5 to account for the lower dust-to-gas ratio in the SMC than in the Galaxy (see a discussion in Section 4.1). The gas-to-dust ratio is a function of position in the SMC; the mean gas-to-dust ratio is a factor of 2 lower in the bar than in the wing (within the area covered by



**Figure 9.** Example SEDs for YSO candidates with different scores that are well-fit with the YSO models. The scores are: 2.6, 1.7, 1.0, and 0.7 for Y103, Y115, Y810, Y819, respectively. Sources Y103 and Y115 are spectroscopically confirmed YSOs (Oliveira et al. 2013) and have different shapes of the SEDs. Source Y115 with the flat SED can be confused with an AGN or QSO (see Figure 14). Sources Y810 and Y819 are examples of sources with low scores, but good YSO model fits. Source Y103 was classified as FIR object by Boyer et al. (2011) and source Y115 is in the Meyssonier & Azzopardi (1993) catalog of emission line objects. The solid black line in each plot shows the best fit YSO model (Robitaille et al. 2006). The dashed line is the central stellar atmosphere corresponding to the best fit model, extinguished by the fitted foreground extinction. Filled circles and triangles are valid flux values and flux upper limits, respectively. The values of a reduced  $\chi^2$  per data point ( $\chi^2/pt$ ) and interstellar visual extinction ( $A_V$ ) for the best-fit model are indicated in the plots. The YSO model names are also provided; these models can be viewed in detail at <http://caravan.astro.wisc.edu/protostars/>. The flux error bars are plotted if larger than the data points.



**Figure 10.** Example SEDs of two YSO candidates that are poorly fit with the YSO models. Both sources have high scores (3.4 and 2.4 for Y674 and Y735, respectively) and were thus classified as high-reliability YSO candidates. The bad YSO model fit to Y674's SED ( $\chi^2/pt \sim 10$ ) is most likely the result of the excess emission in the *J*-band. For Y735, the bad fit ( $\chi^2/pt \sim 6$ ) is caused by the strong PAH emission, which is not taken into account in the models. The symbols and lines are the same as in Figure 9.



**Table 5**  
Physical Parameters of the High-reliability YSO Candidates Well-fitted with the YSO Models

Source Name	$\chi_{\min}^2/pt^a$	$n_{\text{fits}}^b$	$L_{\star} (L_{\odot})$		$T_{\star} (K)$		$M_{\star} (M_{\odot})$		$M_{\text{env}} (M_{\odot})$		$M_{\text{disk}} (M_{\odot})$		Evol. Stage	
			best	ave <sup>c</sup>	best	ave	best	ave	best	ave	best	ave	best	ave
Y1	1.1	5	2825	6327(3202)	4147	20640(4389)	9.2	9.6(1.1)	833.8	166.8(166.8)	0.00	0.13(0.04)	I	I
Y7	2.8	61	83	591(117)	4405	8952(904)	3.2	4.0(0.3)	27.7	5.9(1.8)	0.09	0.09(0.01)	I	I
Y13	0.9	23	90	113(32)	4168	6203(963)	2.3	1.9(0.3)	113.9	22.7(8.8)	0.61	0.12(0.03)	I	I
Y18	1.5	2	4718	3837(880)	8479	8096(383)	11.2	10.6(0.6)	213.1	275.4(62.3)	0.06	0.17(0.10)	I	I
Y20	0.9	42	526	461(88)	15810	12890(850)	5.1	4.5(0.2)	0.3	242.6(142.9)	0.00	0.16(0.04)	II	I
Y24	1.6	18	405	962(154)	4140	8693(1394)	4.6	6.0(0.3)	2.9	137.5(70.4)	0.01	0.13(0.06)	I	I
Y25	0.6	56	47	45(7)	4002	3944(38)	1.3	1.2(0.1)	13.2	72.1(36.1)	0.02	0.10(0.02)	I	I
Y28	0.7	22	405	1179(117)	4162	12820(1725)	4.8	6.3(0.3)	39.8	52.2(19.9)	0.10	0.16(0.05)	I	I
Y31	2.8	40	2960	2490(114)	22660	21870(193)	8.1	7.7(0.1)	0.0	0.0(0.0)	0.00	0.02(0.01)	II	II
Y32	1.4	11	768	953(184)	4744	6077(1196)	8.0	7.1(0.4)	39.5	404.3(355.2)	0.02	0.36(0.15)	I	I

#### Notes.

<sup>a</sup>  $\chi_{\min}^2/pt$  is a  $\chi^2$  per data point for the best-fit model.

<sup>b</sup>  $n_{\text{fits}}$  is a number of fits with  $\chi^2/pt$  between  $\chi_{\min}^2/pt$  and  $\chi_{\min}^2/pt+2$ .

<sup>c</sup> The uncertainties represent standard deviations of the mean:  $s_{\text{mean}} = s/\sqrt{N}$ , where  $s$  is the sample standard deviation,  $s = \sqrt{(1/N-1)\sum_{i=0}^N(x-\bar{x})^2}$ . The uncertainties are listed in separate columns in the electronic table.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

the S<sup>3</sup>MC survey; Leroy et al. 2007). The physical parameters listed above can be rescaled to the local gas-to-dust ratio in the regions of interest. The evolutionary stages are defined as follows: Stage I YSOs have  $\dot{M}_{\text{env}}/M_{\star} > 10^{-6} \text{ yr}^{-1}$ . Stage II YSOs have  $\dot{M}_{\text{env}}/M_{\star} < 10^{-6} \text{ yr}^{-1}$  and  $M_{\text{disk}}/M_{\star} > 10^{-6}$ , and Stage III sources have  $\dot{M}_{\text{env}}/M_{\star} < 10^{-6} \text{ yr}^{-1}$  and  $M_{\text{disk}}/M_{\star} < 10^{-6}$ . Also listed in Table 5 are  $\chi_{\min}^2/pt$  and the number of model SEDs used to determine average properties.

The most robust physical parameters provided by the SED fitting are total luminosity and stellar temperature (except for heavily embedded sources, where the stellar temperature is not well constrained). These quantities are determined directly from the fitting, while the remaining parameters are model-dependent. For instance, stellar mass and age are derived using pre-main sequence evolutionary tracks, while the envelope infall rate is determined because it is directly related to the density profile through the assumption of a free-fall rotational collapse model. Since the determination of the envelope infall rates is uncertain, they are not provided in Table 5.

Using the averaged models (see Table 5), we find that the list of high-reliability YSO candidates well-fit with the models include: 428 (~95%) Stage I, 23 (~5%) Stage II, and 1 (<1%) Stage III sources. For the best-fit models, ~78%, ~15%, and ~8% sources are classified as Stage I, II, and III YSO candidates, respectively. We compare this result to the number of Stage I, II, and III YSO models from Robitaille et al. (2006) that fulfill our color-magnitude criteria (see Figure 3 and Section 4.1) and find a good agreement: ~81%, 19%, and 1% models represent evolutionary Stages I, II, and III, respectively. Approximately 83% YSO candidates with good YSO fits have stellar masses  $> 5 M_{\odot}$ .

## 6. ASSESSMENT OF THE LIST

We compare our list of YSO candidates to the catalogs of previously known YSOs (both spectroscopically confirmed and photometric candidates) and non-YSOs to assess its completeness and contamination. We also examine the *HST* optical image of the star formation region NGC 346 and compare it to *Spitzer* images to learn about the environments of the YSO candidates in this region, source confusion, and possible mismatches between

the optical and mid-IR sources that can produce double-peaked SEDs that are very common for sources in our sample. The analysis presented in this section shows that our method is efficient in the selected regions of color-magnitude space.

### 6.1. Comparison to Previous YSO Studies

Figure 11 shows the spatial distribution of all the YSO candidates from our list along with previously known YSO candidates and (spectroscopically) confirmed YSOs overlaid on the SAGE-SMC IRAC 8.0  $\mu\text{m}$  image. The vast majority of objects we identify as YSO candidates are associated with the main body of the SMC and are well-correlated with the 8.0  $\mu\text{m}$  emission that traces star-forming regions. To determine how many of the sources we select are newly identified YSO candidates, we compare our list with previously known YSO candidates and confirmed YSOs. Such a comparison allows us to test the effectiveness of our YSO identification method. A detailed description of this analysis is provided in the following sections.

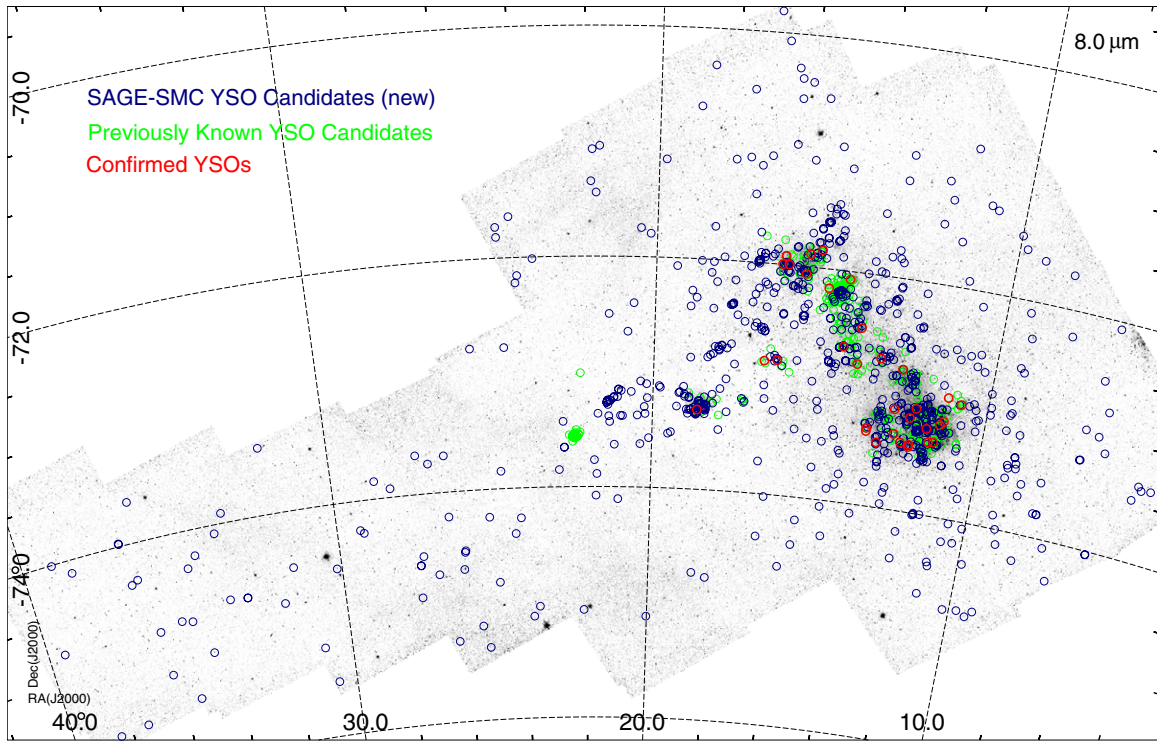
#### 6.1.1. Spectroscopically Confirmed YSOs

Of 33 confirmed YSOs from Oliveira et al. (2013), 31 are in the IRAC catalog, i.e., they fulfill color-magnitude criteria based on catalog fluxes. Two sources do not have IRAC catalog counterparts, but they are present in the MIPS 24  $\mu\text{m}$  catalog (SMCYSO01 and SMCYSO07). Using *Spitzer* fluxes from aperture photometry, we find that these two sources also meet our YSO color-magnitude criteria.

All spectroscopically confirmed YSOs have high scores (see Table 1 and Figure 6) and are classified as high-reliability YSOs. Of the 33 spectroscopic YSOs, 27 are well-fit with YSO models. These results show that our method of the YSO identification is effective.

#### 6.1.2. *Spitzer* Photometric Studies

We compare our list of the YSO candidates to the lists of YSO candidates identified by Bolatto et al. (2007), Simon et al. (2007), Carlson et al. (2007, 2011), and Gouliermis et al. (2007; see Section 3.1). These studies, based primarily on the *Spitzer* photometric data, identified a total of 456 YSO candidates



**Figure 11.** Spatial distribution of the previously known (green) and all newly identified (blue) YSO candidates (including possible YSO candidate clusters), as well as confirmed YSOs (red) in the SMC. The sources are overlaid on the  $8\ \mu\text{m}$  SAGE-SMC image. Previously known YSOs from Bolatto et al. (2007), Simon et al. (2007), Gouliermis et al. (2007), and Carlson et al. (2011). The spectroscopically confirmed YSOs from Oliveira et al. (2013). The distribution of the entire SAGE-SMC sample of YSO candidates (including those previously known) is shown in Figures 15–18. North is up, east to the left.

throughout the main body of the SMC (282 sources; Bolatto et al. 2007 in the  $S^3\text{MC}$  survey), in NGC 346 (111; Simon et al. 2007), and NGC 602 (41 and 22; Carlson et al. 2007, 2011; Gouliermis et al. 2007), respectively. Four sources overlap between Bolatto et al. (2007) and Simon et al. (2007) and 19 between Carlson et al. (2011) and Gouliermis et al. (2007), thus a combined list includes 433 individual YSO candidates.

We investigate how many of these sources our method of YSO identification should have revealed and compare the result to our final list of YSO candidates. Out of 433 previously known YSO candidates, 379 ( $\sim 88\%$ ) have matches in the SAGE-SMC IRAC catalog within a search radius of  $1''$ . It is not surprising that some of the Bolatto et al. (2007), Simon et al. (2007), Gouliermis et al. (2007), and Carlson et al. (2011) sources are missing from the SAGE-SMC catalog due to the higher sensitivities of the data they used— $S^3\text{MC}$  survey and the IRAC data from the *Spitzer* programs on NGC 602, respectively. The sensitivity of the  $S^3\text{MC}$  catalog is higher than that of the SAGE-SMC catalog that incorporates  $S^3\text{MC}$  data due to different data processing (see Appendix A).

Out of 379 sources that are included in the SAGE-SMC catalog, 241 fulfill our color–magnitude criteria based on the catalog fluxes. A significant number of the previously known YSO candidates, particularly those from Simon et al. (2007), Gouliermis et al. (2007), and Carlson et al. (2011), lie outside our color–magnitude selection, as can be seen in Figure 2. This result is not unexpected since our selection criteria are designed to select bright sources. Faint sources could be included reliably by Simon et al. (2007) and by Carlson et al. (2011) because they were inspected not only with *Spitzer*, but also with the high-resolution *HST* data, and contamination by non-YSOs is less likely in these known, active star-formation regions than across the galaxy as a whole. The YSO candidates in NGC 346

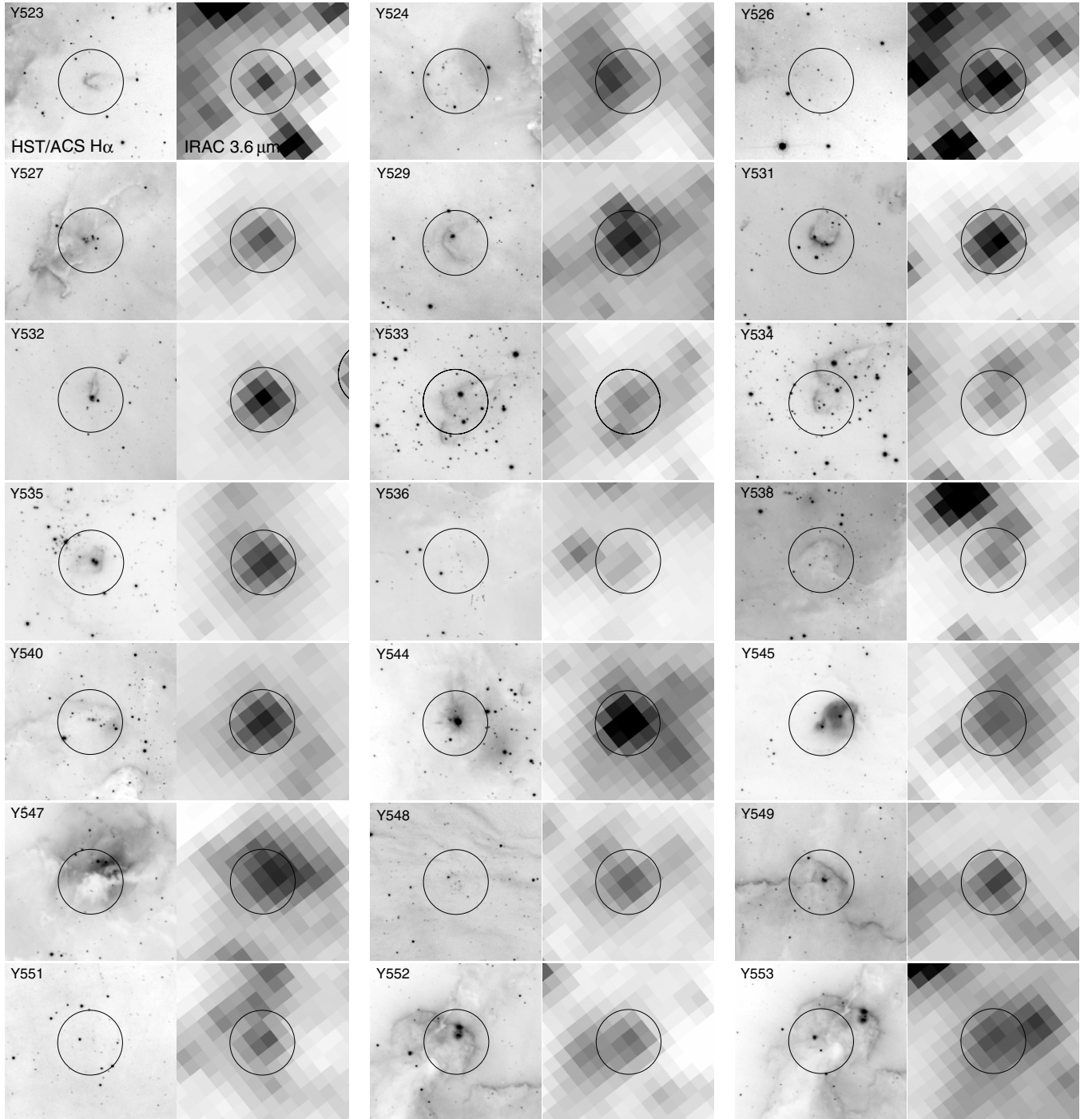
and NGC 602 are not confirmed and hence may still include non-YSOs.

Out of 241 sources that fulfill the criteria, 221 ( $\sim 92\%$ ) are on our final YSO candidate list (210 from Bolatto et al. 2007, 11 from Simon et al. 2007, and 4 from Carlson et al. 2011; three sources overlap). We remove the other 21 sources during visual inspection. Most of these sources are very faint and difficult to separate from background emission. Some are located immediately adjacent to another source, with the other source dominating aperture photometry. Moreover, they all have low scores (mostly because they lack catalog photometry or are faint enough to be located near many galaxies in the CMDs).

Recently, C.-H. R. Chen et al. (in preparation) identified 25 massive YSO candidates in the Magellanic Bridge based on the SAGE-SMC data. They used only one color–magnitude selection criterion combining IRAC 4.5 and  $8.0\ \mu\text{m}$  bands. The region in the color–magnitude space they selected extends to fainter [8.0] magnitudes than that used in this work (see Figure 2). As a consequence, only 11 of the C.-H. R. Chen et al. (in preparation) sources fulfill our color–magnitude selection criteria. Ten sources appear on our final list of candidate YSOs; one was removed as a probable background galaxy.

## 6.2. Optical View of YSO Candidates from Our List in NGC 346

In this case study, we compare the *HST*  $H\alpha$  and IRAC  $3.6\ \mu\text{m}$  images of sources from our YSO candidate list covered by the *HST* observations of the NGC 346 star formation region. The  $H\alpha$  image extracted from the Hubble Legacy Archive was obtained using the Wide Field Channel (WFC) of the *HST*'s ACS (F658N filter; program ID: GO-10248, PI: A. Nota). The spatial resolution of the *HST* at  $H\alpha$  ( $\sim 0''.07$  or  $0.02\ \text{pc}$  at a distance of

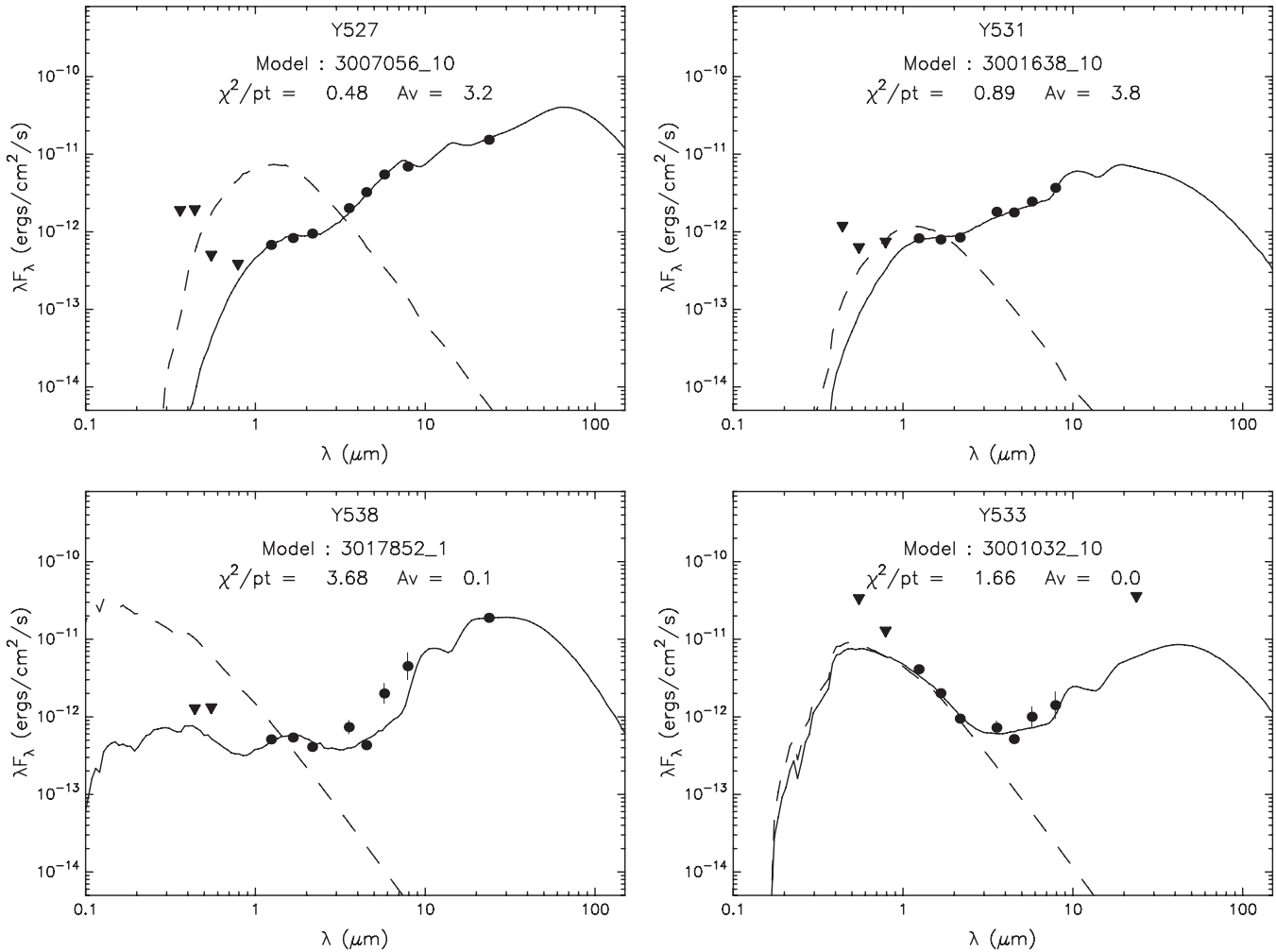


**Figure 12.** Comparison between the *HST*/ACS  $H\alpha$  and IRAC  $3.6\ \mu\text{m}$  images (left and right in each panel, respectively) of the YSO candidates in the NGC 346 star-forming region.

60 kpc) is over 20 times higher than *Spitzer*/IRAC resolution, allowing a detailed examination of the spatial distribution of stars and the nebular structure.

Figure 12 shows that most of the YSO candidates in NGC 346 are associated with  $H\alpha$  emission of diverse morphology. Many of the single or multiple optical counterparts of *Spitzer* YSO candidates can be found at the tip of or inside dusty pillars (e.g., Y531, Y529), while others are surrounded by small  $H\text{II}$  regions (e.g., Y545, Y552). In Figure 13, we show example SEDs for four YSO candidates from NGC 346 with different optical appearance. The *HST*  $H\alpha$  image of source Y527 reveals

a (proto)cluster of at least seven optical sources. Two optical sources located in the dusty pillar correspond to source Y531. Both Y538 and Y533 have single optical counterparts and may be associated with photodissociation regions since their SEDs indicate the presence of PAH emission (the  $4.5\ \mu\text{m}$  dip). The double-peaked SED of source Y533 suggests that the bright optical source identified as an O-type star by Bonanos et al. (2010) may not be a real counterpart of the *Spitzer* source even though they are separated by only  $0''.6$ . SEDs of Y527, Y531, Y538, and Y533 are well-fit with the YSO models (see Figure 13 and Section 5.2.2). The estimated stellar masses of the best-fit



**Figure 13.** Example SEDs for selected NGC 346 YSO candidates shown in the  $H\alpha$  and  $3.6\ \mu\text{m}$  image cutouts in Figure 12. Y527 appears as a protocluster in the  $H\alpha$  image. Y531 corresponds to two optical sources in a dusty pillar, while Y538 appears to be associated with a single optical source. Source Y533 with a double-peaked SED is an example of a mismatch between the *Spitzer* and optical sources. The  $4.5\ \mu\text{m}$  dip in the SEDs of Y538 and Y533 indicates the presence of PAH emission (see Section 4.3). The symbols and lines are the same as in Figure 9.

models shown in Figure 13 are  $12.2$ ,  $8.2$ ,  $8.4$ , and  $7.8\ M_{\odot}$  for Y527, Y531, Y538, and Y533, respectively.

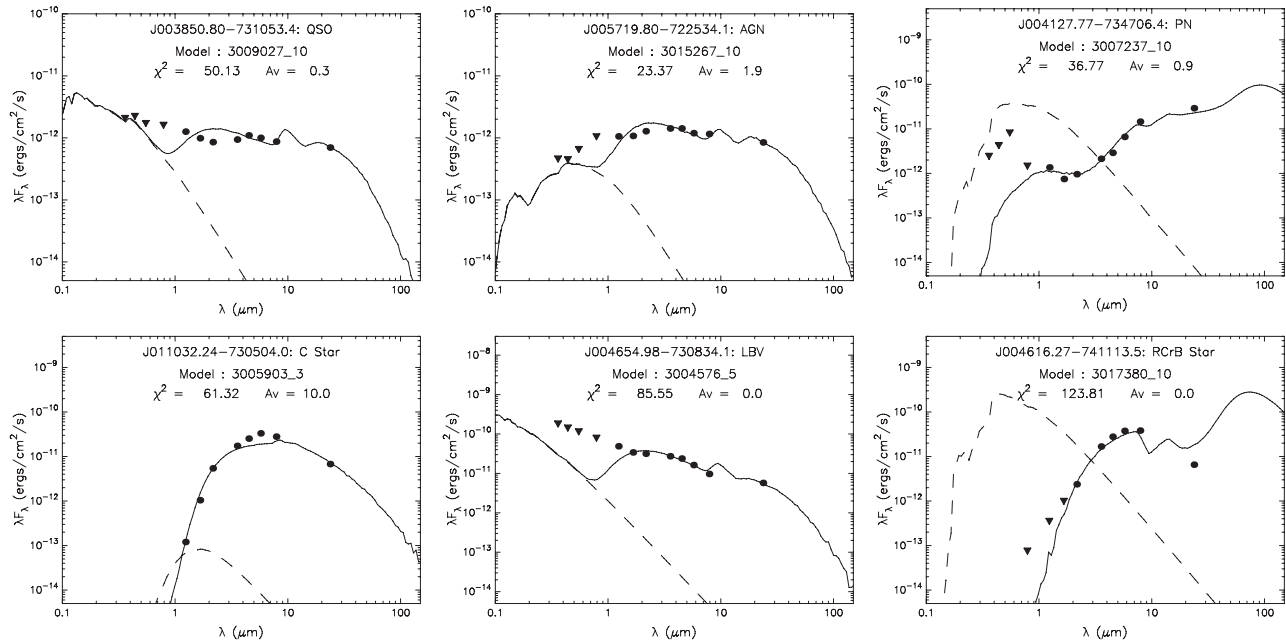
### 6.3. Contamination

Although our color–magnitude selection criteria avoid regions in the CMDs populated by non-YSO sources, Figure 2 shows that there is still an overlap between YSO candidates and other populations, particularly extreme-AGB (x-AGB) stars and PNe. To search for contaminants on the YSO candidate list, we match it to the catalogs of evolved stars (Boyer et al. 2011), massive stars (Bonanos et al. 2010), PNe (G. Jacoby 2009, private communication),  $H\alpha$  emission line stars and small nebulae (Meyssonnier & Azzopardi 1993), dusty OB stars (Sheets et al. 2013), and X-ray sources associated with radio emission (Sturm et al. 2013). The evolved stars, massive stars, and PNe are plotted in Figure 2; all catalogs are described in more detail in Section 3.2 and Appendix C.

Non-YSO sources are easily confused with YSOs because of their SED and/or color similarities. Figure 14 shows example SEDs with the best YSO model fits for sources representing non-YSO populations: a QSO (Q J00388-7310; Véron-Cetty & Véron 2010), an AGN (Q J00573-7225; Véron-Cetty & Véron 2010), a PN (SMP 6; Stanghellini et al. 2009), a carbon star (C star; S 23; Groenewegen & Blommaert 1998; van Loon et al.

2008), a luminous blue variable (LBV; R4 or AzV 16; Bonanos et al. 2010), and an R Coronae Borealis star (R CrB-type star; MSX SMC 014; Kraemer et al. 2005; van Loon et al. 2008). The classification of these objects was confirmed spectroscopically. All of these non-YSOs have scores larger than 1: 1.4, 1.43, 2.66, 1.37, 1.03, 1.45 for the QSO, AGN, PN, C star, LBV, and R CrB-type star, respectively. The AGN and PN are well-fit with the YSO models. PNe overlap with YSOs in the color–magnitude space (Figure 2) and may be confused with YSOs. Their IR spectra are similar to those of YSOs, and thus good fits with YSO models are not unexpected; however, IR-bright PNe have double-peaked SEDs. One peak of the PN’s SEDs is in the optical and the other in the IR as in the example shown in Figure 14. As mentioned above, the shape of the SEDs of AGNs and QSOs may be similar to some YSOs. Evolved stars are easier to distinguish from YSOs than PNe and both normal and active galaxies.

On our list of YSO candidates, we find 53 matches with evolved stars from Boyer et al. (2011) who used the same SAGE-SMC catalog for their identification as we use in this work. Out of 53, 37 sources were classified as far-IR (FIR) objects, i.e., sources with rising SEDs from 8 to  $24\ \mu\text{m}$  whose nature is unclear. These sources may be the dustiest evolved stars, but their characteristics are also consistent with YSOs.



**Figure 14.** SEDs for sources representing non-YSO populations (from top left to bottom right): QSO, AGN, PN, carbon star, LBV, and R CrB-type star (see the text for references). The scores are 1.4, 1.43, 2.66, 1.37, 1.03, 1.45, respectively. The C-rich AGB star and R CrB-type star were classified as x-AGBs in Boyer et al. (2011); it illustrates that the classification of sources based on the photometry alone is uncertain. The symbols and lines are the same as in Figure 9.

Indeed, nine of the far-IR objects from Boyer et al. (2011) have been confirmed spectroscopically as YSOs (Oliveira et al. 2013). Another confirmed YSO was classified as x-AGB by Boyer et al. (2011). Since the Boyer et al. (2011) selection criteria are based solely on color–magnitude cuts, their evolved star classification is not definitive. Thus, 53 sources from Boyer et al. (2011) remain on our list of YSO candidates with an annotation about another possible classification (Table 6). In addition to 37 FIR objects and one x-AGB star spectroscopically identified as YSOs, the Boyer et al. (2011) sources matched to YSO candidates list include one O-rich AGB star, three anomalous O-rich AGB stars (aO-AGBs), two C-rich AGB stars, and nine additional x-AGB stars.

Using a search radius of  $1''$  between our list of candidate YSOs and the catalogs of massive stars, PNe, emission line stars, and dusty OB stars, we find 9, 23, 41, and 4 matches, respectively. We do not find any matches with the catalog of X-ray sources associated with radio emission. Massive stars matching YSO candidates include one O-type star, five early B stars, one late B star, one sgB[e], and one LBV (see Section 4.1). In the sample of 41 emission line stars, 8 are identified as compact or small H II regions, 6 as very low excitation objects (VLE) objects, 17 as PNe, and 1 peculiar H $\alpha$  emission-line star with Fe II and [Fe II] emission. Dusty OB stars matching YSO candidates include three emission line objects (including a YSO and an H II region) and one main-sequence O7 star.

As mentioned in Section 5.2.2, Table 6 provides information about positional coincidence of the SAGE-SMC YSO candidates with non-YSO sources from the catalogs discussed above, as well as with previously known YSO candidates and confirmed YSOs (see Section 6.1). Table 6 lists alternative classifications, source names from other surveys, distances between the SAGE-SMC positions and the positions of the matching sources, and references. Although we consider the matches within  $1''$  between the YSO candidates and non-YSO catalogs the most reliable, sources separated by larger distances can be real counterparts as well. In Table 6, we report matches within  $2''$  rather

than only those within  $1''$  that we discuss above and allow the reader to decide whether the matches at distances larger than  $1''$  are believable.

Table 6 also includes information about SIMBAD matches (sources not confirmed spectroscopically) within  $1''$  to 11 YSO candidates. These include seven Be star candidates, one QSO candidate, two X-ray sources, and one H II region. Another two sources match a Cepheid and a red supergiant, but at a larger distance ( $\sim 1''.7$ ). The classification of these SIMBAD sources is uncertain and thus the matching YSO candidates were not removed from the list.

Background galaxies are another potential contaminant on our YSO candidate list. Although our selection criteria in all five CMDs avoided regions occupied by background galaxies, some of the sources that are brighter than galaxies in one CMD, may overlap with background galaxies in another CMD. We use the data from the SWIRE’s Lockman Hole field to assess the degree to which galaxies may be present on our list. Out of 323,044 sources in the SWIRE’s Lockman Hole field ( $\sim 75\%$  of which were detected at 3.6 and 4.5  $\mu\text{m}$  only), only 323 fulfill our color–magnitude criteria. All of these sources also fulfill our requirement for at least three *Spitzer* detections. We remove sources below SAGE-SMC sensitivity limits and those well-fit with the stellar photosphere models; the resulting list of background galaxy candidates in the Lockman Hole contains  $\sim 280$  sources. About 55% of these sources are detected in all five *Spitzer* bands, and additional 5.4% have all IRAC fluxes but are not detected at MIPS 24  $\mu\text{m}$ . The SAGE-SMC total survey area of  $\sim 36$  square degrees (the main body of the SMC and part of the Magellanic Bridge) is about three times larger than the SWIRE’s Lockman Hole field (11 square degrees). Thus, neglecting obscuration by the interstellar medium of the SMC we estimate that as many as  $\sim 900$  background galaxies are expected to contaminate our color–magnitude selected (pre-inspection) source list. The majority of these sources are bright since our criteria are biased toward bright objects, thus the galaxies’ morphology should be identifiable and removed during

**Table 6**  
Previous Classifications of the YSO Candidates

Source Name	R.A. (J2000) (deg)	Decl. (J2000) (deg)	YSO				Non-YSO			
			Class. <sup>a</sup>	Source ID	Dist. (")	Ref.	Class. <sup>b</sup>	Source ID	Dist. (")	Ref.
Y14	7.04348	-72.97905	...	...	...	...	FIR_object	SSTISAGEMA_J002810.43-725844.5	0	B11
Y30	8.93130	-73.35297	...	...	...	...	FIR_object	SSTISAGEMA_J003543.51-732110.7	0	B11
Y46	9.38775	-72.87917	...	...	...	...	X-AGB	SSTISAGEMA_J003733.06-725245.0	0	B11
Y54	9.97476	-72.85682	...	...	...	...	O-AGB	SSTISAGEMA_J003953.94-725124.5	0	B11
Y57	10.20749	-74.77190	...	...	...	...	aO-AGB	SSTISAGEMA_J004049.79-744618.8	0	B11
Y60	10.28824	-73.11309	...	...	...	...	PN	MA22:PNE_J0041093-730647	0.6	MA93;GJ
Y62	10.30878	-74.19174	...	...	...	...	FIR_object	SSTISAGEMA_J004114.10-741130.2	0	B11
Y63	10.35147	-73.02428	...	...	...	...	FIR_object	SSTISAGEMA_J004124.35-730127.4	0	B11
Y67	10.53679	-73.51904	...	...	...	...	early_B	SSTISAGE1C_J004208.74-733108.4	0.6	B10
Y76	10.75696	-73.34754	YSO_c	S3MC_J004301.63-732051.32	0.2	B07	...	...	...	...

**Notes.**

<sup>a</sup> “YSO\_s” and “YSO\_c” indicate a spectroscopically confirmed YSO and a YSO candidate, respectively.

<sup>b</sup> “PN”—planetary nebula; “em”—emission line object; “VLE”—a very low excitation emission-line object; “H II”—H II region; “FIR\_object”—source with a rising SED from 8 to 24  $\mu$ m, from Boyer et al. (2011); “C-AGB,” “O-AGB,” and “x-AGB” indicate C-rich, O-rich, and extreme AGB stars, respectively, and “aO-AGB” indicates a new class of O-rich AGB stars (Boyer et al. 2011); “early\_B”—early B-type star; “sgB[e]”—B[e] supergiant; “symbiotic\_star”—a symbiotic star from Oliveira et al. (2013); “X-ray”—unclassified X-ray source; “X-ray/AGN”—X-ray source classified as an AGN candidate; “QSO\_candidate”—a quasar candidate; “Cepheid”—a Cepheid variable star; “Be\*” and “Be\*candidate”—confirmed and candidate Be stars, respectively; “RSG”—red supergiant.

**References.** (A07) Atlee & Gould 2007; (B07) Bolatto et al. 2007; (B10) Bonanos et al. (2010) and references therein; (B11) Boyer et al. 2011; (C11) Carlson et al. 2011; (G82) Gatley et al. 1982; (GJ) G. Jacoby (2009, private communication); (G07) Gouliermis et al. 2007; (H56) Henize 1956; (H99) Hummel et al. 1999; (MS93) Meyssonier & Azzopardi 1993; (M08) McGowan et al. 2008; (O13) Oliveira et al. 2013; (P83) Prevot et al. 1983; (S02) Sharpee et al. 2002; (S07) Simon et al. 2007; (S13) Sheets et al. 2013; (vL10) van Loon et al. 2010a; (W06) Wisniewski & Bjorkman 2006.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

the inspection. YSO candidates with luminosities comparable to the luminosities of background galaxy population as a whole (see Figure 2) are redder and separate from them well in the CMDs. Most of these low luminosity sources belong to Population P and are mostly selected based on one CMD – [4.5] versus [4.5]–[5.8]. The PAH emission makes them very red in the [4.5]–[5.8] color.

## 7. FURTHER ANALYSIS

We use the list of the YSO candidates to study a correlation between star formation and molecular (CO) and neutral (H I and H $\alpha$ ) gas in the SMC. We also use the list of well-fit, high-reliability YSO candidates (sources with estimated physical parameters) to determine for the first time in the SMC the SFR based on the YSO counts and compare it to SFRs determined using other methods.

### 7.1. Comparison to Gas Tracers

Figure 15 shows a three-color composite image of the SMC combining the H $\alpha$  (red), H I (green), and CO (blue) images (Section 3.4) with the area covered by the SAGE-SMC survey overlaid. The H $\alpha$ , H I, and CO survey areas are outlined in corresponding colors. The H I survey covers the entire SMC and the tail, while the MCELS H $\alpha$  survey concentrates on the bar and the wing. The NANTEN CO survey is the most spatially limited and only covers the northeast and southwest regions of the bar, and the H II regions N 84 and N 88 in the wing. Several sites of CO emission were identified in the Bridge by Mizuno et al. (2006) and are indicated with blue circles in Figure 15.

Figures 16 and 17 show that YSO candidates appear to be preferentially located in regions of the brightest H I and H $\alpha$  emission, respectively, in the bar and wing. They also appear to be well-correlated with the H I emission in the tail; however, the SAGE-SMC survey area does not cover the whole extent of the H I image. The black contour in Figure 16 corresponds to an H I column density of  $10^{21}$  cm $^{-2}$ , which is roughly the lower limit for column density of H I gas that correlates well with young massive stars (Muller et al. 2003 and references therein). The correlation of the YSO candidates with higher column density H I gas (within the area delineated by the black contour) provides supporting evidence that a source is a YSO. YSO candidates correlated with lower column density H I gas are more likely to be confused with background galaxies. Only high resolution images (e.g., *HST*) and/or spectroscopic observations can differentiate between these two classes of objects.

The NANTEN survey identified 21 giant molecular clouds (GMCs) in the observed area of the SMC with masses of  $\sim 10^4$ – $10^6 M_{\odot}$  (Mizuno et al. 2001). They found a good correlation between GMCs and both young clusters and H II regions, indicating that cluster formation is ongoing in the GMCs. In Figure 18, we compare the positions of YSO candidates from our list to CO emission. The image shows that in general there is a good correlation between CO emission and YSO candidates; however, not all of GMCs are associated with YSOs. For instance, the CO cloud at (R.A., Decl.)  $\sim$  (12 $^{\circ}$ 5,  $-73^{\circ}$ 5) has no significant correlation with YSO candidates. One source at the edge of this CO cloud may be a chance superposition as there are more YSOs spread around that are not associated with bright CO. Out of the seven sites of CO emission in the Bridge, three are associated with the YSO candidates from our sample (see Figure 15). However, a more detailed study of the Magel-

**Table 7**  
H I, CO, and H $\alpha$  Image Pixel Values at the Position  
of YSO and YSO Cluster Candidates<sup>a</sup>

Source Name	H I ( $10^{20}$ cm $^{-2}$ )	CO (K km s $^{-1}$ )	H $\alpha$ ( $10^{-17}$ erg cm $^{-2}$ s $^{-1}$ )
Y800	60.9	<0.45	1994.5
Y801	18.8	...	2924.2
Y802	60.9	<0.45	1459.9
Y803	63.4	0.6	1579.2
Y804	61.0	<0.45	918.9
Y805	46.4	<0.45	3135.9
Y806	45.3	...	116.0
Y807	58.1	<0.45	125.8
Y808	69.3	1.2	2252.2
Y809	69.3	1.6	2537.4

**Note.** <sup>a</sup> Upper limits for the CO and H $\alpha$  pixel values represent the  $3\sigma$  sensitivity limits: 0.45 K km s $^{-1}$  for CO (Mizuno et al. 2001) and  $9 \times 10^{-17}$  erg cm $^{-2}$  s $^{-1}$  (measured in the region off galaxy) for H $\alpha$ . For H I, we use the image that combines the H I observations of the main body of the SMC by Stanimirović et al. (1999) and the Bridge by Muller et al. (2003). Due to the processing of the image (E. Muller 2013, private communication), all but one pixel have values larger than the  $3\sigma$  sensitivity limits of the individual H I observations (H I column density of  $\sim 1.2 \times 10^{19}$  cm $^{-2}$ ).

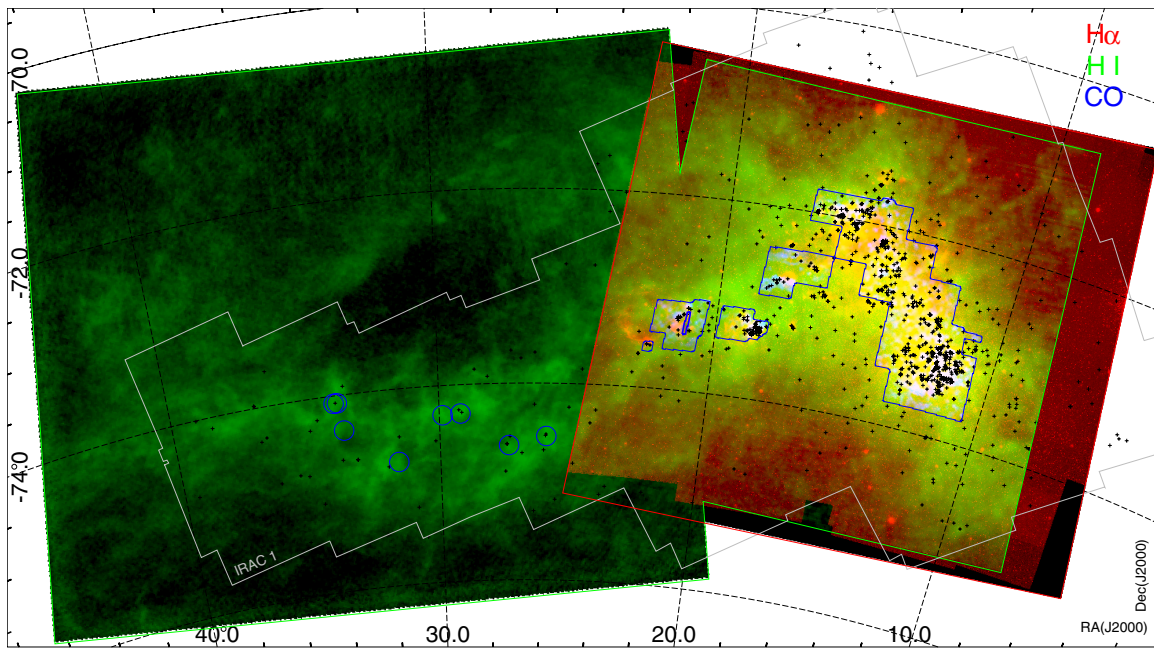
(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

lanic Bridge by C.-H. R. Chen et al. (in preparation) identified a small number of YSO candidates in two of these regions: two embedded and three fainter YSO candidates that did not pass our color–magnitude criteria.

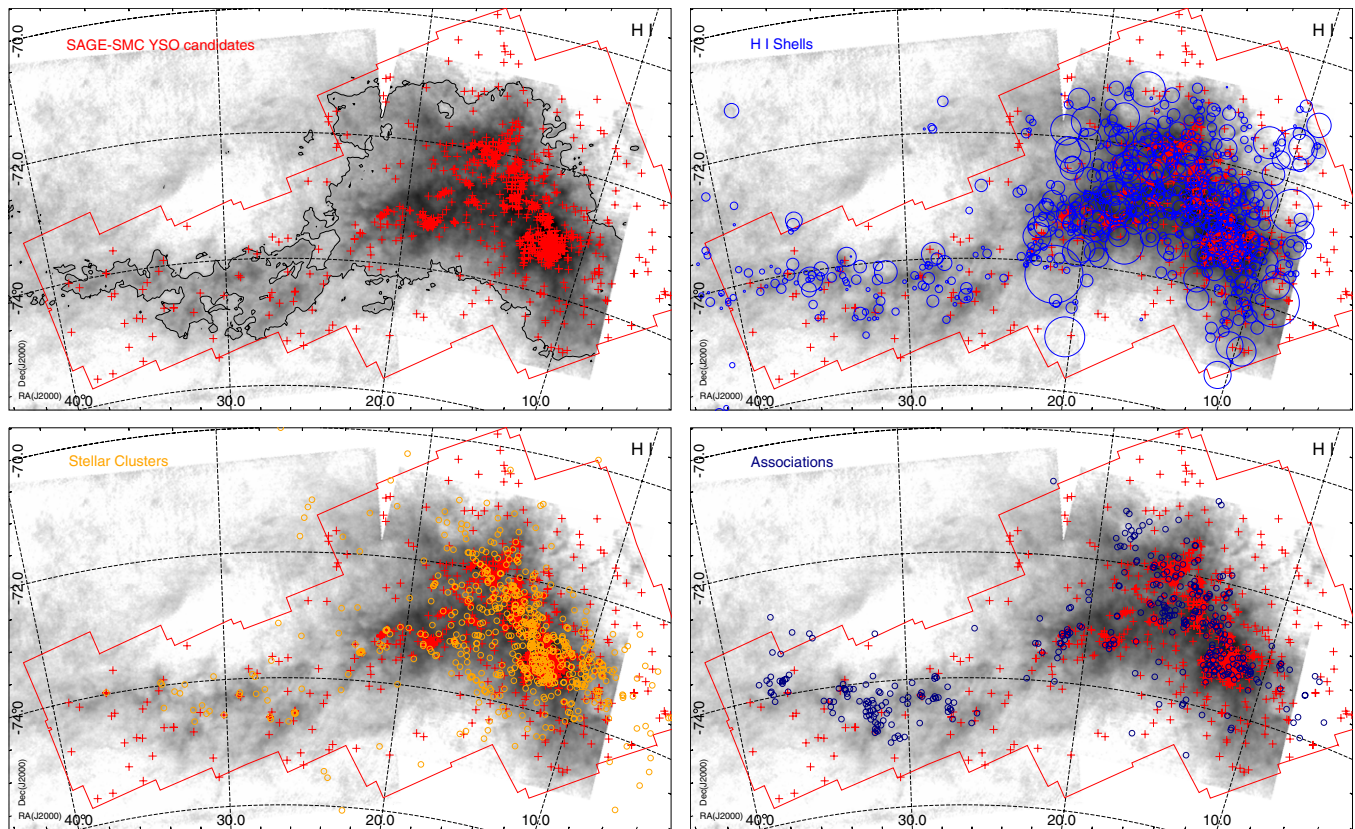
To quantify our results, we construct pixel value histograms for the entire H $\alpha$ , H I, and CO images (Figure 19, shown in white) and compare them to the pixel value histograms for pixels corresponding to YSO candidates' locations in each image (gray; pixel values are listed in Table 7). Since the area covered by the H $\alpha$  and H I images is larger than of the SAGE-SMC survey, we only considered H $\alpha$  and H I pixels that overlap with IRAC coverage. To match YSO selection criteria, we consider only the H $\alpha$ , CO, and H I data covered by at least three different *Spitzer* bands. For each gas tracer in Figure 19, the histogram of all the pixels is divided by the histogram of pixels associated with YSOs; the result is shown as black dots. The leftmost bin in the CO, H $\alpha$ , and H I histograms includes all pixels below the  $3\sigma$  sensitivity limit of the observations as marked with the vertical dashed line. Figure 19 shows that in each plot the distributions of pixels associated with the YSOs peak at high surface densities/brightnesses, confirming that YSO candidates are indeed associated with regions of high H I column densities and high CO and H $\alpha$  surface brightnesses. Note, however, that while regions bright in CO are more likely to be associated with YSOs, 86% of YSOs are located in regions below the  $3\sigma$  sensitivity limit of the observations. This indicates that a significant fraction of SMC star formation occurs in CO-free or CO-dim gas. The peak at the low H I column density in Figure 19 is associated with pixels in the Magellanic Bridge where the H I emission is fainter and subtends larger area than that in the bar and the wing of the SMC.

#### 7.1.1. Extended Distribution and Isolated YSOs

All the sources on our final list are YSO candidates, and they require spectroscopic confirmation. We do not include a criterion

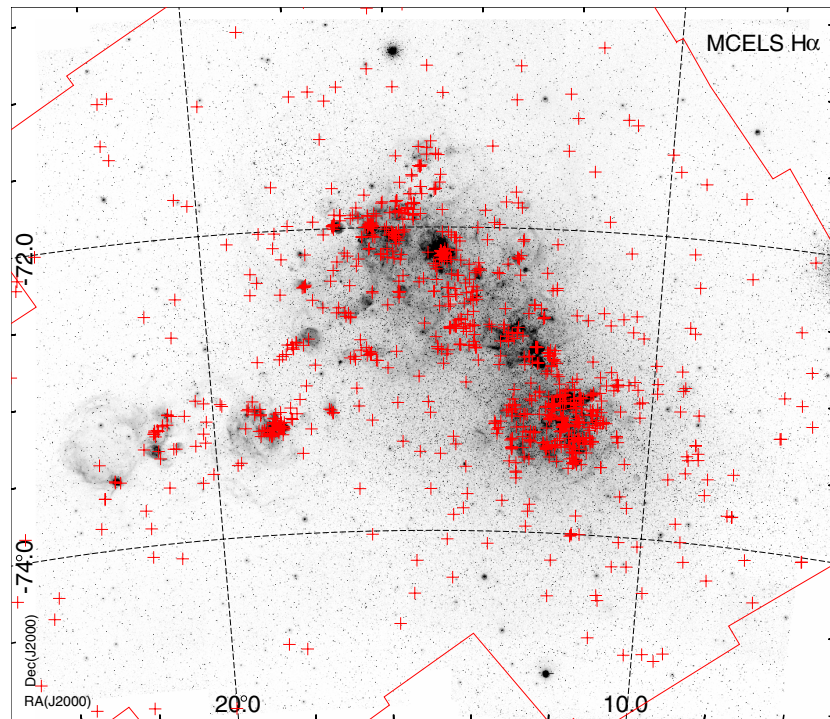


**Figure 15.** Three-color image showing the distribution of the H $\alpha$  (red; H $\alpha$  image is provided by F. Winkler, S. D. Points, R. C. Smith, the MCELS Team and NOAO/AURA/NSF), H I (green; Stanimirović et al. 1999, Muller et al. 2003), and CO (blue; NANTEN, Y. Fukui 2012, private communication) gas. Blue circles indicate the positions of the CO emission sites detected in the Bridge by Mizuno et al. (2006). The YSO candidates are indicated with plus signs (+). The H $\alpha$ , H I, and CO survey areas are outlined in corresponding colors. The area covered by the SAGE-SMC survey with IRAC at 3.6  $\mu\text{m}$  is outlined in grey. North is up, east to the left.

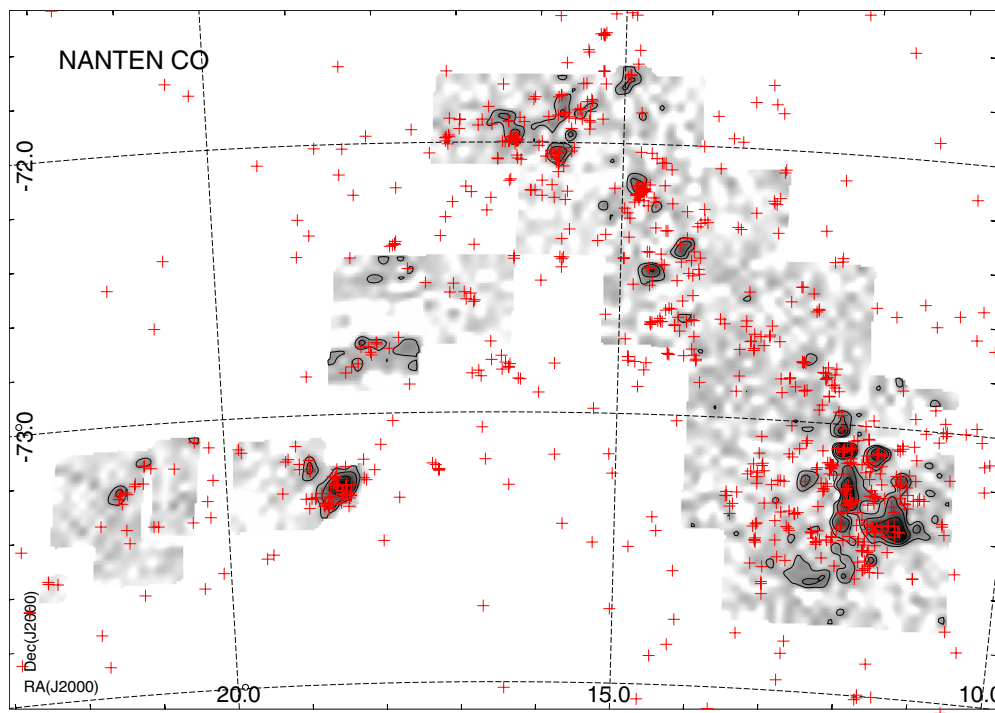


**Figure 16.** Distribution of the YSO candidates (+, upper left; also shown in other images) and other populations of sources (upper right: H I shells and supershells, lower left: stellar clusters; lower right: associations) with respect to the H I emission. The H I column density peak is  $1.4 \times 10^{22} \text{ cm}^{-2}$ . The black contour in the upper left plot corresponds to the H I column density of  $10^{21} \text{ cm}^{-2}$ , which is roughly the lower limit for column density of the H I gas that correlates well with young massive stars (Muller et al. 2003 and references therein). The area covered by the SAGE-SMC survey with IRAC at 3.6  $\mu\text{m}$  is shown as a red solid line. North is up, east to the left.





**Figure 17.** Distribution of the YSO candidates (+) with respect to the  $H\alpha$  emission. The vast majority of the YSO candidates are associated with the brightest regions of  $H\alpha$  emission. North is up, east to the left.

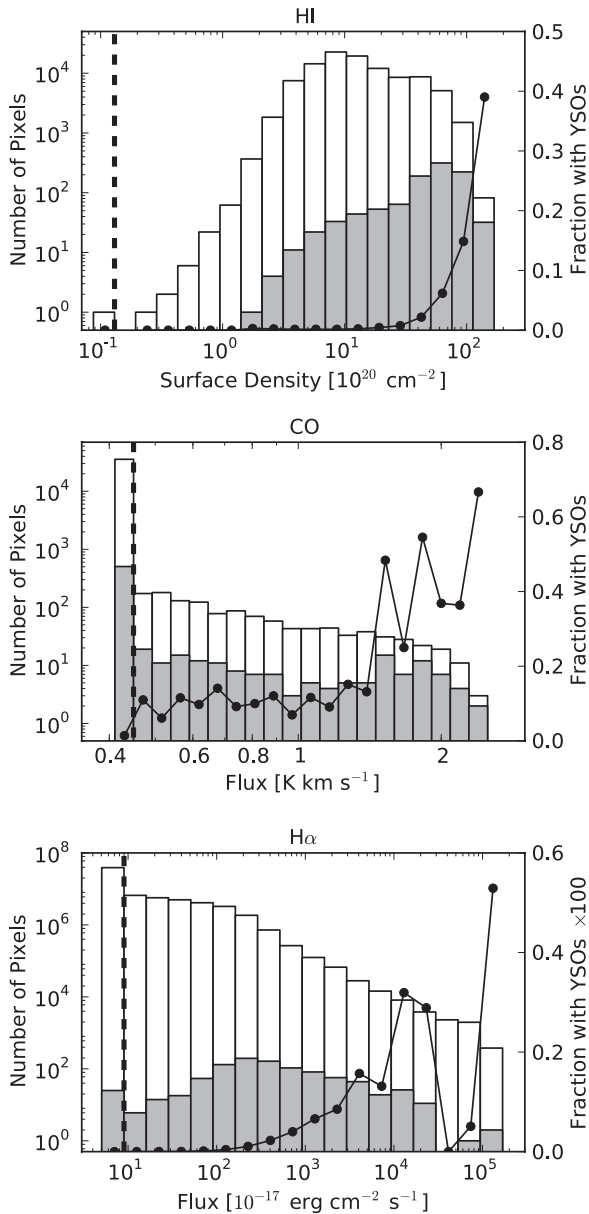


**Figure 18.** Distribution of the YSO candidates (+) with respect to the CO emission from the NANTEN survey. While in general there is a good correlation between CO emission and YSO candidates, not all CO clouds are associated with YSOs. The CO peak flux is  $2.48 \text{ K km s}^{-1}$ . The CO contours are at the levels of 10%, 20%, 40%, and 80% of the peak. North is up, east to the left.

based on location (e.g., the value of  $H\text{I}$  column density) since we do not want to introduce a bias by making assumptions about where the YSOs should be. However, sources in regions with low column density are more likely to be misidentified background

galaxies than those in high density regions. These isolated YSO candidates especially require a spectroscopic follow-up.

The SMC is very extended in  $H\text{I}$  (embedded in a common envelope with the LMC), so its angular size is not well-defined



**Figure 19.** Pixel value histograms of the H I (top), CO (middle), and H $\alpha$  (bottom) images. In each plot, the histogram shown in white includes pixels from the entire image (except areas outside the IRAC coverage in case of H I and H $\alpha$ ), while the gray color indicates pixels at the location of the YSO candidates. Black dots (associated with the right-hand y-axis) are the result of the division of the gray histogram by the white histogram and indicate the fraction of pixels associated with YSOs in each bin. The  $3\sigma$  sensitivity limit of the observations is marked with a vertical dashed line. The histograms show that YSO candidates are preferentially located in the regions of high H I column densities and high CO and H $\alpha$  intensities.

in H I (e.g., Putman et al. 2003; Brüns et al. 2005). The SMC’s stellar populations also extend to large distances from the SMC center. Nidever et al. (2011) revealed that the metal-poor RGB stars in the SMC extend to radii up to  $10^{\circ}6$ , while Casetti-Dinescu et al. (2012) found that the periphery of the SMC is partly surrounded by the OB-type star candidates out to radii of  $3^{\circ}$ – $4^{\circ}$ . Of our YSO candidates associated with the bar and the wing of the SMC (R.A.  $\lesssim 25^{\circ}$ ), 98.4% lie within  $3^{\circ}$  of the kinematic center of the SMC (R.A.  $\sim 01^{\text{h}}05^{\text{m}}$ , Decl.  $\sim -72^{\circ}25'$ ; Stanimirović et al. 2004) and 98.2% within  $3^{\circ}$  of the optical center of the SMC (R.A.  $\sim 00^{\text{h}}52^{\text{m}}44^{\text{s}}$ , Decl.  $= -72^{\circ}49'42''$ ; Mateo 1998).

The spatial distribution of OB-type star candidates shows the well-populated wing of the SMC, and two branches of sources surrounding the main body of the SMC at the northeast and southwest; however, none of these objects lie in the northwestern periphery (in the opposite direction from the wing). The Casetti-Dinescu et al. (2012) selection criteria for OB-type star candidates are very restrictive and thus their sample is incomplete. A number of OB associations from the Bica et al. (2008) catalog lie in this area; it is not completely devoid of young OB-type stars. Casetti-Dinescu et al. (2012) postulated that the OB-type star candidates surrounding the main body of the SMC may be related to the “Counter Bridge”—a structure predicted by the  $N$ -body simulations of Diaz & Bekki (2012).

The structure of the SMC is irregular (much thicker at the northeast compared to the southwest and at smaller distances) and disturbed. The tidal interaction with the LMC  $\sim 200$  Myr ago had a profound impact on the SMC. The Diaz & Bekki (2012) model predicts that the tidal interaction with the LMC during this recent encounter produced two tidal structures: the Magellanic Bridge and the Counter Bridge. The model shows that the Counter Bridge extends away from the SMC (up to  $\sim 20$  kpc) along the line of sight, and thus is hidden behind the SMC. Due to the fact that the SMC is so extended and distorted by tidal interactions with the LMC and the Milky Way, it may not be so unexplored to find young objects at its periphery.

About 7% of YSO candidates from our sample are associated with observed H I column densities  $< 10^{21} \text{ cm}^{-2}$  (or an H I surface density of  $\sim 8 M_{\odot} \text{ pc}^{-2}$ ), with a minimum H I column density of  $\sim 2 \times 10^{20} \text{ cm}^{-2}$  (or  $\sim 1.6 M_{\odot} \text{ pc}^{-2}$ ). All of these sources, which are brighter and/or redder than background galaxies in the [8.0] versus [4.5]–[8.0] CMD, are located at the outskirts of the SMC. Some models (e.g., Schaye 2004; Krumholz et al. 2009) predict a threshold gas surface density of  $3$ – $10 M_{\odot} \text{ pc}^{-2}$  below which the star formation rate drops precipitously. This agrees with the empirically estimated value of  $\lesssim 10 M_{\odot} \text{ pc}^{-2}$  (e.g., Leroy et al. 2008 and references therein). However, models by Ostriker et al. (2010) argue that in order for a galaxy’s gas to maintain pressure equilibrium, the star formation rate will have a non-zero value, even at low surface densities. This theoretical prediction agrees with the observations showing that in the H I dominated outer parts of the spiral galaxies and throughout dwarf galaxies, the star formation efficiency decreases steadily with increasing radius and decreasing H I column density out to and beyond the galaxies’s optical radii (e.g., Bigiel et al. 2008 and Leroy et al. 2008). In fact, star formation has been detected beyond the optical radii of galaxies with the *Galaxy Evolution Explorer* even at H I column densities lower than those observed in the SMC in the area studied in the present paper (e.g., Ferguson et al. 1998; Thilker et al. 2005; Thilker et al. 2005; Gil de Paz et al. 2005). We find a small—but non-zero—population of YSOs in low surface density gas beyond the optical boundaries of the SMC.

The YSO candidates associated with low H I column densities (as low as  $2$ – $3 \times 10^{20} \text{ cm}^{-2}$ ) were also found in the LMC by Whitney et al. (2008) and Gruendl & Chu (2009). Unlike in the SMC, not all of these sources are located at the outskirts of the LMC; some are scattered throughout the galaxy.

## 7.2. Star Formation Rate

Harris & Zaritsky (2004) studied the spatially resolved star formation history of the SMC based on the MCPS survey covering the main body of the galaxy (see Section 2.2). They

found that over 8.4 Gyr ago, there was an epoch of intense star formation when over 50% of stars in the SMC were formed, followed by a long quiescent epoch between 3 and 8.4 Gyr ago when a relatively few stars were formed. The star formation rate increased 3 Gyr ago and has been continuous since that time. This most recent epoch of star formation is characterized by the SFR of roughly  $0.1 M_{\odot} \text{ yr}^{-1}$  with three episodes of enhanced star formation at 2.5, 0.4, and 0.06 Gyr. Since the SFR is calculated in large age bins, the variations of the SFR with time could be significant, but undetected.

The youngest age bin analyzed by Harris & Zaritsky (2004) corresponds to  $\sim 4\text{--}5.3$  Myr. For this age bin, they calculated the SFR in 351 individual subregions in the SMC for two metallicities:  $Z = 0.008$  and  $0.004$ . They estimated a global SFR of  $< 0.01 M_{\odot} \text{ yr}^{-1}$  between  $\sim 4$  Myr and 5.3 Myr by summing the SFRs for two metallicities and all subregions.

Other recent measurements of the global SFR in the SMC were obtained by Wilke et al. (2004) and Bolatto et al. (2011). Bolatto et al. (2011) derived the SFR of  $0.037 M_{\odot} \text{ yr}^{-1}$  based on the  $H\alpha$  and  $24 \mu\text{m}$  emission. The  $H\alpha$  emission traces unobscured UV photons from the young massive stars, while  $24 \mu\text{m}$  emission allows the recovery of UV photons that are absorbed by dust and re-emitted in the IR. Based on the study of the far-IR emission, Wilke et al. (2004) calculated the global SFR of  $\sim 0.05 M_{\odot} \text{ yr}^{-1}$ , assuming that only  $\sim 30\%$  of the bolometric luminosity of young stars is absorbed by dust and re-radiated in the far-IR.

Our work on the resolved YSO population in the SMC allows us to estimate the global SFR directly based on YSO counting. The method of calculating SFR by doing a YSO inventory has been used successfully in Galactic molecular clouds (e.g., Lada et al. 2010) and in the Carina Complex (Povich et al. 2011). Although incomplete, our YSO candidate list can provide a good first order approximation of the SFR, as in Whitney et al. (2008) for the whole LMC and Carlson et al. (2012) for individual star-forming regions in the LMC.

We calculate SFR by estimating the total mass of YSO candidates and dividing it by the YSOs' approximate lifetime. Based on stellar masses and luminosities derived from SED fitting, we construct the mass and luminosity functions for high-reliability Stage I YSO candidates. We base SFR calculations on Stage I YSO candidates only to consider a population of sources roughly at the same age and lifetime. Figure 20 shows that the mass function is incomplete at both low-mass and high-mass end. At the low-mass end, our sample is incomplete due to sensitivity limits and our selection criteria that excluded regions in the color–magnitude space where low-mass YSOs overlap with background galaxies. At the high-mass end, getting a complete census of massive protostars is difficult because they evolve faster than lower-mass YSOs and thus have shorter lifetimes. Due to their short lifetimes, massive protostars are predominantly observed in clusters that may be detected as either extended sources in IRAC and MIPS  $24 \mu\text{m}$  bands and are therefore not included in this study, or single sources that we would identify as massive YSOs. Chen et al. (2010) found that if a *Spitzer* YSO is a multiple system with dominant (massive) components at similar evolutionary stages, model fits to the integrated SED would overestimate the mass of individual components. As a consequence, our sample of YSO candidates may be complete up to the higher-mass YSOs than that suggested by the mass histogram ( $\sim 10 M_{\odot}$ ). We miss massive YSOs that are too young to be detected with *Spitzer* at  $24 \mu\text{m}$ . Stage 0/I YSOs can be identified at longer wavelengths

with the *Herschel Space Observatory* (SewiŁo et al. 2010; J. P. Seale et al., in preparation).

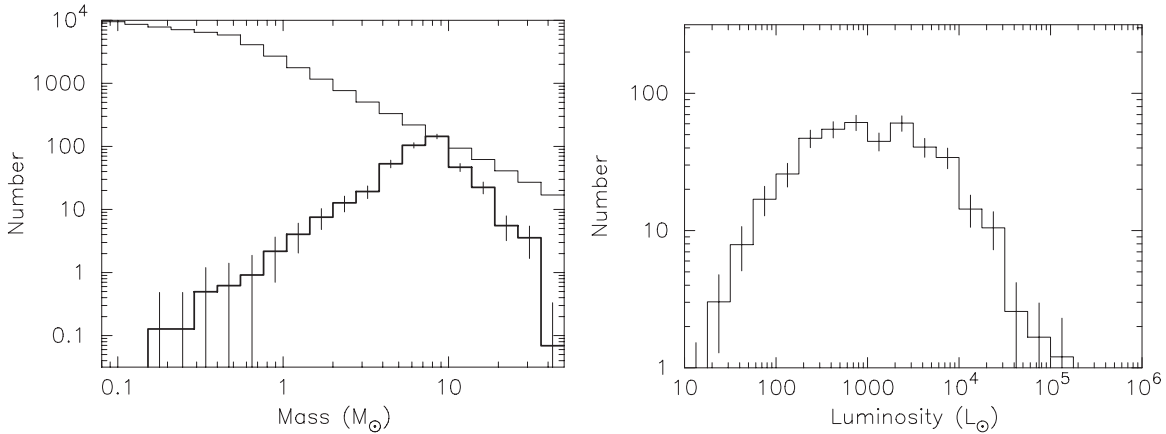
To estimate the total mass, we follow the method developed by Whitney et al. (2008). We match the Kroupa (2001) initial mass function (IMF) to the peak of the observed mass distribution over the mass range where we expect our sample to be approximately complete ( $\sim 7\text{--}10 M_{\odot}$ ). The Kroupa (2001) stellar IMF ( $\zeta(M_{\star}) \propto M_{\star}^{-\alpha}$ ) is a standard two-part function where  $\alpha = 1.3$  for  $0.08 \leq M_{\star}/M_{\odot} < 0.5$  and  $\alpha = 2.3$  for  $M_{\star}/M_{\odot} \geq 0.5$ . Integrating under the IMF over the mass range of  $0.08\text{--}50 M_{\odot}$  results in a lower limit (due to incompleteness) on the total YSO mass. We find a lower limit of  $\sim 31,500 M_{\odot}$  for YSOs throughout the SMC.

The SFR can be estimated by dividing the total mass by the YSOs' lifetime. Evans et al. (2009) used the data from the c2d *Spitzer* Legacy Project for five large, nearby molecular clouds. They based their calculations on a total of 1024 YSOs (the previous studies in individual molecular clouds used 50–100 sources). The uncertainties from small number statistics are decreased to less than 10% effect. Evans et al. (2009) determined a Class I lifetime to be 0.54 Myr (or 0.44 Myr if the photometry is corrected for estimated extinction), assuming the lifetime of Class II phase to be 2 Myr. The vast majority of YSOs from Evans et al. (2009) are low-luminosity YSOs; only  $\sim 50$  sources have luminosities larger than  $5 L_{\odot}$ . Mottram et al. (2011) estimated timescales of massive YSOs (MYSOs) based on the data from the Red MSX Source (RMS) survey (Hoare et al. 2005). They found that the MYSO phase has a duration ranging from 0.4 Myr for  $10^4 L_{\odot}$  to  $\sim 0.07$  Myr at  $10^5 L_{\odot}$ .

We calculate SFRs for two possible Stage I lifetimes: 0.54 Myr and 0.07 Myr, assuming constant star formation over this time. The adoption of a 0.07 Myr lifetime from Mottram et al. (2011) would imply that all sources have  $L \sim 10^5 L_{\odot}$  and is included strictly as a limit. The 0.54 Myr lifetime from Evans et al. (2009) on the other hand is most appropriate for sources with masses well below the  $\sim 7\text{--}10 M_{\odot}$  regime where our study is most complete. These lifetime estimates (and thus SFRs) are very uncertain. The resulting lower limit of SFR is  $0.058 M_{\odot} \text{ yr}^{-1}$  and  $0.45 M_{\odot} \text{ yr}^{-1}$  for lifetimes of 0.54 Myr and 0.07 Myr, respectively. Both results are reasonably consistent with earlier estimates of global SFRs in the SMC within the uncertainties, e.g., due to different tracers used, differences in theoretical and empirical assumptions.

## 8. SUMMARY AND FINAL REMARKS

We identify approximately 1000 YSO candidates and 40 YSO candidate clusters in the SMC using SAGE-SMC catalogs ( $3.6\text{--}24 \mu\text{m}$ ) and images ( $3.6\text{--}70 \mu\text{m}$ ) combined with *JHK<sub>s</sub>* catalogs and images from the IRSF and 2MASS surveys, and optical photometry from the MCPS (*UBVI*) and OGLE-III (*VI*) surveys. Our method of YSO identification incorporates color–magnitude cuts using five CMDs combining all IRAC and MIPS  $24 \mu\text{m}$  bands, detailed visual source inspection using images over a wide wavelength range, and the SED fitting with YSO models. We develop a method to calculate a measure of our confidence that the source is not a non-YSO “contaminant,” but rather a true YSO. This CMD score is a number between 0 and  $\sim 4$ , calculated using the source's location in the color–magnitude space with respect to the non-YSO populations (AGBs, RGBs, background galaxies, and massive stars) in five CMDs used for the initial source selection and scaled by the fraction of CMDs that can be constructed using the source's photometry. We find that spectroscopically confirmed



**Figure 20.** Distribution of stellar mass (left) and luminosity (right) for the YSO candidates. Only high-reliability, Stage I YSO candidates well-fit with the YSO models are included in the calculations. The thin black line in the plot on the left indicates the Kroupa (2001) IMF.

YSOs have high scores, and non-YSOs have low scores. We also show that the Whitney et al. (2008) and Gruendl & Chu (2009) YSO candidates in the LMC have high scores, demonstrating that scores are good indicators of the sources’ classification as YSO candidates. The CMD scores can be used in other galaxies when high-resolution multi-wavelength data resolving stellar populations become available, allowing a quick identification of the most reliable young objects for detailed studies.

Based on the CMD scores and the results of the SED fitting with YSO models, we define two classes of YSO candidates. High-reliability YSO candidates include (1) all sources with high scores ( $>1$ ), and (2) sources with low scores, but well-fit with YSO models. Sources that fulfill these criteria, but have matches in non-YSO catalogs (e.g., evolved stars, PNe, or massive stars; see Section 6.3; Appendix C) are excluded from the high-reliability list. Since the classification of these sources is not definite and there is a possibility of a positional mismatch, we move them to the list of Possible YSO candidates rather than removing them from our list. Possible YSO candidates have low-scores and poor YSO model fits; they do, however, fulfill at least one color–magnitude criterion and their environment supports their young nature. The low score may be the result of incomplete photometry and the SED fitting may be affected by the PAH emission, source confusion or erroneous data point.

Since PAHs contribute to the emission in the  $3.6\ \mu\text{m}$ ,  $5.8\ \mu\text{m}$ , and  $8.0\ \mu\text{m}$  bands, while the  $4.5\ \mu\text{m}$  band remains unaffected, the SED of a source with PAH emission shows a characteristic dip at  $4.5\ \mu\text{m}$ . The current YSO models do not incorporate PAH emission, making SED fitting for sources with PAHs uncertain. We confirm that YSO candidates with PAHs occupy a well-defined region in the  $[3.6]\text{--}[4.5]$  and  $[4.5]\text{--}[5.8]$  CCD, clearly separated from sources with no PAH emission which are coincident with the YSO models. We call these two groups of sources Population P and Population N, respectively. A detailed analysis of these two populations of YSO candidates is beyond the scope of this work and will be discussed in a separate paper.

We provide estimated physical parameters for the YSO candidates well-fit with the YSO models—stellar mass, luminosity, temperature, envelope mass, and disk mass. Good fits indicate that a source’s SED is consistent with YSO model SEDs but do not provide a definite classification. A new YSO model grid and updated fitting tool are under development. The new grid will include half a million models for 10 inclinations, i.e.,  $\sim 5$  million aperture dependent SEDs. These models will be defined in terms of stellar radius and temperature and envelope

density, i.e., parameters that impact the SED directly rather than in terms of model-dependent quantities, e.g., stellar mass and age (derived using pre-main sequence evolutionary tracks), or envelope infall rate (determined because it is directly related to the density profile through the assumption of a free-fall rotational collapse model). This new model grid will not incorporate PAHs and external radiation field. However, the Robitaille (2011) and Whitney et al. (2013) radiation transfer codes include those effects and can be used to model individual sources.

We compare the position of the YSO candidates in the SMC to the distribution of the H I, H $\alpha$ , and CO gas. We find a good spatial correlation between YSO candidates and these gas tracers, which we quantify by comparing histograms of the CO and H $\alpha$  intensities, and H I column density of the pixels in the corresponding images associated with our YSO candidates. Based on the data from the NANTEN CO survey, we find that a significant fraction (86%) of SMC star formation occurs in CO-free or CO-dim gas. However, more sensitive CO observations at higher spatial resolution may detect CO emission at locations where it remained undetected by the NANTEN survey.

The YSO candidates selected based on photometric data needs to be confirmed with spectroscopic observations. While over 300 LMC YSOs are confirmed spectroscopically to date (van Loon et al. 2005; Shimonishi et al. 2008; Oliveira et al. 2009; Seale et al. 2009; van Loon et al. 2010b; Woods et al. 2011), only 33 spectra of massive YSOs in the SMC have been reported in the literature (van Loon et al. 2008, 2010a; Oliveira et al. 2011, 2013). Besides the identification of sources, spectra bring a wealth of information on ice chemistry, excitation conditions, and dust properties in the envelopes of the YSOs. Our list of high-reliability YSO candidates in the SMC will be a valuable source of targets for spectroscopic observations.

Our studies of the YSO candidates in the Magellanic Clouds are laying the groundwork for future studies with the *James Webb Space Telescope (JWST)*. When available, *JWST* instruments will provide great tools to study YSOs. With the Near-Infrared Camera (NIRCam) and the Mid-Infrared Instrument (MIRI) it will be possible to obtain the images and photometry in the wavelength range  $0.6\text{--}5.0\ \mu\text{m}$  (NIRCam) and  $5.0\text{--}27\ \mu\text{m}$  (MIRI). The greatly improved resolution of *JWST* ( $0''.068$  at  $2\ \mu\text{m}$ ) in comparison with *Spitzer* SAGE and SAGE-SMC data ( $\sim 2''$  at IRAC bands) will allow us to study YSO environments in much greater detail. With multiple filters, a well-sampled SED can be constructed for each source. The multi-object Near-Infrared Spectrograph (NIRSpec;  $0.6\text{--}5\ \mu\text{m}$ ) will allow the

efficient collection of spectra for a large number of YSOs. *JWST* will provide the improved angular resolution and sensitivity to study fainter sources in the Magellanic Clouds and will allow us to study resolved stellar populations in more distant galaxies in the Local Group using strategies pioneered for the Magellanic Clouds with *Spitzer* SAGE and SAGE-SMC data.

We thank the anonymous referee for comments and suggestions that improved the clarity of the paper. M.S. acknowledges financial support from the NASA ADAP award NNX11AG50G. This paper is based largely on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.

## APPENDIX A

### NOTES ON THE RELEVANT SAGE-SMC DATA PROCESSING

A detailed description of the SAGE-SMC data products can be found in the documentation for the SAGE-SMC survey available through the *Spitzer* Science Center (Gordon et al. 2011a) and in Gordon et al. (2011b).

Our primary photometry comes from the IRAC SMP (Single Frame + Mosaic Photometry) Archive. This is a combination of mosaic photometry and the single frame photometry Epoch 0+1+2 Archive, where Epoch 0 is the S<sup>3</sup>MC data (re-processed by the SAGE-SMC IRAC pipeline), and Epochs 1 and 2 are two epochs of the SAGE-SMC survey. The single frame photometry Epoch 0+1+2 Archive was derived from doing photometry on individual IRAC frames (single-epoch images), and then doing an error-weighted average of those results for each band. The mosaic photometry was done on the mosaiced images (combined Epochs 0, 1, and 2). The mosaic photometry results were merged with the Epoch 0+1+2 single frame photometry results and the combined *JHK<sub>s</sub>* (1.2, 1.6, and 2.2  $\mu\text{m}$ ) data from the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006) and the 2MASS 6X Deep Point Source Catalog (6X2MASS; Cutri & 2MASS Team 2004). Generally, single frame photometry was used for the brighter sources and mosaic photometry was used at the fainter end.

The S<sup>3</sup>MC image data were included in the SAGE-SMC mosaic image to provide the highest quality image for that region. However, photometry was done in an automated process that sub-tiles the mosaic image into smaller workable images. Due to the irregular shape of the S<sup>3</sup>MC area, and the much larger overall size of the full SAGE-SMC area, the criteria for source detection were set at the typical SAGE-SMC image exposure and not at the higher exposure time for the limited S<sup>3</sup>MC area. This led to a less deep list of sources in the S<sup>3</sup>MC area than optimally possible, but treated the full SAGE-SMC area in the same systematic fashion.

## APPENDIX B

### DATA FROM THE POINT SOURCE CATALOGS

To provide photometry for the SAGE-SMC IRAC catalog sources (already matched to MIPS 24  $\mu\text{m}$  catalog) over a broader wavelength range (from optical to 24  $\mu\text{m}$ ), the IRAC and MIPS catalogs were cross-matched to the following optical and near-IR point source catalogs (see Section 2.2).

1. The Magellanic Clouds Photometric Survey (MCPS; Zaritsky et al. 2002), covering a  $4^\circ \times 4.5^\circ$  region in the SMC in the *UBVI* filters. The MCPS catalog is 50% complete to  $m_V = 21\text{--}22$  mag, depending on the local crowding conditions. The catalog includes at least *B* and *V* photometry for over 6 million SMC stars. Typical seeing and pixel scale are  $\sim 1''.5/\text{pixel}$  and  $0''.7/\text{pixel}$ , respectively.
2. The OGLE-III Photometric Maps of the Small Magellanic Cloud containing *V* and *I* photometry of about 6.2 million stars from 41 OGLE-III fields in the SMC, covering about  $14 \text{ deg}^2$  (Udalski et al. 2008). The pixel scale is  $0''.26/\text{pixel}$ . The median seeing of the *I*-band images is  $1''.2$ . Completeness of the photometry is  $I \sim 21$  mag and  $V \sim 21.5$  mag.
3. Infrared Survey Facility (IRSF) Magellanic Clouds Point Source Survey (Kato et al. 2007), covering an  $11 \text{ deg}^2$  area of the SMC and a  $4 \text{ deg}^2$  area of the Magellanic Bridge in the *J*, *H*, *K<sub>s</sub>* filters. The IRSF catalog contains  $\sim 2.8$  million sources in the SMC and  $\sim 435,000$  in the Magellanic Bridge with the  $10\sigma$  limiting magnitudes of 18.8, 17.8, and 16.6 mag at *J*, *H*, and *K<sub>s</sub>*, respectively. IRSF has a pixel scale of  $0''.45/\text{pixel}$  and the average seeing of  $1''.3$ ,  $1''.2$ ,  $1''.1$  in the *JHK<sub>s</sub>* bands, respectively. The IRSF catalog has a higher resolution and is significantly deeper than the 2MASS catalog. An average seeing for 2MASS is  $2''.5$  and limiting magnitudes are 15.8, 15.1 and 14.3 mag at *J*, *H*, and *K<sub>s</sub>*, respectively. 2MASS has a pixel scale of  $2''/\text{pixel}$ .

The 2MASS/6X2MASS data are the integral part of the SAGE-SMC IRAC catalog (Gordon et al. 2011a, 2011b). The 2MASS/6X2MASS and all *Spitzer* magnitudes and fluxes were extracted from the SAGE-SMC catalogs (see Section 4). The MCPS, OGLE-III, and IRSF catalogs provide magnitudes only. We derived fluxes using the following zero magnitude fluxes: (1) 1790, 4063, 3636, and 2416 Jy for the MCPS bands *U*, *B*, *V*, and *I*, respectively; (2) 3636 and 2416 Jy for the OGLE-III bands *V*, *I*; and (3) 1594, 1024, and 666.7 Jy for the IRSF bands *J*, *H*, and *K<sub>s</sub>*, respectively. All magnitudes are calibrated relative to Vega.

In Table 8, we present photometry for a wavelength range from *U* to MIPS 24  $\mu\text{m}$  for the YSO candidates from all the surveys used in this work. The photometry is listed separately for each survey as opposed to Tables 1 and 2 where the photometry from different surveys is combined for each band if more than one flux measurement is available in a given band (e.g., *J*-band flux from 2MASS and IRSF; see Section 2.2). Table 8 includes magnitudes and fluxes for the following bands.

1. *UBVI* from The Magellanic Clouds Photometric Survey (MCPS; Zaritsky et al. 2002)
2. *VI* from The Optical Gravitation Lensing Experiment (OGLE-III; Udalski et al. 2008)
3. *JHK<sub>s</sub>* from Infrared Survey Facility (IRSF) Magellanic Clouds Point Source Survey (Kato et al. 2007)
4. *JHK<sub>s</sub>* from Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006) or 2MASS 6X Deep Point Source Catalog (6X2MASS; Cutri & 2MASS Team 2004)
5. IRAC 3.6, 4.5, 5.8, and 8.0, and MIPS 24  $\mu\text{m}$  from the *Spitzer* survey Surveying the Agents of Galaxy Evolution in the Tidally Stripped, Low Metallicity Small Magellanic Cloud (SAGE-SMC; Gordon et al. 2011b)

Each magnitude and flux measurement is provided with the uncertainty except for the OGLE-III *V*- and *I*-band data

**Table 8**  
Data from the MCPS, OGLE-III, IRSF, 2MASS, and *Spitzer* SAGE-SMC Point Source Catalogs

Column	Name	Description	Null
1	Source Name	The identification number of the YSO candidates	...
2	R.A.(J2000)	Right Ascension, J2000 (deg)	...
3	Decl.(J2000)	Declination, J2000 (deg)	...
4–5	OGLEID_map, OGLEID_DBno	a filename of the OGLE-III map of the field the source is located in and the identification number in the OGLE-III database; together they allow a unique identification within OGLE	null
6	IRSFID	IRSF source designation	null
7	2MASSID	identification number for a 2MASS source (cntr in the IRSA/GATOR)	null
8	IRACDesignation	SAGE-SMC IRAC SMP source name: “SSTISAGEMA JHHMMSS.SS±DDMMSS.S”	...
9	MIPSDesignation	SAGE-SMC MIPS 24 $\mu$ m source name: “SSTM1SAGE1 JHHMMSS.SS±DDMMSS.S” (Epoch 1) or “SSTM1SAGE1 JHHMMSS.SS±DDMMSS.S” (Epoch 2)	null
10–17	magi_MCPS,dmagi_MCPS	<i>UBVI</i> ( $i = 1–4$ ) magnitudes and magnitude uncertainties from MCPS	99.999, 99.999
18–19	magi_OGLE	<i>VI</i> ( $i = 1, 2$ ) magnitudes from OGLE-III (mag)	99.999, 99.999
20–25	magi_2MASS,dmagi_2MASS	<i>JHK<sub>s</sub></i> ( $i = 1–3$ ) magnitudes and magnitude uncertainties from 2MASS/6X2MASS (mag)	99.999, 99.999
26–31	magi_IRSF,dmagi_IRSF	<i>JHK<sub>s</sub></i> ( $i = 1–3$ ) magnitudes and magnitude uncertainties from IRSF (mag)	99.999, 99.999
32–39	magi, dmagi	SAGE-SMC IRAC 3.6, 4.5, 5.8, and 8.0 $\mu$ m ( $i = 1–4$ ) magnitudes and magnitude uncertainties (mag)	99.999, 99.999
40–41	mag24, dmag24	SAGE-SMC MIPS 24 $\mu$ m magnitude and magnitude uncertainty (mag)	99.999, 99.999
42–49	fluxi_MCPS,dfluxi_MCPS	<i>UBVI</i> ( $i = 1–4$ ) fluxes and flux uncertainties from MCPS (mJy)	–999.9, –999.9
50–51	fluxi_OGLE	<i>VI</i> ( $i = 1, 2$ ) fluxes from OGLE-III (mJy)	–999.9, –999.9
52–57	fluxi_2MASS,dfluxi_2MASS	<i>JHK<sub>s</sub></i> ( $i = 1–3$ ) fluxes and flux uncertainties from 2MASS/6X2MASS (mJy)	–999.9, –999.9
58–63	fluxi_IRSF,dfluxi_IRSF	<i>JHK<sub>s</sub></i> ( $i = 1–3$ ) fluxes and flux uncertainties from IRSF (mJy)	–999.9, –999.9
64–71	fluxi, dfluxi	SAGE-SMC IRAC 3.6, 4.5, 5.8, and 8.0 $\mu$ m ( $i = 1–4$ ) fluxes and flux uncertainties (mJy)	–999.9, –999.9
72–73	flux24, dflux24	SAGE-SMC MIPS 24 $\mu$ m flux and flux uncertainty (mJy)	–999.9, –999.9
74	IracMipsDist	a distance between IRAC and MIPS 24 $\mu$ m matching sources (arcsec)	–9.999
75–76	IracMCPSDist,MipsMCPSDist	IRAC-MCPS and (MIPS 24 $\mu$ m)-MCPS distances (arcsec)	–9.999
77–78	IracOGLEDist,MipsOGLEDist	IRAC-OGLEIII and (MIPS 24 $\mu$ m)-OGLEIII distances (arcsec)	–9.999
79–80	IracIRSFdist,MipsIRSFdist	IRAC-IRSF and (MIPS 24 $\mu$ m)-IRSF distances (arcsec)	–9.999

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

(uncertainties were not provided in the OGLE-III catalog). We provide source names from the original catalogs for all the surveys except for MCPS that does not include source identification numbers. Table 8 also includes distances between sources from different surveys (IRAC-MIPS24, IRAC-MCPS, MIPS24-MCPS, IRAC-OGLEIII, MIPS24-OGLEIII, IRAC-IRSF, MIPS24-IRSF) that are considered to be a match—distances were required to be  $\leq 1''$  (see Section 2.2 for details).

Table 8 can be found in its entirety in the online journal.

## APPENDIX C

### LISTS OF NON-YSOs

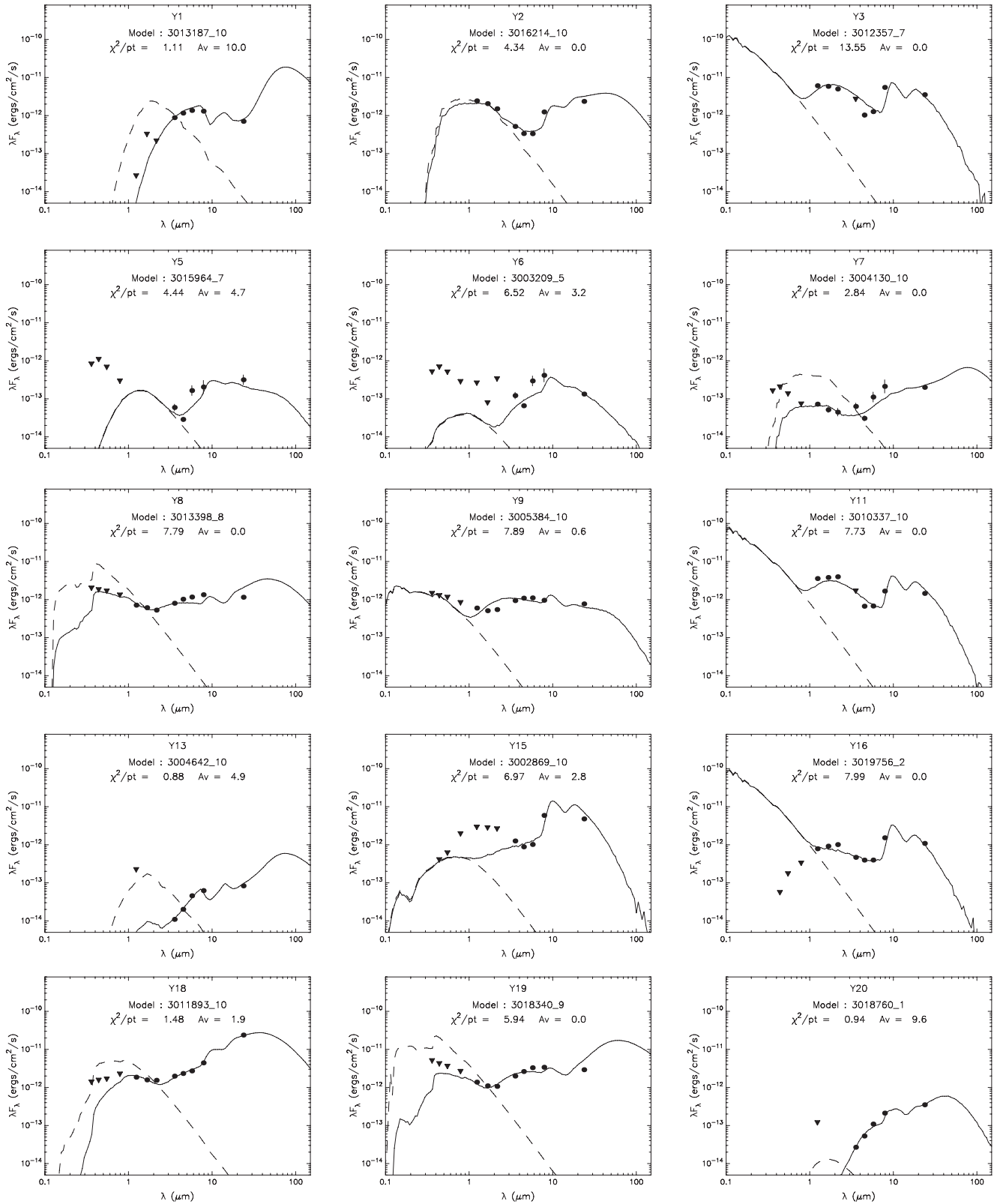
We use the data from non-YSO catalogs from literature to identify regions in the color–magnitude space where YSOs can be confused with other categories of sources. This information will allow us to select the initial YSO candidates list that is as free of contaminants as possible. We will also use the non-YSO catalogs to assess the residual contamination from non-YSOs on our final list of YSO candidates.

The following data from literature were used.

- Asymptotic giant branch stars (AGBs): C-rich, O-rich, extreme (x-), and aO- (anomalous O-rich). AGBs were identified by Boyer et al. (2011) using near-IR (2MASS/IRSF) photometry for C-rich and O-rich AGBs, and a combination of near-IR and mid-IR (SAGE-SMC) photometry for x-AGBs and aO-AGBs. The selection of the AGB star candidates was based on color–magnitude cuts.
- Massive stars: O stars, early (B0–B2.5) and late (B3–B9) B stars (the majority of these have supergiant or giant luminosity classifications), spectral type A, F and G type (AFG) supergiants, K and M red supergiants (RSGs), Wolf–Rayet

(WR) stars, B[e] supergiants (sgB[e]), confirmed luminous blue variables (LBVs), and Be/X-ray binaries. The mid-IR data were extracted from the Bonanos et al. (2010) photometric catalog of spectroscopically confirmed massive stars covering a wavelength range 0.3 – 24  $\mu$ m. The Bonanos et al. (2010) catalog was made based primarily on the SAGE-SMC data.

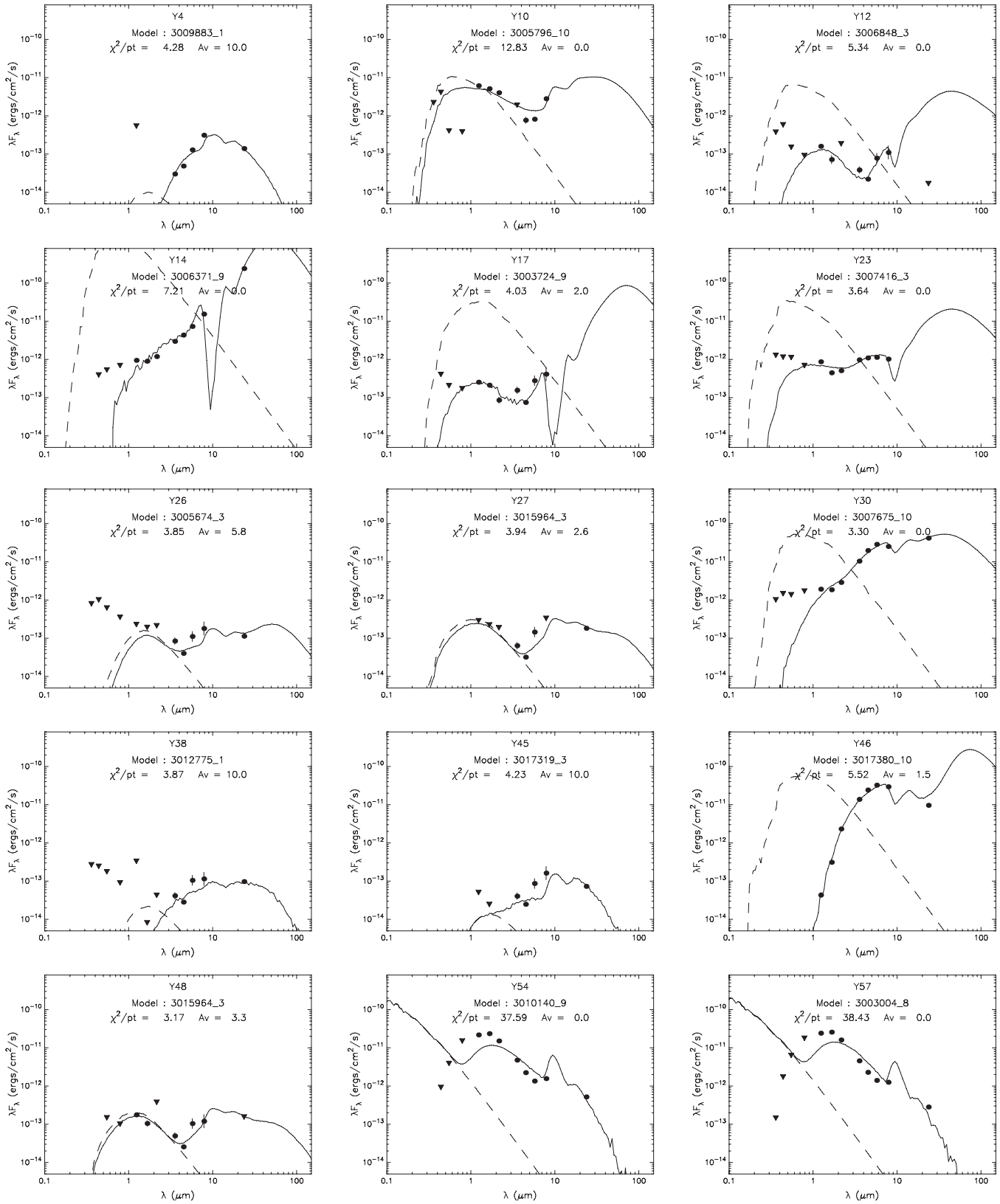
- Planetary Nebulae (PNe): We matched the sources from the catalog of PNe (G. Jacoby 2009, private communication) to the SAGE-SMC IRAC SMP Archive to extract their mid-IR fluxes. We required a distance between PNe and their SAGE-SMC counterparts to be  $\leq 1''$ .
- H $\alpha$  emission line stars and small nebulae: Meyssonnier & Azzopardi (1993) catalog of H $\alpha$  emission line stars and small nebulae in the SMC is based on the H $\alpha$  + [N II] survey with the 0.9 m Curtis Schmidt telescope at Cerro Tololo Inter-American Observatory under seeing better than 2''. The catalog contains 1898 emission-line objects with 62 identified as PNe (11 are very low excitation (VLE)) and 81 as compact H II regions.
- Dusty OB stars: The catalog of 125 dusty OB stars with optical spectra from Sheets et al. (2013). This is a subset of 193 objects that were originally identified in the S<sup>3</sup>MC survey as showing an excess emission at 24  $\mu$ m (Bolatto et al. 2007). Based on the optical spectroscopy (a wavelength range from  $\sim 385$  to 870 nm), Sheets et al. (2013) classified 87 out of 125 sources as normal stars without line emission, mostly late-O/early-B type main-sequence stars. The remaining sources include emission line stars (17), possible emission line stars (7), objects with forbidden-line emission (5), as well as 2 PNe, a YSO, and 2 X-ray binaries.



**Figure 21.** SEDs and YSO fits for high-reliability YSO candidates. The symbols and lines are the same as in Figure 9. (The complete figure set (50 images) is available in the online journal.)

6. Background galaxies: We use the IRAC and MIPS 24  $\mu$ m catalogs from The *Spitzer* Wide-area InfraRed Extragalactic Survey (SWIRE; Lonsdale et al. 2003). SWIRE surveyed  $\sim$ 49 square degrees area in 6 fields. We selected the

‘Lockman Hole’ field (11.1 square degrees) that is the furthest away from the Galactic plane. We removed sources that were well-fit with the stellar photosphere models (see Section 5.2 for a discussion of the SED fitting). The



**Figure 22.** SEDs and YSO fits for possible YSO candidates. The symbols and lines are the same as in Figure 9. (The complete figure set (16 images) is available in the online journal.)



remaining sources are most likely background galaxies; we show a randomly selected subsample of this list ( $\sim 10\%$ ) in the plots (Figure 2).

7. X-ray point sources associated with radio emission: Sturm et al. (2013) catalog of 88 XMM-Newton X-ray sources associated with unique radio counterparts. Out of 88 sources, 75 are classified as AGNs. The remaining sources include eight galaxies, one confirmed quasar, one foreground star, two radio sources within clusters of galaxies, and one pulsar wind nebula candidate.

## APPENDIX D

### CMD SCORE CALCULATION

The CMD score is calculated by comparing a source's position in color–magnitude space to the known positions of other types of non-YSO objects. The score makes the basic assumption that a source is more likely to be a certain type of object (e.g., evolved star or background galaxy) if it is similar in color and magnitude to objects of that type. Practically, the process involves identifying how similar a source's color and magnitude is to objects that are not YSOs. The most securely classified YSOs are those that are most dissimilar in color and magnitude from non-YSOs. To calculate the score, we first must identify the positions in color–magnitude space of the possible non-YSO contaminating source types. Because they occupy similar regions of the CMD as YSOs, we consider evolved stars (AGBs and red giant branch stars—RGBs), background galaxies, and massive stars as the most likely contaminants to the candidate YSO list. We obtained large catalogs of photometry for each of these object types from the literature (AGBs and RGBs; Boyer et al. 2011; galaxies, SWIRE/Lonsdale et al. 2003; massive stars, Bonanos et al. 2010). Through the calculation of the score, this photometry of non-YSOs is then compared to that of YSO candidates. The scores are calculated using the catalog photometry for YSO candidates when available with aperture photometry filling in the missing bands. Each YSO candidate's CMD score is computed separately as follows.

1. For a particular CMD used to identify YSO candidates for inspection ( $\text{CMD}_i$ , where  $i = 1, 2, 3, 4, 5$  denotes the 5 CMDs used), a circle is centered at the photometric position of the source, and the smallest radius that contains 5% of a particular non-YSO type is determined ( $R_{i,j}$ , where  $j = \text{AGB}, \text{RGB}, \text{galaxy}, \text{massive star}$  denotes the non-YSO type). The catalogs of non-YSO photometry contain large numbers of sources (9124 AGBs, 135437 RGBs, 323043 background galaxies, and 3654 massive stars), such that 5% is large enough to contain a statistically significant number of sources but small enough to allow enough differentiation between sources within and outside of non-YSO regions in a CMD. Large radii signify sources isolated in color–magnitude space from non-YSOs, while small numbers signify sources in non-YSO-dominated areas.
2. Step 1 is repeated for each  $\text{CMD}_i$ , and the average value of  $R_{i,j}$  for all available CMDs,  $R_j$ , is calculated:  $R_j = (\sum_{i=0}^5 R_{i,j})/5$
3. Steps 1–2 are repeated for each non-YSO species (AGB, RGB, galaxy, and massive stars) so that each source has an associated  $R_{\text{AGB}}, R_{\text{RGB}}, R_{\text{galaxy}}$ , and  $R_{\text{massive}}$ .
4. The smallest  $R_j$  from Step 3 is adopted as the radius,  $R$ , particular to that source (a measure of the distance in the color–magnitude space from the source to the most similar non-YSO population).

5.  $R$ , the result of steps 1–4 is a measure of where a source is located in color–magnitude space relative to non-YSOs, but it does not take into account the certitude that the source is a YSO. Recall that not all sources are detected in every *Spitzer* waveband, meaning some sources' classifications are based upon less data than others'. If a source only has enough photometry to produce 2 CMDs, we are less certain of its classification than for a source with 5 CMDs. In other words, a source with all 5 CMDs has more photometric evidence supporting its classification. To account for this,  $R$  is scaled by the fraction of CMDs the source's photometry can produce. A source with all 5 CMDs is scaled by a factor of  $5/5 = 1.0$ , those with 4 are scaled by  $4/5 = 0.8$ , and so forth. Theoretically, a source with no CMDs would have a  $0/5 = 0.0$  scale factor, meaning we can have no certainty in its classification. This scaled value is the final CMD score and accounts for both the source's proximity in color–magnitude space to non-YSOs and the certitude of its classification.

## APPENDIX E

### THE SEDs AND YSO FITS FOR YSO CANDIDATES

We present the SEDs with the best YSO model fits for 742 high-reliability (Figure 21) and 238 possible (Figure 22; 4 sources were not fitted due to too few data points) YSO candidates. The SED fitting and the fitting results are discussed in Section 5.2.

## REFERENCES

- André, P., Ward-Thompson, D., & Barsony, M. 1993, *ApJ*, 406, 122
- Atlee, D. W., & Gould, A. 2007, *ApJ*, 664, 53
- Bagheri, G., Cioni, M.-R. L., & Napiwotzki, R. 2013, *A&A*, 551, A78
- Bica, E., Bonatto, C., Dutra, C. M., & Santos, J. F. C. 2008, *MNRAS*, 389, 678
- Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, 136, 2846
- Bolatto, A. D., Leroy, A. K., Jameson, K., et al. 2011, *ApJ*, 741, 12
- Bolatto, A. D., Simon, J. D., Stanimirović, S., et al. 2007, *ApJ*, 655, 212
- Bonanos, A. Z., Lennon, D. J., Köhlinger, F., et al. 2010, *AJ*, 140, 416
- Boyer, M. L., Srinivasan, S., van Loon, J. T., et al. 2011, *AJ*, 142, 103
- Brüns, C., Kerp, J., Staveley-Smith, L., et al. 2005, *A&A*, 432, 45
- Carlson, L. R., Sabbi, E., Sirianni, M., et al. 2007, *ApJL*, 665, L109
- Carlson, L. R., Sewillo, M., Meixner, M., Romita, K. A., & Lawton, B. 2012, *A&A*, 542, A66
- Carlson, L. R., Sewillo, M., Meixner, M., et al. 2011, *ApJ*, 730, 78
- Casetti-Dinescu, D. I., Vieira, K., Girard, T. M., & van Altena, W. F. 2012, *ApJ*, 753, 123
- Chen, C., Chu, Y., Gruendl, R. A., Gordon, K. D., & Heitsch, F. 2009, *ApJ*, 695, 511
- Chen, C.-H. R., Indebetouw, R., Chu, Y.-H., et al. 2010, *ApJ*, 721, 1206
- Cutri, R. M., & 2MASS Team., 2004, *BAAS*, 36, 1487
- Diaz, J. D., & Bekki, K. 2012, *ApJ*, 750, 36
- Evans, N. J., II, Dunham, M. M., Jørgensen, J. K., et al. 2009, *ApJS*, 181, 321
- Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, *ApJS*, 154, 10
- Ferguson, A. M. N., Wyse, R. F. G., Gallagher, J. S., & Hunter, D. A. 1998, *ApJL*, 506, L19
- Gatley, I., Hyland, A. R., & Jones, T. J. 1982, *MNRAS*, 200, 521
- Gil de Paz, A., Madore, B. F., Boissier, S., et al. 2005, *ApJL*, 627, L29
- Gordon, K. D., Babler, B., Block, M., et al. 2011a, “The SAGE-SMC Data Description: Delivery 3” document available at the Spitzer Science Center website: [http://data.spitzer.caltech.edu/popular/sage-smc/20110429\\_enhanced/documentation/](http://data.spitzer.caltech.edu/popular/sage-smc/20110429_enhanced/documentation/)
- Gordon, K. D., Bot, C., Muller, E., et al. 2009, *ApJL*, 690, L76
- Gordon, K. D., Meixner, M., Meade, M. R., et al. 2011b, *AJ*, 142, 102
- Gouliermis, D. A., Quanz, S. P., & Henning, T. 2007, *ApJ*, 665, 306
- Groenewegen, M. A. T., & Blommaert, J. A. D. L. 1998, *A&A*, 332, 25
- Gruendl, R. A., & Chu, Y. 2009, *ApJS*, 184, 172
- Harris, J., & Zaritsky, D. 2004, *AJ*, 127, 1531
- Haschke, R., Grebel, E. K., & Duffau, S. 2012, *AJ*, 144, 107
- Henize, K. G. 1956, *ApJS*, 2, 315

- Hilditch, R. W., Howarth, I. D., & Harries, T. J. 2005, *MNRAS*, **357**, 304
- Hoare, M. G., Lumsden, S. L., Oudmaijer, R. D., et al. 2005, in IAU Symp. 227, *Massive Star Birth: A Crossroads of Astrophysics*, ed. R. Cesaroni, M. Felli, E. Churchwell, & M. Walmsley (Cambridge: Cambridge Univ. Press), 370
- Hummel, W., Szeifert, T., Gässler, W., et al. 1999, *A&A*, **352**, L31
- Kato, D., Nagashima, C., Nagayama, T., et al. 2007, *PASJ*, **59**, 615
- Kozłowski, S., Kochanek, C. S., & Udalski, A. 2011, *ApJS*, **194**, 22
- Kraemer, K. E., Sloan, G. C., Wood, P. R., Price, S. D., & Egan, M. P. 2005, *ApJL*, **631**, L147
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, *ApJ*, **699**, 850
- Lada, C. J. 1987, in IAU Symp. 115, *Star Forming Regions*, ed. M. Peimbert & J. Jugaku (Cambridge: Cambridge Univ. Press), 1
- Lada, C. J., Lombardi, M., & Alves, J. F. 2010, *ApJ*, **724**, 687
- Lee, J.-K., Rolleston, W. R. J., Dufton, P. L., & Ryans, R. S. I. 2005, *A&A*, **429**, 1025
- Leroy, A., Bolatto, A., Stanimirović, S., et al. 2007, *ApJ*, **658**, 1027
- Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, **136**, 2782
- Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, *PASP*, **115**, 897
- Mateo, M. L. 1998, *ARA&A*, **36**, 435
- McGee, R. X., & Newton, L. M. 1986, *PASAu*, **6**, 471
- McGowan, K. E., Coe, M. J., Schurch, M. P. E., et al. 2008, *MNRAS*, **383**, 330
- Meixner, M., Gordon, K. D., Indebetouw, R., et al. 2006, *AJ*, **132**, 2268
- Meyssonnier, N., & Azzopardi, M. 1993, *A&AS*, **102**, 451
- Mizuno, N., Muller, E., Maeda, H., et al. 2006, *ApJL*, **643**, L107
- Mizuno, N., Rubio, M., Mizuno, A., et al. 2001, *PASJ*, **53**, L45
- Mottram, J. C., Hoare, M. G., Davies, B., et al. 2011, *ApJL*, **730**, L33
- Muller, E., Staveley-Smith, L., Zealey, W., & Stanimirović, S. 2003, *MNRAS*, **339**, 105
- Nidever, D. L., Majewski, S. R., Muñoz, R. R., et al. 2011, *ApJL*, **733**, L10
- Oliveira, J. M., van Loon, J. T., Chen, C., et al. 2009, *ApJ*, **707**, 1269
- Oliveira, J. M., van Loon, J. T., Sloan, G. C., et al. 2011, *MNRAS*, **411**, L36
- Oliveira, J. M., van Loon, J. T., Sloan, G. C., et al. 2013, *MNRAS*, **428**, 3001
- Ostriker, E. C., McKee, C. F., & Leroy, A. K. 2010, *ApJ*, **721**, 975
- Povich, M. S., Smith, N., Majewski, S. R., et al. 2011, *ApJS*, **194**, 14
- Prevot, L., Martin, N., Rebeiro, E., Maurice, E., & Rousseau, J. 1983, *A&AS*, **53**, 255
- Putman, M. E., Staveley-Smith, L., Freeman, K. C., Gibson, B. K., & Barnes, D. G. 2003, *ApJ*, **586**, 170
- Rieke, G. H., Young, E. T., Engelbracht, C. W., et al. 2004, *ApJS*, **154**, 25
- Robitaille, T. P. 2008, in ASP Conf. Ser. 387, *Massive Star Formation: Observations Confront Theory*, ed. H. Beuther, H. Linz, & T. Henning (San Francisco, CA: ASP), 290
- Robitaille, T. P. 2011, *A&A*, **536**, A79
- Robitaille, T. P., Whitney, B. A., Indebetouw, R., & Wood, K. 2007, *ApJS*, **169**, 328
- Robitaille, T. P., Whitney, B. A., Indebetouw, R., Wood, K., & Denzmore, P. 2006, *ApJS*, **167**, 256
- Rolleston, W. R. J., Dufton, P. L., McErlean, N. D., & Venn, K. A. 1999, *A&A*, **348**, 728
- Rolleston, W. R. J., Venn, K., Tolstoy, E., & Dufton, P. L. 2003, *A&A*, **400**, 21
- Russell, S. C., & Dopita, M. A. 1992, *ApJ*, **384**, 508
- Schaye, J. 2004, *ApJ*, **609**, 667
- Seale, J. P., Looney, L. W., Chu, Y.-H., et al. 2009, *ApJ*, **699**, 150
- Sewillo, M., Indebetouw, R., Carlson, L. R., et al. 2010, *A&A*, **518**, L73
- Sharpee, B., Stark, M., Pritzl, B., et al. 2002, *AJ*, **123**, 3216
- Sheets, H. A., Bolatto, A. D., van Loon, J. T., et al. 2013, *ApJ*, **771**, 111
- Shimonishi, T., Onaka, T., Kato, D., et al. 2008, *ApJL*, **686**, L99
- Simon, J. D., Bolatto, A. D., Whitney, B. A., et al. 2007, *ApJ*, **669**, 327
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Stanghellini, L., Lee, T.-H., Shaw, R. A., Balick, B., & Villaver, E. 2009, *ApJ*, **702**, 733
- Stanimirović, S., Staveley-Smith, L., Dickey, J. M., Sault, R. J., & Snowden, S. L. 1999, *MNRAS*, **302**, 417
- Stanimirović, S., Staveley-Smith, L., & Jones, P. A. 2004, *ApJ*, **604**, 176
- Sturm, R., Drašković, D., Filipović, M. D., et al. 2013, arXiv:1308.1218
- Thilker, D. A., Bianchi, L., Boissier, S., et al. 2005, *ApJL*, **619**, L79
- Udalski, A., Soszyński, I., Szymański, M. K., et al. 2008, *AcA*, **58**, 329
- van Loon, J. T., Cohen, M., Oliveira, J. M., et al. 2008, *A&A*, **487**, 1055
- van Loon, J. T., Oliveira, J. M., Gordon, K. D., Sloan, G. C., & Engelbracht, C. W. 2010a, *AJ*, **139**, 1553
- van Loon, J. T., Oliveira, J. M., Gordon, K. D., et al. 2010b, *AJ*, **139**, 68
- van Loon, J. T., Oliveira, J. M., Wood, P. R., et al. 2005, *MNRAS*, **364**, L71
- Véron-Cetty, M.-P., & Véron, P. 2010, *A&A*, **518**, A10
- Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, *ApJS*, **154**, 1
- Whitney, B. A., Robitaille, T. P., Bjorkman, J. E., et al. 2013, arXiv:1307.0561
- Whitney, B. A., Sewillo, M., Indebetouw, R., et al. 2008, *AJ*, **136**, 18
- Whitney, B. A., Wood, K., Bjorkman, J. E., & Cohen, M. 2003a, *ApJ*, **598**, 1079
- Whitney, B. A., Wood, K., Bjorkman, J. E., & Wolff, M. J. 2003b, *ApJ*, **591**, 1049
- Wilke, K., Klaas, U., Lemke, D., et al. 2004, *A&A*, **414**, 69
- Wisniewski, J. P., & Bjorkman, K. S. 2006, *ApJ*, **652**, 458
- Woods, P. M., Oliveira, J. M., Kemper, F., et al. 2011, *MNRAS*, **411**, 1597
- Zaritsky, D., Harris, J., Thompson, I. B., Grebel, E. K., & Massey, P. 2002, *AJ*, **123**, 855