THE EVOLUTION OF MASSIVE YOUNG STELLAR OBJECTS IN THE LARGE MAGELLANIC CLOUD. I. IDENTIFICATION AND SPECTRAL CLASSIFICATION

JONATHAN P. SEALE[1], LESLIE W. LOONEY[1], YOU-HUA CHU[1], ROBERT A. GRUENDL[1], BERNHARD BRANDL[2], C.-H. ROSIE CHEN[3], WOLFGANG BRANDNER[4], AND GEOFFREY A. BLAKE[5]

[1] Astronomy Department, University of Illinois, Department of Astronomy, MC-221, 1002 West Green Street, Urbana, IL 61801, USA
[2] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2300 RA Leiden, Netherlands
[3] Astronomy Department, University of Virginia, Department of Astronomy, P.O. Box 400325, Charlottesville, VA 22904-4325, USA
[4] Max-Planck-Institut für Astronomie, Planet & Star Formation, Königstuhl 17, 69117 Heidelberg, Germany
[5] Astronomy Department, California Institute of Technology, Caltech, MC 105-24, 1200 East California Boulevard, Pasadena, CA 91125, USA
Received 2009 February 6; accepted 2009 April 20; published 2009 June 10

ABSTRACT

We present and categorize *Spitzer* infrared spectrometer spectra of 294 objects in the Large Magellanic Cloud (LMC) to create the largest and most complete catalog of massive young stellar object (YSO) spectra in the LMC. Target sources were identified from infrared photometry and multiwavelength images indicative of young, massive stars highly enshrouded in their natal gas and dust clouds. Several objects have been spectroscopically identified as non-YSOs and have features similar to more-evolved stars such as red supergiants, asymptotic giant branch (AGB), and post-AGB stars. Our sample primarily consists of 277 objects we identify as having spectral features indicative of embedded YSOs. The remaining sources are comprised of seven C-rich evolved sources, eight sources dominated by broad silicate emission, and one source with multiple broad emission features. Those with YSO-like spectra show a range of spectral features including polycyclic aromatic hydrocarbon emission, deep silicate absorption, fine-structure lines, and ice absorption features. Based upon the relative strengths of these features, we have classified the YSO candidates into several distinct categories using the widely used statistical procedure known as principal component analysis. We propose that these categories represent a spectrum of evolutionary stages during massive YSO formation. Using our catalog we put statistical constraints on the relative evolutionary timescale of processes involved in massive star formation. We conclude that massive pre-main-sequence stars spend a majority (possibly as high as 90%) of their massive, embedded lives emitting in the UV. Half of the sources in our study have features typical of compact H II regions, suggesting that massive YSOs can create a detectable compact H II region half-way through the formation time present in our sample. This study also provides a check on commonly used source-selection procedures including the use of photometry to identify YSOs. We determine that a high success rate (>95%) of identifying objects with YSO-like spectra can be achieved through careful use of infrared color–magnitude diagrams, spectral energy distributions, and image inspections.

Key words: galaxies: individual (LMC) – infrared: stars – instrumentation: spectrographs – Magellanic Clouds – stars: evolution – stars: formation

Online-only material: extended figure, figure set, machine-readable table

1. INTRODUCTION

During all stages of their lives, stars play an important role in the evolution of their host galaxy. Massive stars, in particular, have a strong influence on galactic properties (e.g., Kennicutt 2005), as they are the principal source of heavy elements and UV radiation. A single star can influence the structure and chemistry of its local interstellar medium (ISM; e.g., van Dishoeck & Blake 1998; Kwon et al. 2006; Seale & Looney 2008), and the combined effects of many evolving stars can shape galactic superstructures spanning thousands of parsecs (e.g., McCray & Kafatos 1987; Tenorio-Tagle & Bodenheimer 1988). More specifically, during their formation, massive stars inject energy into the ISM via powerful jets and outflows (e.g., Shepherd & Churchwell 1996). As they evolve toward and onto the main sequence, massive stars become sources of copious amounts of UV radiation that energizes the surrounding medium creating compact H II regions (see reviews by Garay & Lizano 1999; Beuther et al. 2007). OB associations, groups of massive stars, can generate enough UV radiation to ionize large regions of the ISM. Finally, massive stars end their relatively short lives with supernovae that stir up the ISM by their strong explosions. The combined effects of supernovae and fast stellar winds can form large superbubbles or giant interstellar shells that propagate star formation to distances of hundreds or thousands of parsecs (e.g., McCray & Kafatos 1987; Chu et al. 2005; Chen et al. 2009; Book et al. 2009). In addition, it is now commonly speculated that most stars—including our own Sun (e.g., Looney et al. 2006)—form in clusters, and within these clusters, massive stars dominate the overall luminosity, greatly affecting any nearby star or planet formation. Clearly, massive stars play a considerable role in galactic ISM structure and overall stellar content.

To understand the formation of massive stars, it is necessary to identify massive young stellar objects (YSOs) throughout a galaxy. YSOs are surrounded by dense envelopes of gas and dust that serve as the young stars' mass reservoirs for accretion (see reviews by Zinnecker & Yorke 2007; McKee & Ostriker 2007). Photospheric optical and UV radiation from the embedded central star absorbed by hundreds of magnitudes of extinction from the circumstellar gas and dust envelope is reradiated in the infrared (IR). Massive YSOs spend a large fraction of their relatively short lives in these embedded stages and are therefore best identified by their point-source IR radiation (Churchwell et al. 2002). The observed IR emission, reprocessed stellar radiation, reveals information about the state of the circumstellar material (van Dishoeck & Blake 1998).

While the formation of low-mass stars can be studied in great detail and at high spatial resolution in nearby Galactic star-forming regions (e.g., Hartmann et al. 2005; Luhman & Rieke 1999; Tobin et al. 2007; Seale & Looney 2008), massive stars are much rarer, making observational studies more difficult. High-mass star formation studies within the Milky Way are plagued by numerous problems—a low number of observable sources, dust obscuration along the Galactic plane, and distance ambiguities leading to high-luminosity uncertainties.

The formation of individual massive stars in other galaxies could not be resolved until recently. The advent of the *Spitzer Space Telescope* (Werner et al. 2004) has enabled large surveys to identify YSOs in the Galaxy. Its superb angular resolution makes it possible to identify YSOs even in nearby galaxies, such as the Large and Small Magellanic Clouds (LMC and SMC). The LMC is an excellent laboratory to study high-mass star formation as it is located at a small and known distance of 50 kpc, where $1''$ corresponds to 0.25 pc (Feast 1999); furthermore, its low foreground extinction and nearly face-on orientation provides a clear view throughout the entire galaxy. Thus, surveys of high-mass YSOs in the LMC have been conducted by Whitney et al. (2008) and Gruendl & Chu (2009).

We have used the *Spitzer*'s infrared spectrometer (IRS; Houck et al. 2004) to obtain mid-IR spectra of nearly 300 massive YSO candidates in the LMC to assess and confirm their nature. These observations represent the most extensive and complete collection of massive YSO spectra to date. These IR spectra are rich with spectral features from a wide variety of species, some of which are unobservable at other wavelengths. Among the most interesting features are ice-phase molecular absorptions, deep silicate absorptions, polycyclic aromatic hydrocarbon (PAH) emission, and ionic fine-structure emission lines.

Such a large spectral survey allows meaningful comparisons with theoretically or empirically derived formation and evolutionary paths of massive stars, as reviewed by Zinnecker & Yorke (2007), Beuther et al. (2007), Krumholz & Bonnell (2007), and Garay & Lizano (1999). Star formation typically begins with the collapse of a high-density core within a large clump of a molecular cloud. A protostar forms and acquires its mass by accreting from its natal cloud via a circumstellar disk. The forming star heats the dense massive core, turning it into a hot dense massive core. Eventually, hydrogen burning will begin. It is commonly accepted that massive stars, unlike their low-mass counterparts, begin hydrogen burning while still accreting material. At this stage, the circumstellar material including the accretion disk may become photoionized, a possible origin of the so-called hypercompact H II region (Keto 2007). Initially, the accretion flow is able to quench the ionizing flux, but once accretion slows or becomes geometrically unable to contain the UV flux, the ionizing radiation can escape into the surrounding regions creating ultracompact/compact H II regions and eventually, larger classical H II regions. Over time, stellar winds and radiation will disperse the circumstellar material, revealing a newly formed high-mass star.

In this paper, we will use the term "YSO" to describe the young central object–disk–envelope system that will eventually become a main-sequence star. The term "protostar" is used more specifically to describe an object that will eventually become a main-sequence star but has yet to begin hydrogen-burning. The sample of YSOs studied here may all be hydrogen-burning YSOs that are still highly embedded and may be accreting.

Additionally, a large majority of them probably have massive central objects ($\gtrsim 10\,M_\odot$; see Section 3.5). Our sample of sources have mid-IR emission from warm dust, and thus are highly embedded but already contain a central heating source. True high-mass protostars, i.e., pre-hydrogen-burning accreting sources with masses above $10\,M_\odot$, likely only exist very briefly. Since the YSOs in our *Spitzer* IRS spectral catalog are likely high mass, most (possibly all) have already begun hydrogen-burning.

The spectra of our objects show great variations that may reflect an evolutionary sequence during a high-mass YSO's formation. This paper discusses the spectral features present in the IRS spectra and ties them to a possible evolutionary scheme. In Section 2, we introduce the catalog of sources and describe the observations used in this spectroscopic study. The results of a spectral categorization are presented in Section 3 along with our classification method. Finally, in Section 4, we discuss our proposed connection between spectral classification and evolutionary stage.

2. OBSERVATIONS AND REDUCTION

2.1. Targets

Given the benefits of studying the LMC, it is of no surprise that this galaxy has been the focus of intense star formation research in recent years. Of particular interest is the *Spitzer* survey of the LMC Surveying the Agents of Galaxy Evolution (SAGE; Meixner et al. 2006), which has enabled two catalogs of high-mass YSOs in the LMC, produced independently by Whitney et al. (2008) and Gruendl & Chu (2009, hereafter GC09). Their different selection criteria and approaches have resulted in significant discrepancies between these two catalogs (see GC09 for details). We adopt the results of GC09, a precursor of the IRS observations reported in this paper. A brief description of GC09's methodology follows. GC09 have extracted point-source photometry from all available observations of the LMC made with *Spitzer*'s Infrared Array Camera (IRAC; Fazio et al. 2004) in the 3.6, 4.5, 5.8, and 8.0 μm bands and Multiband Imaging Photometer for *Spitzer* (MIPS; Rieke et al. 2004) in the 24 and 70 μm bands. The [4.5]−[8.0] versus [8.0] color–magnitude diagram (CMD) of all LMC sources was used to identify objects with IR excesses. The nature of each object with IR excess was then assessed by simultaneously examining its images and spectral energy distribution (SED) from optical to mid-IR wavelengths. These IR objects were classified into six broad categories: asymptotic giant branch (AGB) and post-AGB stars, planetary nebulae, background galaxies, diffuse sources, normal stars, and YSOs. Objects that could not be classified unambiguously were usually assigned a primary (with higher probability) classification and a secondary (possible but with lower probability) classification. The YSOs are distributed throughout the galaxy and are seen to both cluster near some of the numerous large H II complexes in the LMC and to form in more isolated environments (see Figure 14 of GC09).

For the *Spitzer* IRS observations, we have selected 294 targets from the GC09 catalog. Of these, 269 were identified as "Definite" or "Probable" YSOs including all of the catalog's "Definite" or "Probable" YSOs with [8.0] < 8.0. In addition, we selected 10 objects with primary classification as AGB/post-AGB stars, and two that were probably diffuse sources. The latter 12 targets were included in our IRS observing program because they had properties that are unusual for their primary classifications, making classification as YSOs an attractive

alternative. For example, seven of the observed AGB/post-AGB sources belong to a subgroup dubbed extremely red object (ERO) by GC09. These EROs are redder than sources identified as typical AGB and post-AGB stars, but their SEDs are also unlike the YSOs in that they peak between 8 and 24 μm. Our IRS spectra have shown that these EROs are Extreme Carbon Stars, and are reported in detail in a discovery paper by Gruendl et al. (2008). Finally, 13 of the targets were previously uncategorized because they lacked IRAC photometry in the 4.5 μm band, but they are bright sources in the MIPS 24 and 70 μm bands and are in gas/dust-rich environments, suggesting that these sources are possibly YSOs.

2.2. Observations

The IRS on board *Spitzer* is composed of four modules, two of which provide low spectral resolution ($R \sim 60$–127) between 5.2 and 37.9 μm, and two provide high spectral resolution ($R \sim 600$) between 9.9 and 36.8 μm. To cover the entire $\sim$5 to 37 μm wavelength range, we used the IRS modules SL (short wavelength, low-resolution), SH (short wavelength, high-resolution), and LH (long wavelength, high-resolution). The SL module has two possible configurations: SL1 for the wavelength range 7.6–14.6 μm and SL2 for 5.2–8.4 μm. The high-resolution modules, SH and LH, each consist of 10 spectral orders and cover the wavelength ranges 9.9–19.3 and 18.9–36.9 μm, respectively. A more detailed description of the IRS is presented by Houck et al. (2004).

In order to efficiently observe as many LMC YSOs as possible, we used the GC09 YSO catalog to identify groups of targets with similar 8 and 24 μm brightnesses within $2°$ of one another. Each group of targets could then be observed with the same setup during a single astronomical observation request (AOR) that treated the group as a fixed "cluster." To ensure that a background estimate was available to remove the zodiacal light contribution from the SH and LH modules, we added one off-position to each cluster. This off-position was chosen by searching for the point with lowest surface brightness in the *IRAS* 12 and 25 μm maps within the $2°$ constraint imposed by the targets. The exposure times for the SL1, SL2, SH, and LH modules were chosen such that the observations would produce a contiguous spectrum from 5 to 37 μm with a signal-to-noise ratio of $\sim$30.

Seven sources were observed with only the SH and LH modules by our program. Five of these (053142.4−683454.3, 053943.3−693854.6, 053943.8−693834.0, 053944.3−693847.5, and 053945.9−693839.2) did not include the SL modules because the detectors would have saturated for the shortest possible exposures. The two remaining sources, 053941.9−694612.0 and 053959.3−694526.3, had previous observations with the SL1 and SL2 modules available in the *Spitzer* archive and could be observed as part of the same cluster. For these latter two sources, we have downloaded and combined the archival data with our SH and LH observations.

2.3. Data Reduction

The observations were processed by the *Spitzer* Science Center's pipeline reduction software. We began each spectral extraction process with the basic calibrated data (BCD) product. The campaign-specific rogue pixel masks were applied to the BCD data and removed using the interactive IDL IRSCLEAN package (version 1.9). We further identify and remove transient rogue pixels by examining variations between frames, also using

the IRSCLEAN package. Multiple exposures at each slit position were then median averaged together.

The subtraction of a background was carried out differently for the high- and low-resolution modules because of their different slit sizes: $3''.7 \times 57''$ for SL1, $3''.6 \times 57''$ for SL2, $4''.7 \times 11''.3$ for SH, and $11''.1 \times 22''.3$ for LH. For SL1 and SL2, the slit is long enough that a local background along the slit can be determined and subtracted from the source. This is achieved by differencing the two nod positions. For the SH and LH modules, the slit is too short for a source-independent background; thus, the background subtraction is performed by subtracting the cleaned BCD data from a nearby background position.

High-resolution spectra were extracted from the full SH and LH apertures using the SMART software package (Higdon et al. 2004). For a full aperture extraction, any flux within the bounds of each order at a given wavelength is extracted and summed. The low-resolution SL spectra were extracted using the tapered column point source extraction package within SMART. Because of the large number of spectra, an automated procedure was used to detect a source in a cleaned, background-subtracted spectral image and extract its spectrum with an extraction window whose width varies to accommodate the wavelength-dependent point spread function (i.e., wider at longer wavelengths). Fifty-eight sources could not be detected by the automated procedure either because they appeared extended or resided in a crowded environment, resulting in multiple sources within the slit. These sources were extracted manually using the interactive portion of SMART. Here the position and FWHM of the source on the spectral image were visually determined and the spectrum extracted using a tapered column.

To combine (or "stitch" together) spectra from different modules into a single contiguous spectrum, several corrections need to be made. First, the fringes in the high-resolution one-dimensional spectra, originating from the detectors of the SH and LH modules, are removed using the IDL package IRSFRINGE (version 1.1). Next, flux discontinuities between modules are corrected by applying scaling factors to spectra from different modules. Flux discontinuities may arise from differences in slit size and orientation, background-subtraction, and spectral extraction. To account for the latter two effects, we assume that the local background-subtraction in the SL1 and SL2 spectra are more reliable, and further scale SL1 to SL2 (scaling factor is $\sim$1; average of 0.98 with standard deviation of 0.14). The LH spectrum is then combined with the SH spectrum after applying a scale factor determined from the ratio of fluxes in the wavelength region where the two overlap ($\sim$18.9–19.3 μm). Finally, the SH + LH combined spectrum is scaled to the SL1 + SL2 combined spectrum based on the ratio of fluxes in the region where the two spectra overlap ($\sim$9.9–14.6 μm). Typical values of the net scaling factors range between 0.5 and 2.0.

2.4. Cautionary Remarks on Stitching Spectra

The scaling factors can be used to produce a contiguous spectrum, but may also introduce inadvertent effects. Several considerations need to be born in mind when interpreting these spectra.

If the full aperture were uniformly illuminated in both the SH and LH modules, the scale factor between them would simply be the ratio of their aperture extraction areas. If the observed source is point-like, then no scaling factor would be required. The fact that a scaling factor is required suggests an extended

and dusty nature of the objects or a significant contribution from background emission. Furthermore, the scaling factor is determined from the dust continuum, but the lines may behave differently. For example, the fine-structure lines may originate from a small compact core, while the continuum is extended; therefore, the lines may not have required scaling.

The necessity of module stitching is partly a result of each module's different aperture size and orientation which result in different spatial regions being probed. The stitching process is further complicated by possible differences in the spatial extent of the regions emitting at different wavelengths. For example, the emission at longer wavelengths may originate from cooler, more extended gas and dust surrounding a smaller, warmer region which dominates the emission at shorter wavelengths. Alternatively, since the spectrum at shorter wavelengths contains many spectral emission features associated with dust, the spatial extent of the emitting region at short wavelengths may be equal or even larger than that of the region emitting at longer wavelengths. To get a sense of how strongly the stitching is affected by spatial differences in emitting areas, we compared $F_{24}^{\rm phot}/F_{24}^{\rm spec}$ (the ratio of MIPS photometric 24 μm fluxes to IRS 24 μm spectroscopic fluxes) to the scaling factor applied to the 24 μm wavelength region. We find that any source with $F_{24}^{\rm phot}/F_{24}^{\rm spec} \gtrsim 1.5$ has a scaling factor less than 1. In other words, the long-wavelength portion of the spectrum was scaled down to match the short-wavelength portion causing the 24 μm spectroscopic flux value to underestimate the photometric value. This is consistent with the picture described above where emission contributing to the short-wavelength portion of the spectrum originates from a smaller region than the long-wavelength portion.

3. RESULTS AND CLASSIFICATION

The spectra are first separated by eye into two primary categories that reflect their level of "embedded-ness." The first category's members all have continua that rise into the red end of the spectrum, while the second category's spectra are much flatter or even begin to fall at long wavelengths. The second category also contains sources with slightly rising spectra with a strong emission feature at 10 μm. The first category are likely highly embedded objects such as YSOs while the second are likely either less-embedded, more-evolved young stars or highly evolved stars such as AGBs, post-AGBs or red supergiants (RSGs). The embedded sources are additionally organized via principal component analysis (PCA; see Section 3.2) into six groups according to their dominant spectral features. Note that the details of the use of PCA follows in Sections 3.2 and 3.3, but first, we will discuss the spectra and their features more generally. The less-embedded object spectra are further subdivided by eye into three groups based upon their very different dominant features. It is more efficient to classify these spectra by eye since the spectra are very different among the three groups and do not suffer from the same by eye biases that the embedded source classifications do. In this manner, the surveyed sources are separated into nine different spectral classifications, and a sample of the spectra for each category is presented in Figures 1(a)–(c) and 2(a). The spectral catalog can be accessed in its entirety (Figures 1(d)–(ee) and 2(b) and (c)) in the electronic version of this paper. A summary of the source classification is given in Table 1.

Our categories closely resemble a portion of a classification system originally developed by Kraemer et al. (2002) to classify

spectra from the Infrared Space Observatory (ISO; Kessler et al. 1996) short-wavelength spectrometer (SWS), hereafter referred to as the KSPW system. In the KSPW system, objects are first classified by their continuum shape into six broad groups with subclasses assigned based upon specific spectral features. Under the original KSPW system, Kraemer et al. (2002) classified over 900 objects of various astronomical origin into a total of 59 classes. All the sources we have identified as embedded would be broadly classified into the KSPW Group 5, sources with SEDs dominated by dust emission with a peak wavelength longer than 45 μm. Since *Spitzer* IRS spectra cover a different wavelength range than ISO, and because our catalog focuses on a narrow portion of the large spectrum of categories established in the KSPW system, we have developed a modified version of the categories originally developed by Kraemer et al. (2002). In Table 1, we list both our own classification of each source along with the corresponding classification in the KSPW system. Classification by the KSPW scheme was performed by eye, and every source was classified three times with no knowledge of prior categorization to ensure an unbiased classification. There is generally a one-to-one correlation between the two classifications, but because classifications in our system are performed automatically via PCA for the embedded sources while the KSPW system is done by eye, there are some discrepancies.

3.1. Embedded YSO Spectral Features

The mid-IR is rich in spectral features including bands due to molecular ices, PAHs, silicates, and atoms, thus spectroscopic observations provide an abundance of information about the physical and chemical properties of a forming star and its environment. The spectra of the coldest and therefore youngest YSOs peak around 100 μm (e.g., van Dishoeck 2004), corresponding to a blackbody temperature of several tens of Kelvins. In environments at high densities and low temperatures such as these, absorption from ices (particularly CO_2) and silicates are detected in the spectra. Older YSOs begin to show signs of UV radiation from a central ionizing source including UV-pumped fluorescing PAHs and fine-structure lines from compact H II regions.

The six groups of spectra in the first category, whose continua rise with wavelength throughout the IRS spectral range, are described below and examples are shown in Figures 1(a)–(c) with major spectral features marked. (1) The S Group, consisting of 12 sources, are dominated by silicate absorption features including the prominent 10 μm absorption feature and sometimes including the 18 μm absorption feature. (2) The SE Group, consisting of five sources, have both strong silicate absorption features (similar to S Group) and strong fine-structure emission lines from ionized gas. (3) The P Group, consisting of 100 sources, show prominent PAH emission features. The sources in this group may also show some absorption from silicates, particularly at 10 μm, but PAH emission features at 6.2, 7.7, 8.6, and 11.3 μm make it difficult to identify the silicate absorption unambiguously. For example, strong PAH emission features on either side of a weak 10 μm silicate feature may mimic a strong silicate absorption feature. (4) The PE Group, consisting of 142 sources, show both strong PAH emission features and fine-structure lines. Similar to the previous group, silicate absorption features are also difficult to identify in the presence of strong PAH emission features. (5) The E Group consists of only two sources whose dominant features are fine-structure lines originating from ionized gas. This group appears
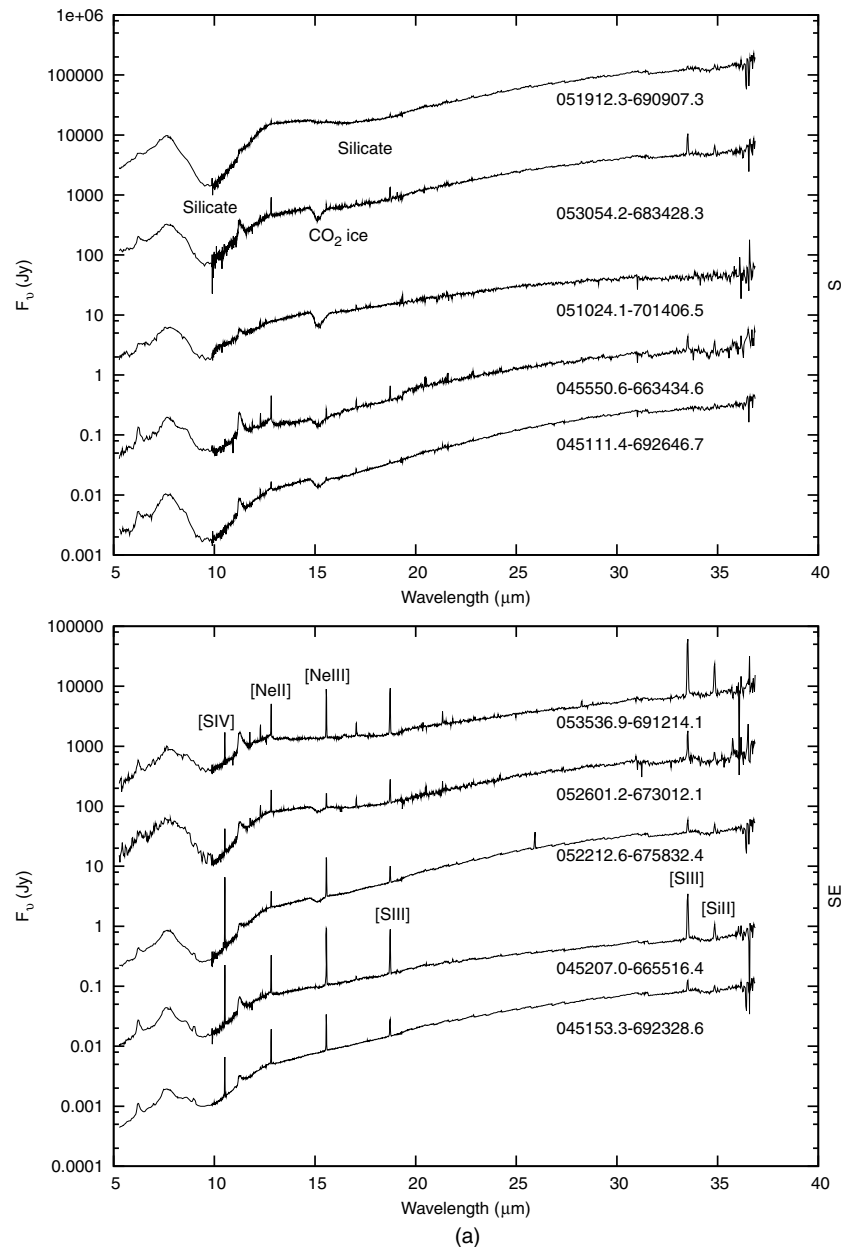
Figure 1. (a) Representative spectra of the S and SE Groups of the embedded YSO sources. The source identification is given below each curve. Top: the broad silicate absorption features at 10 and 18 μm and the 15.2 μm CO_2 ice band are identified. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 0.05, 1, 30, 800, and 10^4. Bottom: prominent fine-structure lines are indicated including [S III], [S IV], [Ne II], [Ne III], and [S III]. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 0.001, 0.1, 1, 10^3, and 8×10^3. (b) Representative spectra of the P and PE Groups of the embedded YSO sources. The source identification is given below each curve. Top: the prominent PAH emission features are identified. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 0.05, 1, 10, 800, and 3×10^4. Bottom: prominent fine-structure lines are indicated along with three PAH emission features. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 1, 20, 600, 6000, and 2×10^5. (c) The two spectra of the E Group and the five spectra of the F Group with complete spectral coverage. The source identification is given below each curve. Top: prominent fine-structure lines are indicated including [S III], [S IV], [Ne II], [Ne III], and [S III]. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 1 and 10. Bottom: the spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 1, 140, 24000, 10^5, and 10^7.

(The complete figure set (31 images) is available in the online journal.)

similar to the PE Group and may be an artifact of the thresholds used in the PCA. (6) The F Group, consisting of 11 sources, do not meet the criteria for the above five groups, but are spectrally similar to embedded YSOs. They may show weak indications of PAHs, silicates, or fine-structure lines but are generally nearly featureless.

The largest of the six groups is the PE Group, which is characterized by both strong PAH and fine-structure line

emission. PAH emission features at 6.2, 7.7, 8.6, 11.3, 12.7, 14.2, and 16.2 μm, first identified in the mid 1970s in the spectra of compact H II regions and planetary nebulae (e.g., Soifer et al. 1976), originate from the C–C and C–H stretching and bending modes of PAHs excited by UV radiation. Fine-structure emission lines such as [S IV] 10.5 μm, [Ne II] 12.8 μm, [Ne III] 15.5 μm, [S III] 18.7 μm and 33.5 μm, and [S III] 34.8 μm originate from ionized gas and can be used to assess the
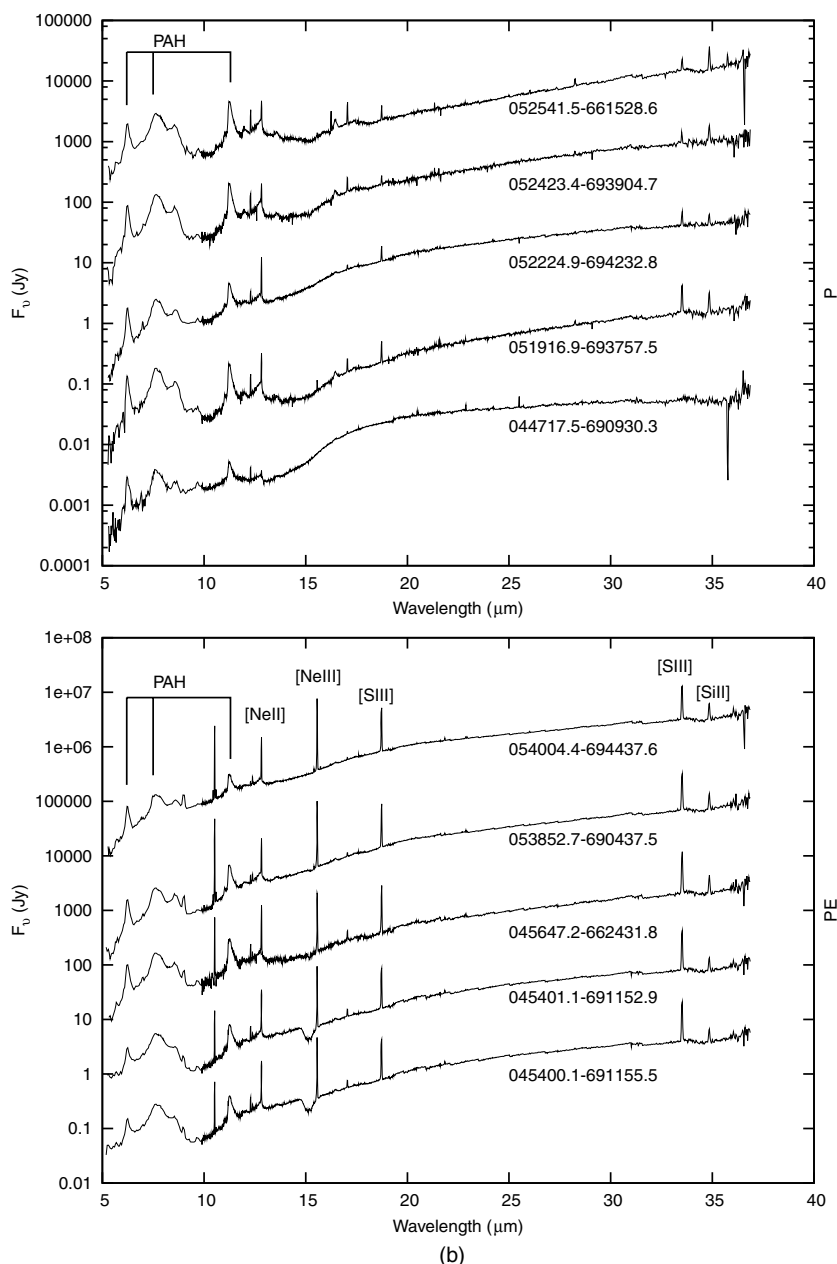
Figure 1. (Continued)

hardness of the UV radiation field. The presence of both PAH and fine-structure line emission clearly suggests the presence of an ionizing source with UV radiation.

The deep silicate feature at 10 μm and the shallower 18 μm feature in the S and SE Groups are generally smooth, indicating silicate dust in its amorphous form; however, several sources show a more structured 10 μm feature (Figure 3), possibly a result of dust processing (van Dishoeck 2004). There appears to be a range of strengths to the structure in our sample, possibly representing evolutionary changes from environmental heating. However, there are also known emission lines (e.g., H2 S(3) at 9.64 μm) which may be contributing to the non-smooth silicate profile. This phenomenon will be further investigated in the future. A majority of the sources dominated by silicate absorption also have a prominent solid CO_2 absorption at 15.2 μm. The IR vibrational bands of CO_2 are known to be very sensitive to the environment, and are therefore excellent tracers of ther-

mal and chemical conditions within the embedding molecular core. Similar to the 10 μm silicate band, the CO_2 absorption feature can be either double-peaked in a smoother single-peaked form (Figure 3). The relative strength of the double/single peaked components is indicative of the environment the ice was formed in (van Dishoeck 2004) and will be explored in future papers.

3.2. Principal Component Analysis

For the embedded sources, we organize the spectra into groups with common spectral features by using PCA to quantify the strength of the various spectral features described above. PCA is a technique that allows for the recognition of patterns in spectra and can be used to express a set of spectra in ways that highlight both their similarities and their differences (Murtagh & Heck 1987). PCA has previously been used in the spectral
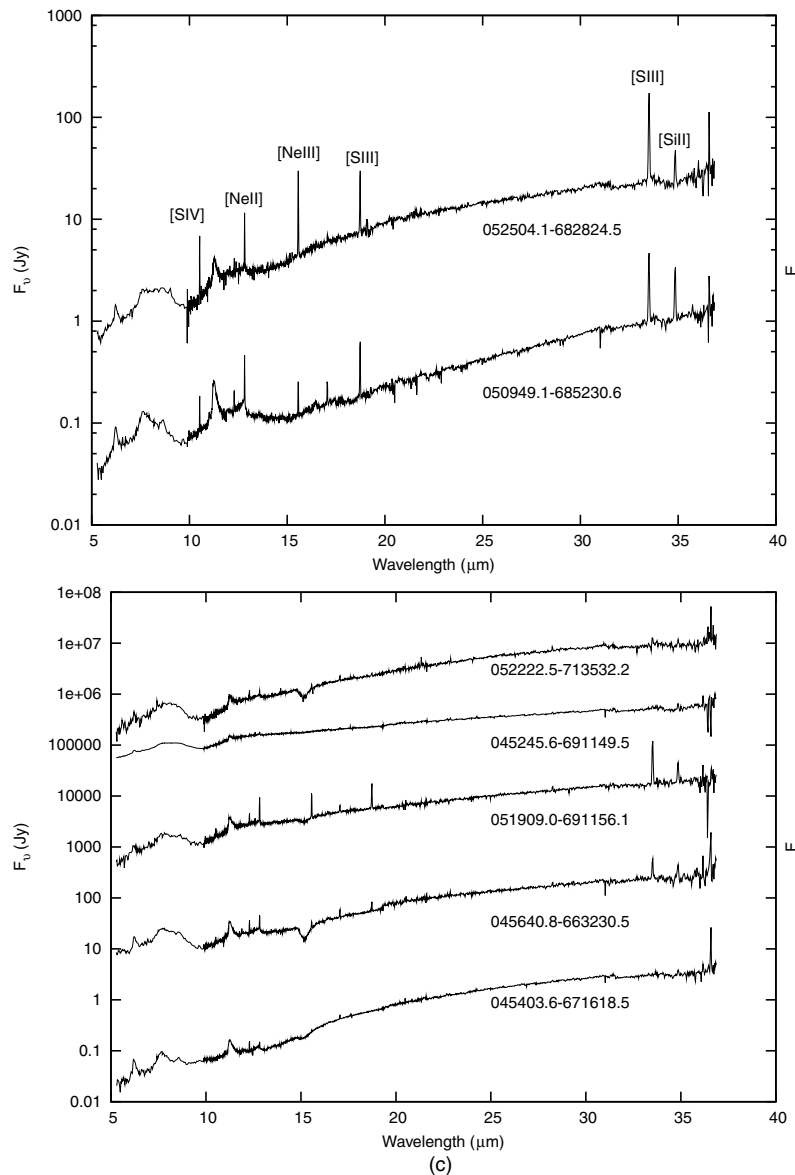
Figure 1. (Continued)

classification of stars (e.g., Murtagh & Heck 1987), QSOs (e.g., Francis et al. 1992), and galaxies (e.g., Connolly et al. 1995). Given the large number of spectra in our sample and the great number of possible spectral features in every spectrum, it can be difficult to visually recognize patterns in the absence, presence, or strength of spectral features. PCA removes by eye classification guesswork and can quantify the differences between spectra, making it an impartial and unbiased method to assign classifications to spectra. We use PCA to identify and quantify the strength of several spectral features including silicate absorption, PAH emission, and fine-structure lines. We then classify the spectra into several groups based upon the strengths of these features. What follows is the PCA recipe used in this study. For a more complete description of the process and proper implementation of PCA, see Murtagh & Heck (1987).

Suppose we have a catalog of N spectra ($i = 1, 2, \ldots, N$), each covering the same wavelength range sampled at M ($j = 1, 2, \ldots, M$) identical wavelengths. Each spectrum can be represented by a vector S_i in an M-dimensional space containing the flux at M wavelengths where i corresponds to the specific

source. We can then construct an $N \times M$ matrix, A, in which each row contains the M flux measurements for each of the N sources. In this manner, each row represents a single point in the M-dimensional A space. If M is a small number, i.e., if A were a low-dimensional space, it may be possible to recognize groups and classifications of spectra by their location in A space. But when covering a large wavelength range, with even moderate resolution many spectral features, the number of dimensions becomes too great to visualize without the aid of a statistical method. PCA allows us to reconstruct the N-dimensional space into a projection of only a handful of principal coordinates/ components. PCA is an orthogonal linear transformation of data (in this case, spectra) to a new coordinate system such that the greatest variance in the data can be attributed to the first coordinate, the second most variance to the second coordinate, and so on. Each new coordinate is called a component, with those contributing the most variance being the principal components. The new coordinates can be expressed graphically by treating each coordinate as a spectrum, in the same way an object's spectrum can be represented as a vector.
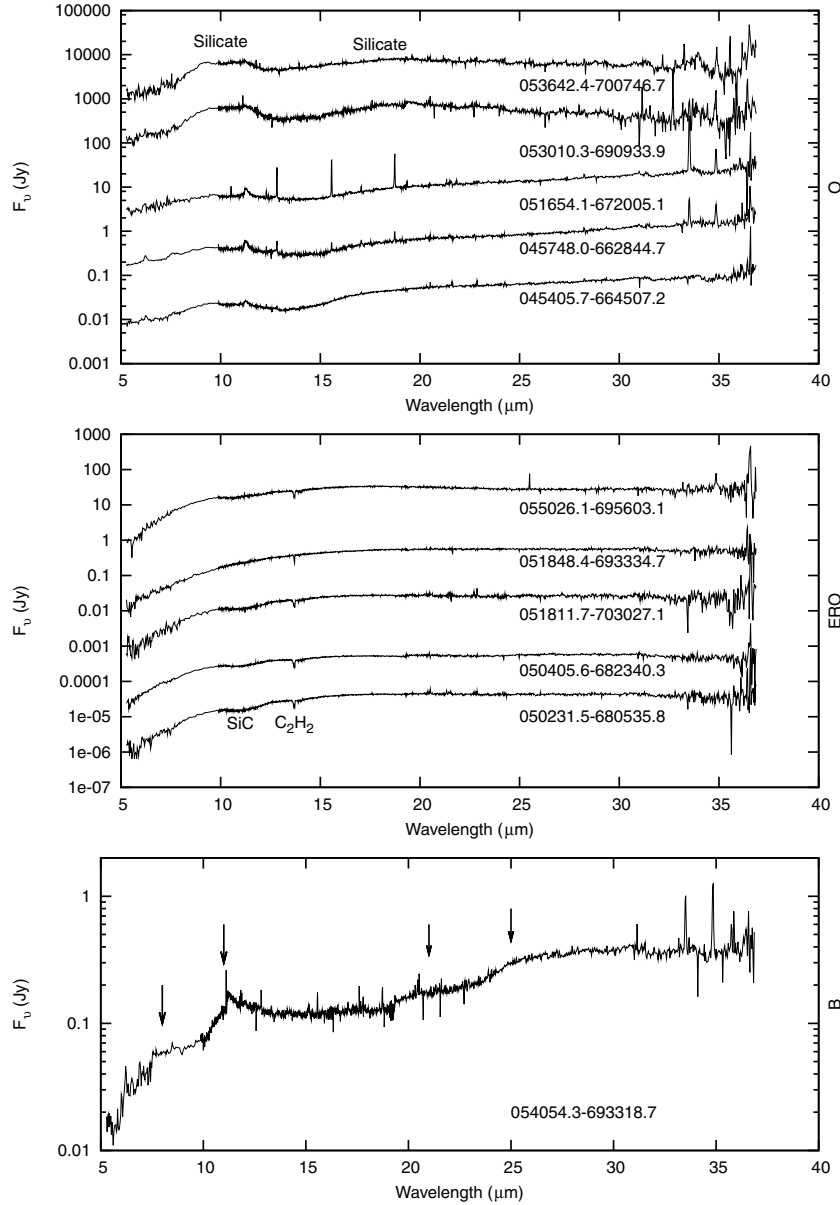
Figure 2. (a) Representative spectra of the O, ERO, and B Groups in top, middle, and bottom panels respectively. The source identification is given below each curve. Top: silicate emission features at 10 and 18 μm are indicated. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 0.1, 1, 100, 4000, and 8×10^4. Middle: absorption features coincident with known SiC and C_2H_2 bands are indicated. The spectra from bottom to top have been scaled by the following multiplicative factors for clarity: 10^{-4}, 0.1, 1, 100, and 10^{-3}. Bottom: arrows indicate the locations of unidentified broad emission features that create the step-like spectral shape. (b) and (c) Spectra of the complete source catalog from which selected sources are shown in Figure 2. Sources are ordered as in Table 1, and source identifications are given below each curve.

(An extended version of this figure is available in the online journal.)

To implement PCA, we must normalize all the spectra to isolate variances due to spectral features and minimize variances from bolometric luminosity differences. The first step is to scale all the spectra to the same total integrated luminosity over the wavelength range, 5.5 to 37 μm. Next, we use a spline interpolation to place all spectra in the same wavelength grid. A final normalization is performed to create a data set whose mean is zero. To do so, the mean of each dimension is subtracted from each flux measurement in that dimension. If f_{ij} is the flux measurement for the ith source in the jth dimension, the full normalized spectra matrix A is constructed as follows:

$$A_{ij} = f_{ij} - \frac{1}{N} \sum_{i'=1}^{N} f_{i'j}.$$

We can then compute the $M \times M$ covariance matrix, C, of the matrix containing all the spectral data:

$$C_{jk} = \frac{1}{N} \sum_{i=1}^{N} A_{ji} A_{ki}$$

where both j and k range between 1 and M, inclusive.

We can then calculate M eigenvectors and the M corresponding eigenvalues of this covariance matrix. The eigenvectors e_m and eigenvalues λ_m obey the identity equation

$$C e_m = \lambda_m e_m.$$

The eigenvector with the largest absolute value of the paired eigenvalue represents the dimension containing the largest

Table 1
Source Classification

Object ID	KSPW Class[a]	GC09 Class	Silicate Absorption (Projection[b])	PAH$_{5-10\mu m}$ Emission (Projection[b])	PAH$_{10-13\mu m}$ Emission (Projection[b])	Fine-structure Line Emission (Projection[b])	Possible Multiple	Molecular Cloud	Compact Hα	Diffuse Hα ([S III] Concentration Ratio[c])	Diffuse Dust
					S Group						
045111.39−692646.7	5.SA	C	Yes (0.546)	No (0.402)	No (0.411)	No (0.268)	✓	✓	✓	✓	✓
045550.62−663434.6	5.U	C	Yes (0.361)	No (0.674)	No (0.372)	No (0.375)	✓	✓	✓	✓	✓
					SE Group						
045153.29−692328.6	5.SAe	C	Yes (0.361)	No (0.835)	No (0.571)	Yes (0.175)	✓	✓	✓	✓ (2.04)	✓
045206.97−665516.4	5.SAe	C	Yes (0.424)	No (0.817)	No (0.454)	Yes (−0.293)	✓	✓		✓ (0.98)	✓
					P Group						
044717.50−690930.3	5.U	C	No (−0.236)	Yes (−0.031)	Yes (0.316)	No (0.3)		✓	✓		✓
044854.41−690948.3	5.UE	C	No (−0.039)	Yes (0.187)	No (0.544)	No (0.287)	✓	✓		✓	✓
					PE Group						
044849.85−690930.6	5.UE	C	No (−0.106)	Yes (−0.315)	Yes (−0.386)	Yes (0.103)	✓	✓		✓ (1.05)	✓
044927.29−691206.3	5.UE	C	No (0.015)	Yes (−0.124)	Yes (0.062)	Yes (−0.383)	✓	✓		✓ (1.99)	✓
					E Group						
050949.11−685230.6	5.UE	C	No (−0.51)	No (0.832)	No (0.349)	Yes (0.124)	✓	✓		✓ (1.4)	✓
052504.10−682824.5	5.UE	C	No (−0.16)	No (0.607)	No (0.631)	Yes (−0.019)		✓		✓ (1.26)	✓
					F Group						
045201.15−692007.7	5.UE	C	No (−0.5)	No (0.482)	No (0.608)	No (0.415)		✓		✓	✓
045245.62−691149.5	5.SA	CA	No (−0.201)	No (0.829)	No (0.562)	No (0.285)		✓	✓		✓
					Embedded objects with unknown classification						
053142.44−683454.3	5.UE	C	⋯	⋯	⋯	⋯	✓		✓	✓	✓
053943.26−693854.6	5.E	C	⋯	⋯	⋯	⋯	✓	✓		✓	✓
					O Group						
045405.74−664507.2	4.SE	C	⋯	⋯	⋯	⋯	✓	✓	✓		✓
045747.96−662844.7	4.SE	U	⋯	⋯	⋯	⋯		✓		✓	✓
					ERO Group						
050231.49−680535.8	4.CR	E	⋯	⋯	⋯	⋯					
050405.60−682340.3	4.CR	E	⋯	⋯	⋯	⋯					
					B Group						
054054.25−693318.7	4.CT	C	⋯	⋯	⋯	⋯	✓			✓	✓

Notes.

[a] Primary GC09 classifications for the entire source sample. When applicable, the primary classification is followed by the secondary classification. GC09 classifications are as follows: C – candidate YSO; A – AGB or post-AGB star (non-ERO); D – diffuse source; G – background galaxy; S – normal star; E – extremely red object. Sources without a GC09 classification are assigned a U – unknown classification.

[b] Dot product of the source's spectrum with the eigenvector containing the pertaining spectral feature.

[c] Ratio of the equivalent width per pixel of the [S III] 18.7 μm fine-structure line of the full slit extraction to the equivalent width per pixel of the same line taken at the 1 pixel wide "off-source" slit position. Sources without values do not have a measurable [S III] line.

(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.)

variance, the first principal component; the eigenvector with the second largest eigenvalue is the dimension with the second largest variance, the second principal component, and so on. The fractional variance from each principal component is $\frac{\lambda_i}{\sum_i \lambda_i}$. A matrix of all *M* eigenvectors represents a complete set of orthogonal axes which completely describe the entire data set. If the summed fractional variance of the first few principal components is ∼1, then the spectra can be well represented by projecting the spectra onto these first few principal components. We find that often 90% or more of the variance can be described by the first few principal components.

There are several weaknesses to using PCA including its assumption of linearity. This problem is amplified by the nature of our source catalog. Since many of our embedded sources are likely small clusters of YSOs, features from several different sources are mixed together into a single spectrum. In fact, within the H II complex N44, Chen et al. (2009) found 65% of the candidate YSOs to be resolved multiples. Therefore, features that may show a strong correlation within a single source may not have as strong a correlation when originating from a clustered object. In this study, we are using PCA to extract and quantify spectral features. Since we are only using PCA to broadly categorize spectra by their major spectral features, we accept the assumption of linearity as valid to first order.

3.3. Automated Classification of Embedded YSOs

In order to use PCA to quantify the strength of different spectral features, four wavelength regions are chosen that contain the strongest components of a particular species type— one region highlights the strong 10 μm absorption feature, two highlight PAH emissions between 5 and 10 μm and between 10 and 13 μm, and a final region between 10 and 20 μm encompasses four ion fine-structure lines. PCA is performed on each range to quantify the contribution from various features. Each of these four wavelength regions goes through the procedure described above to find the principal components in that region. The five sources without complete spectral coverage were excluded from this analysis.

The wavelength segment between 7 and 15 μm is used to determine the strength of each spectrum's 10 μm silicate absorption. The second principal component of this wavelength range (Silicate PC of Figure 4), primarily responsible for
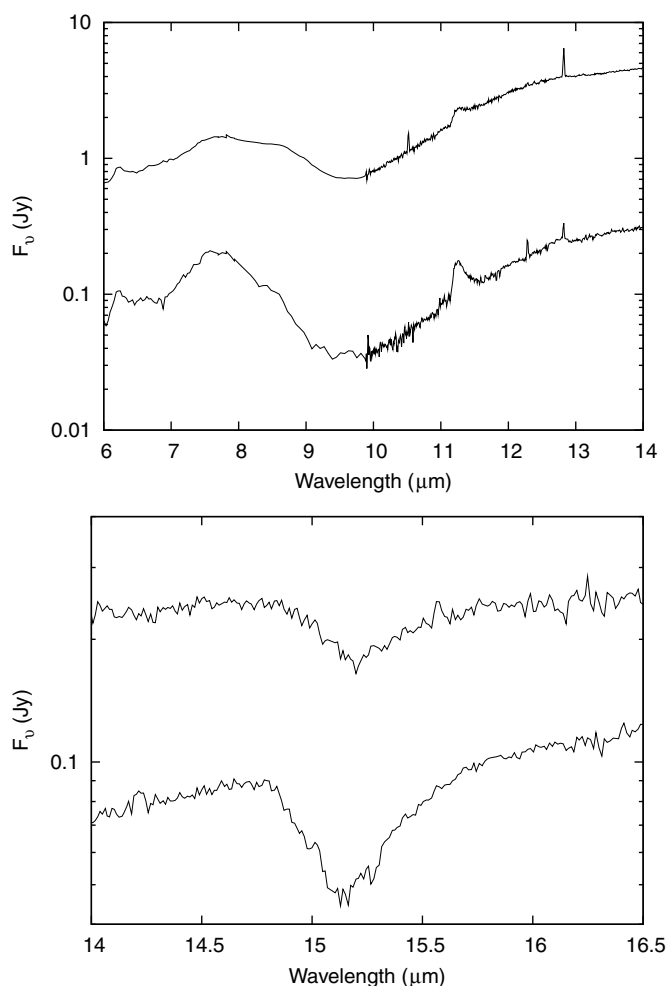
Figure 3. Top panel: spectra of 052646.61−684847.2 (top) and 045111.39−692646.7 (bottom) showing the variation in silicate absorption band profiles. Bottom panel: spectra of 053754.82−693435.8 (top) and 045842.47−660835.7 (bottom) showing the variation in CO_2 ice absorption band profiles. The flux values of 053754.82−693435.8 have been scaled by a factor of 2 for clarity.

silicate absorption at 10 μm, accounts for 21% of the variance in this wavelength range, and is used to assess the strength of a spectrum's silicate absorption. To get a measure of the amount of silicate absorption in a single spectrum, we do a projection of the vector form of the spectrum into the principal component containing the silicate feature, in this case the second principal component of the 7–15 μm segment. Figure 5 displays what we call the Silicate Projection of each source, dot products of the embedded sources' spectra with the Silicate PC. A value of zero means there is no contribution from this eigenvector in the spectrum (the spectrum's vector is perpendicular to the component's vector), while positive and negative values indicate positive and negative contributions, respectively; more positive dot products mean larger silicate absorptions. Because PC2 is primarily the result of silicate absorption, we define all sources with a Silicate PC projection greater than 0.2 to have strong silicate absorption.

There are two tests implemented to establish the presence of PAH emission. PCA is applied to two wavelength regions, each containing strong bands of PAH emission. The first, between 5 and 10 μm, contains PAH features at 6.2, 7.7, and 8.6 μm, and the second, between 10 and 13 μm, has a feature at 11.3 μm.

Between 5 and 10 μm, the second principal component, which contains the negative of three PAH bands, is responsible for 26% of the spectral variance. Similarly, component 2 of the 10 to 13 μm range contains the negative of a PAH feature which is responsible for 33% of the variation. These components are shown in Figure 4. We again project the spectra of the sources into these components to determine the contribution from each eigenvector. Since the principal components with PAHs contain negatives of the features, objects with strong PAH features should have small or negative projections onto both these principal components. We set a cutoff of 0.33 for both ranges, and define any source with a second component projection of less than 0.33 in either wavelength range to have strong PAH emission. Figure 5 shows the dot products (PAH Projections) of all the embedded source spectra onto both the PAH principal components. Note that while most objects that demonstrate strong PAH emission from one test also pass the second PAH emission test, a minority of sources pass only one test. The relative strengths of different PAH features will depend on the specific composition of the emitting PAH (Tielens et al. 2004) and will be explored in future studies.

A final PCA test is implemented to test for spectral contributions from fine-structure lines. Between 10 and 20 μm, there are four lines attributed to [S IV], [Ne II], [Ne III], and [S III]. The third principal component in this wavelength range, which contains the negative of the fine-structure lines and is plotted in Figure 4, accounts for only 5% of the spectral variance. We define sources showing strong fine-structure emission lines as having a projection into the third component of less than 0.2. Figure 5 plots the Fine-structure Line Projections, the projections onto the third eigenvector, and indicates the boundary between sources we define as with and without fine-structure lines.

To test the ability of PCA to correctly identify related spectral features and produce their proper shape, we have compared the principal components to sources with known spectral origins. The silicate principal component has a direct counterpart in the spectrum of the embedded massive YSO W33 A. The silicate principal component along with W33 A's 7–15 μm spectrum are shown in Figure 4. The spectral data of W33 A (Gerakines et al. 1999) were obtained from the ISO archive via the NASA Astrophysics Data System's[6] online data function. The ISO data were reduced using SMART, multiple exposures were averaged, and the spectrum was smoothed to the resolution of the *Spitzer* IRS. The IR spectra of W33 A—a YSO embedded within a cold, dense envelope—is known to be dominated by deep ice and silicate absorption features. Note that the broad feature at 10 μm in the principal component matches in location and breadth the deep silicate absorption seen in the spectrum of W33 A. We further compared the PAH principal components to the ISO spectrum of HD 44179. Spectra were obtained and processed for HD 44179 in the same manner as for W33 A. HD 44179, associated with the Red Rectangle nebula, is the prototype object showing PAH emission features. Now identified as a post-AGB star, HD 44179 illuminates the surrounding nebula and fluoresces PAH grains in the IR. The structured peaks seen in the PAH principal components match well features at 6.2, 7.7, 8.6, and 11.3 μm identified as PAH emission in the Red Rectangle (Figure 4).

Applying the four PCA cutoffs allowed us to quantitatively assess the contribution different spectral features have to a
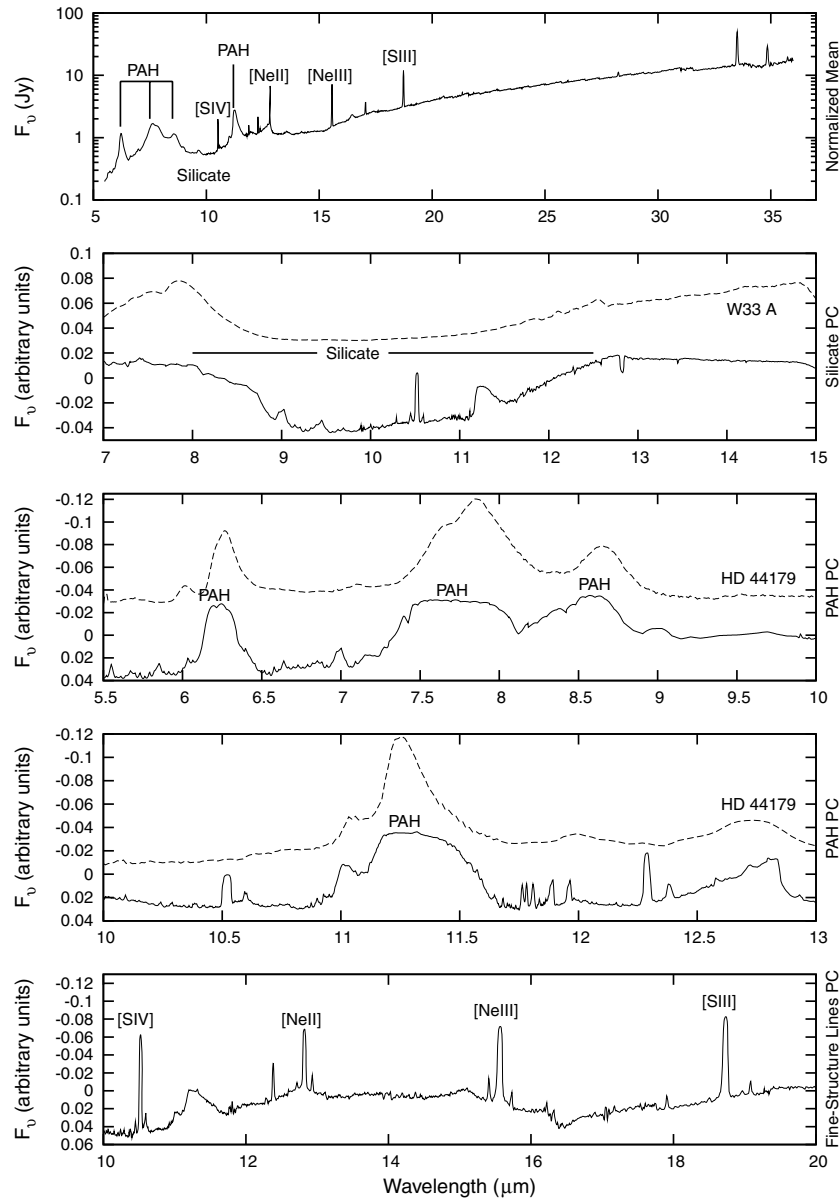
6 http://adswww.harvard.edu/

Figure 4. Top panel: full normalized mean spectra of all the embedded YSOs. Bottom panels: the solid curves show the four principal components used in our analysis. The dashed curves are smoothed ISO data from objects with IR spectra dominated by the species represented within each particular principal component. The silicate panel shows an ISO spectrum from the highly embedded protostar W33A, while the PAH panels have the baseline-subtracted ISO spectrum of HD 44179, the post-AGB star at the center of the Red Rectangle protoplanetary nebula. Prominent features are identified in each panel.

spectrum. The results of the analysis are given in Table 1. There are 54 sources which pass the 10 μm silicate absorption test, 243 which pass at least one PAH test, and 152 which pass the fine-structure line test. Using the results of the PCA, we divide the embedded objects into the six groups described earlier: (1) sources displaying only strong silicate absorption, (2) sources with only strong silicate absorption and fine-structure lines, (3) those with dominant PAH features, (4) objects with both strong PAH features and fine-structure lines, (5) sources with only fine-structure emission lines, and (6) largely featureless spectra. Note that although 54 sources pass the 10 μm silicate absorption test, only 12 and 5 objects are in the S and SE Groups, respectively. This is due to the earlier-stated reason that the blending of PAH emission with silicate absorption makes the two features difficult to distinguish. Any source identified as having PAH emission by definition can only belong to the P or PE Groups. The five embedded sources which do not have full spectral

coverage cannot be properly classified into any particular group because PCA cannot be performed. Therefore, despite showing fine-structure lines, they have been left unclassified and are represented in Table 1 as embedded objects with unknown classification.

3.4. Classification of Non-embedded Sources

By eye classification is implemented for the less-embedded sources whose spectra are largely flatter or fall at longer wavelengths (Figure 2). The first non-embedded group (Group O) all have a broad emission feature at 9 − 11 μm superimposed on a cool dust continuum. The ∼10 μm feature corresponds to emission from silicates typical of oxygen-rich objects including O-rich AGBs and the disks of Herbig Ae/Be stars. The slope of the spectrum at the long-wavelength end of the spectrum should be an indicator for how much gas and dust is around
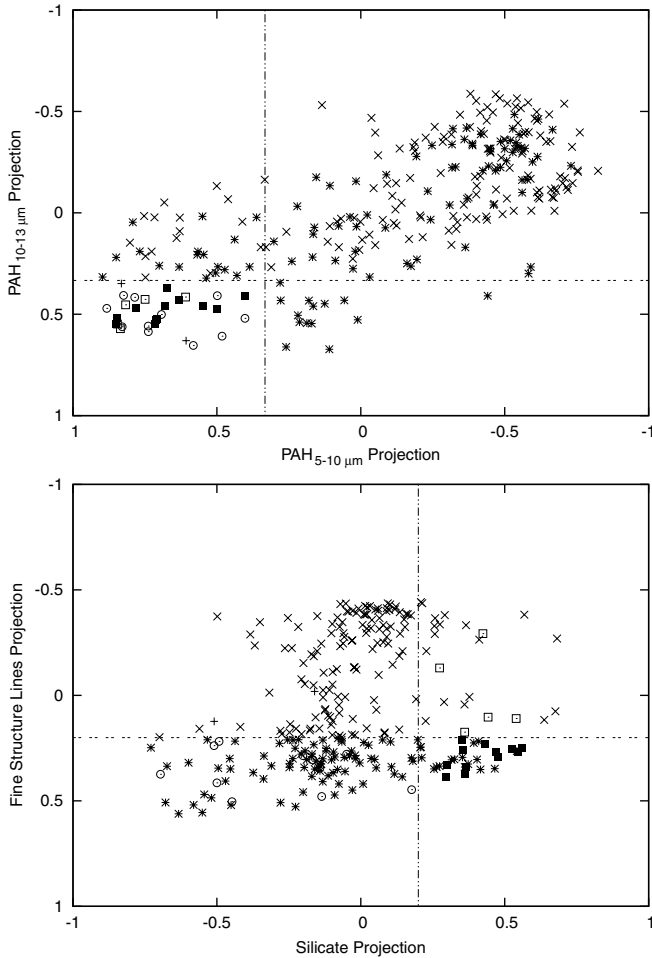
Figure 5. Top: projections of each embedded YSO's spectrum onto both the PAH$_{5-10\,\mu m}$ and PAH$_{10-13\,\mu m}$ principal components. S Group sources are indicated by solid squares, SE objects by open squares, P Group by stars, PE Group by "×"s, E Group are crosses, and F objects are open circles. The vertical and horizontal dotted lines indicate the cutoffs between sources which pass and fail the PAH$_{5-10\,\mu m}$ and PAH$_{10-13\,\mu m}$ tests, respectively. Objects with projections to the right of the vertical line or above the horizontal line are considered to have PAH emission. Bottom: projections of each embedded YSO's spectrum onto both the silicate and fine-structure line principal components. The vertical and horizontal dotted lines indicate the cutoffs between sources which pass and fail the silicate and fine-structure line tests, respectively. Objects with projections to the right of the vertical line have silicate absorption, while those above the horizontal line are considered to have fine-structure line emission. Symbols are identical to the top panel.

the source; since the long-wavelength part of the mid-IR is dominated by thermal radiation from dust, spectra which are flat or rise into the far-IR must be surrounded by neighboring gas and dust. Objects without much circumstellar dust would be expected to have spectra which fall into the far-IR. Indeed, of the eight O Group sources, two of the three with spectra that fall at long wavelengths were suggested by GC09 to be either possible AGBs or stars with an IR excess, stars which should no longer be associated with much circumstellar material. We probed the dusty environment of these sources by inspecting their IRAC images which are sensitive to both PAH emission and thermal dust emission and assessed their locations with respect to molecular clouds with CO observations from the NANTEN CO($J = 1$–0) survey (Fukui 2002). Source 053642.42−700746.7, located on the outer edge of a molecular cloud, is the only source with a falling spectrum located within a dusty complex and a molecular cloud. This is in great contrast

Table 2
Comparison Between GC09 and Our Classification

Our Classification Group	Total	GC09 Primary and Secondary Classification						
		YSO	AGB (Non-ERO)	Galaxy	Diffuse	Stellar	ERO	Unknown
Total	⋯	**269**	**3 (1)**	**0 (1)**	**2 (2)**	**0 (5)**	**7**	**13**
YSO	**277**	262	1 (1)	0 (1)	2 (2)	0 (4)	0	12
S	**12**	12	0	0	0	0	0	0
SE	**5**	5	0	0	0	0	0	0
P	**100**	96	1	0	0	0	0	3
PE	**142**	131	0	0 (1)	2 (2)	0 (3)	0	9
E	**2**	2	0	0	0	0	0	0
F	**11**	11	0 (1)	0	0	0 (1)	0	0
U[a]	**5**	5	0	0	0	0	0	0
O	**8**	6	1	0	0	0 (1)	0	1
ERO	**7**	0	0	0	0	0	7	0
B	**1**	1	0	0	0	0	0	0
Miss	**1**	0	1	0	0	0	0	0

Notes. GC09 primary classification counts are given for every GC09 class. When applicable, secondary classification counts are provided in parentheses. The total number of sources in the spectral groups assigned in this paper are provided in the column labeled "Total," and the total number of sources in each GC09 class are provided in the row labeled "Total." Unbolded numbers show the overlap between the two classification schemes.
[a] Objects in the U Group are not able to be classified due to incomplete spectral coverage.

to embedded sources which are found to reside in molecular clouds ~80% of the time and virtually always sit in a larger cloud of dust. The five O Group sources with flatter spectra may be highly evolved pre-main-sequence stars such as a Herbig Ae/Be stars—all are associated with either dusty filaments or larger dusty structures, and four of the five are located within molecular clouds, suggesting they still reside with the clouds they formed from.

The seven objects in the second non-embedded group are the previously mentioned EROs. Their spectra are a mostly featureless blackbody continuum consistent with warm dust emission. Six of the spectra show a broad absorption feature at 11 μm coinciding with the SiC feature seen in Galactic extreme carbon stars. All the EROs additionally have a C$_2$H$_2$ absorption feature at 13.7 μm, confirming the C-rich nature of these objects (Gruendl et al. 2008).

The final group (Group B) contains only one source, 054054.25−693318.7. It has a step-like appearance produced by broad emission features at 8, 11, 21, and 25 μm superimposed on a continuum that peaks at ~30 μm. Previously identified sources with similar spectra have been classified as F or G supergiants or possible protoplanetary nebulae (Kraemer et al. 2002). Its identification as a non-YSO is strengthened by the fact it is not located within a molecular cloud.

3.5. Comparison to Previous Classifications

The catalog from which our sources were chosen (GC09) was developed via multiwavelength photometry and image inspection. We compared our classifications with those originally proposed by GC09. Table 2 summarizes the results of this comparison. Of the 294 sources in this study, 269 had been identified as Definite or Probable YSOs by GC09. Of the 269 candidate GC09 YSOs, 262 of them have here been confirmed spectroscopically to have spectra indicative of embedded YSOs. Six

remaining sources are members of the O Group which may be Herbig Ae/Be stars, AGBs, or post-AGBs, and the seventh is a member of the B Group. Most LMC AGBs are excluded from the GC09 candidate YSO catalog based upon mid-IR colors bluer than YSOs. GC09 identified other candidate AGBs by their SEDs which tend to peak in the IRAC bands. We find that the spectra of members of the O Group are generally flat, explaining why they may have not been excluded from the GC09 candidate YSO catalog. Clearly, with a success rate of $\gtrsim 95\%$, the CMD, SED, and image inspection criteria used by GC09 to distinguish sources with YSO-like spectra from non-YSOs constitute a robust procedure.

GC09 gave sources 052219.71−654319.0, 052335.54−675235.6, and 054033.97−692509.9 the primary classification of AGB star; additionally, 045245.62−691149.5 received a secondary classification of AGB. Source 052335.54−675235.6 has strong PAH features and is likely an embedded YSO. This object has been previously identified in the literature as IRAS 05237−6755, is listed in the online database SIMBAD[7] as a star, and has been studied previously as a candidate AGB. We agree with previous studies by Reid et al. (1990) and Loup et al. (1997) that concluded via color criterion that this source is not an AGB. Source 045245.62−691149.5, a member of the F Group, has a mostly featureless spectrum with extremely weak silicate absorption and even weaker PAH emission. GC09 classified it as a "Probable" YSO but could not rule out the possibility of it being an AGB. Previously identified as an AGB by Wood et al. (1992), we find that it has spectral properties more consistent with a massive embedded YSO. This is further supported by its location within both a molecular cloud and a complex of dusty material seen prominently in the 5.8 and 8.0 μm IRAC images. GC09 gave 054033.97−692509.9 a primary classification of AGB, and we can confirm that it has the spectroscopic properties consistent with an evolved star. Finally, the spectrum of 052219.71−654319.0 is likely a result of a mispointing by *Spitzer*; its spectrum is completely featureless and is unlike any other previously identified spectra.

Five of our observed objects are GC09 candidate YSOs with a secondary classification of normal star. All five of these sources have now been spectroscopically confirmed to have spectra indicative of embedded YSOs. One GC09 candidate YSO, 045713.47−661900.1, has a secondary classification of background galaxy. Here, we have classified it in the PE Group. The 13 sources we observed left unclassified by GC09 due to the lack of data in at least one IRAC band are classified as follows: three are classified in the P Group, nine in the PE Group, and one in the O Group. All seven EROs were photometrically identified by GC09 and have here been confirmed spectroscopically to represent a rare class of evolved star.

Two sources have GC09 primary classifications of "diffuse" source. Diffuse sources have SEDs and images consistent with local enhancements of dust emission (see Section 5.4 of GC09). Possible characteristics of diffuse objects include extended or non-point-source IRAC images and the absence of a near-IR or MIPS counterpart. Since diffuse sources have IRAC photometry that is dominated by PAH emission, it is unsurprising that the four diffuse sources were here classified in the PE Group. It is probable that both these sources truly are only local enhancements of the ISM and not YSOs. Sources 045205.39−665513.8 and 052204.85−675800.3 have no optical or near-IR counterparts, and both are located on prominent dust filaments seen

in IRAC images and are extended in the direction of the filament. Two other sources have GC09 secondary classifications of diffuse sources. The two sources, 053554.94−693903.0 and 054137.57−711902.0 are spatially extended in the mid-IR, although both have tentative near-IR counterparts—a weak K-band counterpart exists for the former, and a K-band point source with a coincident compact Hα emission region is slightly offset from the latter. We concur with GC09 that 045205.39−665513.8 and 052204.85−675800.3 are likely diffuse sources, while 053554.94−693903.0 and 054137.57−711902.0 may be YSOs. It is important to note that these likely diffuse objects are spectroscopically indistinguishable from YSOs based solely upon their IRS spectra. Clearly, a careful inspection of SEDs, multiwavelength images, and spectra is required to separate sources which may be YSOs from those that are not.

The GC09 candidate YSOs are broadly separated into two luminosity categories—those with $[8.0] > 8.0$ and those with $[8.0] < 8.0$. The 8.0 μm magnitude is likely a good first-order proxy for a YSO's mass so that the brightest sources ($[8.0] < 8.0$) are high mass, while the dimmer sources are less massive. This is supported by several recent studies. GC09 found that the luminosity function of the $[8.0]$ band of their Definite and Probable YSOs with $[8.0] < 8.0$ is consistent with recent power-law determinations of the high-mass part of the initial mass function (see Figure 17 of GC09). Chen et al. (2009) found in a case study of massive YSOs in the LMC's N44 that SEDs of YSOs with $[8.0] < 8.0$ can be well fit by radiative transfer models where the central YSO has a mass $\gtrsim 9$ $M_\odot$ (Table 7 of Chen et al. 2009). But their results also demonstrate that good fits to the SEDs can be achieved by using a variety of central source masses. Thus, the direct application of YSO masses based upon luminosities is likely most valid when considering a large sample of sources. We observed a total of 38 objects with $[8.0] > 8.0$. Of those, 36 were classified as definite YSOs by GC09 and two as probable YSOs. The dimmest sources do not show any large biases toward a particular spectral feature. The 38 $[8.0] > 8.0$ sources are classified by PCA as follows: one source in the S Group (3%), 16 in the P Group(42%), and 21 in the PE Group (55%). This distribution is consistent within the limits of small number statistics with the entire PCA-based classification catalog which is composed of 4% S Group, 2% SE Group, 37% P Group, 52% PE Group, 1% E Group, and 4% F Group sources.

4. SPECTRAL EVOLUTION OF AN EMBEDDED MASSIVE YSO

Since massive YSOs can take thousands of years to form, it is observationally impossible to watch a single star develop. Instead, studies are done statistically on a large sample of sources. Our catalog of sources represents the most reliable census of massive YSOs in the LMC, and given the large number of sources and the unbiased nature of their identification, we have captured YSOs at various stages of evolution. This allows us to develop a scheme for the evolution of massive YSO spectra and to tie the spectra to the astrophysics involved in star formation. Here, we primarily focus on the evolution of the most massive YSOs, and so are only concerned with those embedded sources with $[8.0] < 8.0$. We refer to these sources as massive based upon the rough mass lower limit described in Section 3.5. This mass cutoff is certainly not exact, and it should be noted that when we refer to massive YSOs, we are referring not to a mass-based cutoff, but a luminosity-based one.

---

[7] http://simbad.u-strasbg.fr/simbad/

4.1. Complications of the Spectral–Evolutionary Connection

The assessment of a YSO's evolutionary state based upon its IRS spectra is primarily complicated by the possibility of spectral features originating from regions external to the YSO's immediate star–disk–envelope system. Given the large size of the IRS slits (as wide as 11″.1–2.6 pc at the LMC distance), any radiation from an object's environment that falls within the slit will contribute to its spectrum. As evidenced by the apparent spatially extended nature of these objects, we suspect this type of spectral contamination from the environment may play an important role.

In particular, there may be contamination from fine-structure lines that originate external to the YSOs. To better assess the origin of the fine-structure lines, we have compared the strength of a fine-structure line in the spectrum of the IRS's full-slit to the strength of the same line when a spectrum is extracted at the position on the slit farthest from the source. The SH module has a slit size of 5×2 pixels; in the first nod position, the source is centered on the first 2×2 pixels, while in the second nod position the source is centered on the last 2×2 pixels. We have therefore extracted a spectrum for every object with fine-structure lines along a 1 pixel wide strip on the side of the slit opposite the source. At the distance of the LMC, the "off-source" position is ∼2 pc from the object. Nominally, the 1×2 pixels at these slightly "off-source" positions would not detect emission from a point source. Consequently, if the target source is coincident with a compact fine-structure line-emitting object, then there is a concentration of fine-structure emission on the source, and the equivalent width per pixel of a fine-structure line from the full slit spectrum will be larger than the equivalent width per pixel at the "off-source" slit position. However, if the ratio of the fine-structure line's equivalent width per pixel of the full slit and the "off-source" position is equal to or less than one, the full slit spectrum used in our classification is likely contaminated by strong fine-structure line emission external to the source. The [S III] line at 18.7 μm is ideal for this investigation since its location on a relatively flat region of the spectrum allows for an easy determination of the equivalent width. We have calculated this ratio for every fine-structure line source (SE, PE, and E Groups) and have provided this [S III] concentration ratio in Table 1. It may be that objects with [S III] concentration ratios of nearly or less than unity are being misidentified as YSOs capable of ionizing their surroundings and producing fine-structure lines. Given the spatial resolution of the IRS, determining the exact nature of each source may require the less contaminated spectra that will be possible with the *James Webb Space Telescope* (Gardner et al. 2006).

To test the severity of environmental contamination from the presence of external dust within the slit, we compared the flux values at 24 μm from background subtracted MIPS photometry and IRS spectroscopy. Comparing the 24 μm fluxes in this manner can be a measure of the radiative contamination from thermally emitting dust. We find that the sources with the smallest photometric to spectroscopic ratios are the dimmest sources. This implies an expected finding, that environmental contamination from dust is the most severe for the dimmest sources.

A second way the environment can affect an object's spectra is through the external pumping of UV radiation onto the outer regions of the object, a situation that is unrecognizable using the [S III] concentration ratio. Such a scenario is certainly the case for the previously discussed diffuse sources 045205.39−665513.8 and 052204.85−675800.3. We demonstrated that these sources identified as local enhancements of dust—which are without a central heating source—can have spectra that mimic YSOs with photodissociation regions and compact H II regions. The PAHs within the local dust knot are most likely being UV-pumped from neighboring O and B stars, while the fine-structure lines likely originate from the larger H II complexes in which they reside. Similar external radiation effects almost certainly contaminate the spectra of many objects in our sample.

Given that a majority of our sources are located within or on the edges of large H II complexes, spectral contamination from ionized gas is expected to be prevalent. This problem is difficult to identify due to the limited spatial resolution of the observations and is harder to correct since the radiation is a phenomenon local to the YSO and cannot be corrected by background subtraction. For every source, we have investigated its environment through careful examination of Hα, IRAC, MIPS, and CO maps. We obtained the Hα images from the Magellanic Cloud Emission Line Survey (MCELS; Smith & The MCELS Team 1999) and downloaded available data from the NANTEN CO($J = 1$–0) survey of the LMC (Fukui 2002). The results of this examination are given in Table 1. Each source is checked for four different environmental situations: whether or not it is located within a molecular cloud delineated by CO data, possible coincidence with a small (<1 pc radius) Hα source, if it is located within a larger diffuse Hα emission region, and whether or not it is located within large-scale dust complexes including filaments, rings, or more irregular shapes. We find that ∼80% are located within CO clouds, approximately 1/3 of the sources are coincident with a compact H II region, nearly every embedded source is located within a dusty structure, and ∼3/4 are within regions of diffuse Hα emission. A complete assessment of an object's true nature must include considerations of its spectrum along with possible environmental sources of spectral contamination.

It is important to note that many of the 277 embedded object spectra likely are the combined spectra of several YSOs in a tight cluster that are spatially unresolvable by *Spitzer* IRS. Should there be one source that dominates the IR-spectrum, then the following analysis is fairly strong, but if there are multiple equally contributing YSOs, the analysis is less robust. We have inspected near-IR images of every object to look for signs of multiplicity. We obtained J and K_s band observations from the Two Micron All Sky Survey (2MASS; Skrutskie et al. 2006) and with the higher resolution Infrared Side Port Imager (ISPI; van der Bliek et al. 2004) camera on the Cerro Tololo Inter-American Observatory (CTIO) Blanco 4 m telescope. Table 1 reports our results. Sources are considered candidates for multiplicity if they clearly show several point sources or are extended in the J and K_S bands where IRAC images show a single source. A total of 162 embedded sources (58%) are identified as being possible multiple sources, in good agreement with the 65% finding of Chen et al. (2009) within the N44 complex. However, a majority of the possible multiple sources appear to have a single source that dominates the luminosity, a finding that is consistent with the known steep power law between mass and luminosity that allows massive stars to dominate the luminosity of a cluster. We conclude that treating the YSOs as single massive YSO systems is valid to first order.

4.2. Evolutionary Groups

The objects in the S Group represent the earliest stages of a massive YSO's life after attaining enough mass to be included

in our sample. The extremely deep absorption from silicates indicates that the sources are highly embedded in dense envelopes of material, with the deepest absorptions originating from the most-embedded sources. An alternate view is that the strength of the silicate absorption depends on the sight line from which the YSO is being viewed. Because of the inherent high radiation pressure that can repel accreting material, massive star formation is poorly described by a spherically symmetric envelope model. Instead, the most recent models have included axis-symmetric (Yorke & Sonnhalter 2002; Whitney et al. 2003a, 2003b) or even asymmetric, "clumpy" envelope geometries (Krumholz et al. 2009; Indebetouw et al. 2006). In the case of axis-symmetric models, the strength of the silicate feature will depend on the viewing angle with respect to a YSO's bipolar outflow. Edge-on (parallel to the plane of the disk) sight lines show prominent silicate absorption profiles, while highly inclined sight lines down the "throat" of a bipolar outflow show little or no silicate absorption (Robitaille et al. 2006). Similarly, Indebetouw et al. (2006) showed that the 10 μm silicate feature can vary greatly in strength between different lines of sight through a clumpy envelope. In fact, the voids between the clumps allow near IR radiation to escape the envelope, resulting in little or no silicate absorption along a sight line through such a void. Given that recent modeling predicts that a significant percentage of sight lines produce silicate absorption features, the small number of objects with silicate absorption relative to the number without it suggests that silicate absorption strengths are likely not exclusively line of sight dependent. Although it is possible that S and SE Group objects may be identified as P, PE, or E Group objects if viewed from a different location, it should be noted that the most embedded YSOs, those enshrouded in the most material, are more likely to have sight lines through either dense clumps (in the asymmetric models) or the densest regions of the envelope not cleared by outflows (in the axis-symmetric models).

Nine of the 11 massive objects in the S stage additionally have a prominent 15.2 μm CO_2 ice feature, likely originating from the cold outer region of the YSO's envelope. These absorption features resemble those seen in galactic sources W33 A (Gibb et al. 1998) and Elias 29 (Boogert et al. 2000), both YSOs deeply embedded in the clouds from which they formed. While silicate absorptions can arise from diffuse cloud spectra, features identified with ices are seen exclusively in dense clouds; with ice sublimation temperatures of $\sim$100 K or less, the absorption must occur in the outer, colder regions of the stellar envelope (van Dishoeck 2004). There is a large variation in the strength of the ice band, and it is therefore an excellent diagnostic of environmental changes. In particular, the large number of S Group sources with ice bands and the scarcity of sources in other groups with ice features demonstrate a gradual heating of the protostellar envelope. While 70% of the S and SE sources have a prominent CO_2 absorption, only $\lesssim$10% of the later-stage objects display the same feature. This finding supports the idea of gradual envelope heating due to an increase in the ratio of stellar radiation to circumstellar mass resulting from the dissipation of the envelope from accretion, outflow, and stellar wind. There also appears to be differences between sources in the line shape of the ice feature, a variation that may reflect evolutionary changes within the S Group itself.

Members of the SE Group likely represent a more-evolved form of the highly embedded sources in the S Group. Fine-structure lines likely originate from compact or ultracompact H ii regions. Small H ii regions can form around single massive YSOs after they begin producing a copious amount of UV radiation. SE objects are still highly embedded in circumstellar material but also show indications of a compact H ii region. This suggests that the creation of a compact or ultracompact H ii region can begin very early in the life of a YSO, before it has cleared much circumstellar material. The strength of SE Group fine-structure lines varies by several orders of magnitude between sources. If the emission is originating from a small H ii region surrounded by a larger gas and dust shell, then the line emission may be attenuated by the intervening material by an amount determined by the optical depth of the shell. We have compared six different fine-structure line strengths for the five SE sources to the strength of the silicate feature, a measure of the column density of the surrounding material. Silicate feature optical depths are determined by fitting the spectra to model silicate spectra using the IDL PAHFIT code (Smith et al. 2007). For SE objects, we find a very strong inverse correlation between the fine-structure line strengths and the silicate optical depth, suggesting that the fine-structure emission is being screened and attenuated by the surrounding dust shell. However, given the small number of sources in the SE Group (5), the implicit assumption of equal intrinsic line strengths (intrinsic line strengths will vary between sources with differing ionization fluxes, densities, region diameters, and abundances), and the fact that the correlation disappears after including sources of all groups, the correlation may not represent a true phenomenon. If the fine-structure–silicate correlation within the SE Group is real, then the lack of a correlation among the other groups implies that either the optical depths of the dust shells are not large in these groups, or that the fine-structure emission in the other groups does not originate from within a dust shell at all.

Because SE objects show both the characteristics of very young (S Group) and more-evolved (E and PE Group) objects, SE sources could represent the transition from a highly embedded to a less-embedded YSO, objects that are beginning to clear away their parent molecular cloud. Since members of the S Group show no signs of UV radiation, they are either true massive protostars (accreting, pre-hydrogen-burning massive YSOs) or are accreting at a rate and with a geometry that is able to contain the UV radiation. In either scenario, the massive YSO is unable to substantially disturb its natal cloud in the S stage. Although there are only 17 total sources in either the S or SE Groups, 52 sources ($\sim$20%) pass the PCA test for silicate absorption. A majority of these objects are members of the P and PE spectroscopic groups, both of which have features indicative of UV-emitting central objects. As discussed earlier, several PAH emission bands are superimposed on the 10 μm silicate absorption feature, making the strength of the silicate feature difficult to quantify in P and PE spectra. Even in the scenario that PCA is overestimating the number of silicate absorption sources by improperly separating the blended silicate and PAH features, sources with strong silicate absorption represent a minority of sources. Assuming a typical formation time of 10^5 years (Zinnecker & Yorke 2007), we estimate that it takes a massive YSO on the order of a few thousand to a few ten thousand years to clear its surroundings to the point where the silicate absorption feature is no longer apparent in its spectrum.

Three of the five SE objects are coincident with small compact/ultracompact H ii regions seen in the Hα images of MCELS. The presence of a compact Hα source strengthens our assertion that the fine-structure lines present in the spectra originate from a small region ionized by the source itself.

The two objects without corresponding compact H II regions, 045206.97−665516.4 and 053536.88−691214.1, are both located within areas of high Hα emission from a larger, more extended structure. It is therefore possible that the fine-structure lines do not originate from the source itself and were not properly background subtracted. In fact, the two sources without a visible coincident compact Hα source also have the lowest [S III] concentration ratios of the SE Group. Object 045206.97−665516.4 has a ratio less than 1, indicating that the strength of the [S III] line is stronger at distances farther from the source and that the fine-structure line emission is likely at least in part either an environmental contamination or the result of an extended fine-structure line-emitting structure. Although the sample size is small, there seems to be a correlation between the presence of a compact Hα source and the existence of fine-structure lines in the spectra—60% of the SE sources have accompanying compact H II region, compared with only 25% of the S Group objects. Of the 10 massive S objects, only three, 045111.39−692646.7, 045550.62−663434.6, and 052646.61−684847.2 show any sign of compact Hα emission on the source, and the emission is relatively faint.

The P, PE, and E Groups likely represent the later stages of a YSO's embedded lifetime. Of the 239 massive embedded objects, 217 of them have either PAH emission, fine-structure lines, or both. Since both PAH and fine-structure lines are indicative of a UV-emitting central object, massive YSOs spend a substantial portion of their lives emitting in the UV. Objects with UV-produced features represent ∼90% of our catalog, suggesting that massive embedded YSOs emit in the UV during a substantial proportion of their lives, possibly as high as 90%. However, note that given the possibility of environmental spectral contamination, the length of this stage may in fact be shorter.

The strengths of fine-structure lines vary by up to 6 orders of magnitude between sources, and PAH emission power can vary by as much as 5 orders of magnitude. If fine-structure line and PAH band photons travel along a line of sight through dense circumstellar material, the intrinsic line strengths may be attenuated by a factor set by the optical depth at each line's wavelength. Given a high enough optical depth, objects with inherently strong PAH or fine-structure line emission may have the features completely attenuated by surrounding material, causing them to be improperly categorized in our spectral classification scheme. As previously discussed, we find no correlation within the P, PE, or E Groups between any of the fine-structure lines and the optical depth. Similarly, there appears to be no recognizable correlation between the 7.7 μm and 11.2 μm PAH bands' strengths and the strength of the silicate absorption feature. We fit these PAH features using Lorentzian profiles on a first-order polynomial baseline. This suggests that the optical depth of dust surrounding the fine-structure line and PAH-emitting regions are generally not large. We conclude and that viewing angle does not have a large effect on the strength of PAH bands or fine-structure lines, and should not effect classification.

Among the 144 PE and E Group objects, the [S III] line concentration ratios range between 0.9 and 408 with a median value of 1.4. A median value greater than 1 indicates that for a large majority of our spectra, the fine-structure line is stronger at slit positions closer to the source, suggesting that most the PE and E objects have fine-structure line peaks on the source and are less likely to be solely environmental in nature. However any object with a concentration ratio near or less than 1 is more likely to be affected by the local environment. There are a total of four sources with [S III] concentration ratios less than 1.0; three of these are located within large complexes of strong Hα emission and lack a compact Hα object. This implies that the fine-structure lines of these sources are more likely environmental in nature. Of all the sources located within large H II complexes, the median [S III] concentration ratio is 1.36. Conversely, sources not located within larger H II complexes tend to have larger [S III] concentration ratios: the median [S III] concentration ratio of sources not located within large H II complexes is 3.37. The strong correlation between a source's location within an H II complex and its [S III] concentration ratio suggests that this ratio is a good measure of the amount of environmental fine-structure line contamination. Note that ratios greater than 1.0 are strongly favored, even with sources within H II complexes. This implies that although environmental fine-structure line contamination likely plays a heavy role, most sources have ratios consistent with being coincident with a compact fine-structure line-emitting object. PE and E Group objects with compact source-like [S III] concentration ratios represent approximately one-half the sample of massive YSO candidates in this study. A compact or ultracompact H II region is likely to have formed around half of the YSOs in our catalog, suggesting that massive YSOs can create a region whose fine-structure lines are *Spitzer* IRS-detectable approximately halfway through the massive stages of its formation.

The spectra of P and PE Group objects resemble those typical of photodissociation regions (PDRs); in our sample of massive YSOs, the PAH emission pumped by the YSO's own UV radiation probably originates from such a region around the forming star. High spatial resolution ISO observations of ρ Oph W within the ρ Oph molecular cloud indicate that the PAH bands and fine-structure lines are common to hot dust illuminated by O or B-type stars (Boulanger et al. 2000). The central object heats its surroundings to high temperatures, dissociating molecules, and exciting PAH vibrational levels leading to fluorescent PAH emission. The objects in the P stage, when embedded YSOs display only strong PAH emission, may be younger than those of the PE or E stages. Although both PAHs and fine-structure emission lines are indicators of UV radiation, PAHs can absorb and reprocess longer wavelength radiation than is necessary to create an H II region. In fact, PAH spectra are produced most efficiently when the incident UV radiation is intense but relatively soft—even below the Lyman limit. P Group objects may then represent YSOs who are hot enough to produce soft UV radiation, but have not yet become hot enough to produce many higher energy photons that can ionize its surroundings. The presence of PAH emission and lack of fine-structure lines can also be explained if a hot, hard UV-producing star's radiation is being quenched by a high accretion rate, preventing a detectable compact H II region from forming. In either scenario, objects in the P stage therefore are the transition objects that bridge objects that cannot form detectable compact H II regions and those that can. Externally produced PAH emission is likely a large contaminant to both the P and PE spectra, but even given this environmental effect on the spectra, the absence of fine-structure lines in the P Group indicates a source unable to produce a detectable compact H II region.

Seventeen of the 100 total P objects passed the PCA test for silicate absorption, indicating they may have a silicate feature at 10 μm. These may be the youngest of the P Group, and could represent a subgroup of sources that have just begun clearing

their surroundings and are still embedded in a thick molecular envelope. Alternatively, these objects could in fact be examples of objects whose line of sight passes through a dense region of the envelope, creating a strong silicate feature. The fact that there is no recognizable correlation between either PAH or fine-structure line strength and optical depth also suggests that the PAH features and fine-structure lines are not formed in a region inside a shell of circumstellar material as would be expected in a spherically symmetric model. This is consistent with previously described models which allow radiation to scatter and escape via outflows or voids between clumps. In such a scenario, the $10\,\mu$m absorption feature originates from a dense cloud or clump that is both self-screening and attenuating radiation originating from behind the cloud/clump, while PAH and fine-structure line emission can originate from regions external to this line of sight (e.g., outflow cavity walls, other clumps, extended H II regions) that are being illuminated/ionized by the central source.

PE Group objects represent the most evolved form of YSO in our proposed evolutionary scheme. PAH emission and fine-structure lines indicate that a high level of UV radiation is originating from the central source and has created a compact H II region surrounding the object. While the fine-structure lines originate from this H II region, the PAH emission likely originates from the PDR surrounding the H II region. PAHs are depleted within the boundary of the H II region, destroyed by the intense hard-UV radiation. P Group objects may be an extreme PE source—where a full-sized compact H II region cannot form due to high accretion rates but a PDR can form outside the very small central ionized region. E Group objects likely represent a type of PE object, ones with strong fine-structure lines and relatively weak PAH emission. Although the two objects in the E Group do not pass either PCA test for PAH emission, by eye inspections of both spectra show clear signs of PAH emission. Both the objects may be better classified into the PE Group or as extreme PE sources where the PDR is relatively small.

The F Group is comprised of 11 objects that did not pass any spectral feature's PCA test. Seven may show signs of weak silicate absorption at $10\,\mu$m, seven may have fine-structure lines, and all 11 may have weak PAH emission. It may be that these sources are better classified into other groups, but because PCA was used to quantify the strengths of spectral features, they have been dubbed "featureless." Since they represent a small fraction of the total sample (6%) and do not appear biased toward a particular feature, we do not consider their contribution to the above evolutionary analyses statistically important. The five sources without full spectral coverage are assigned to the U Group.

5. CONCLUSION AND FUTURE DIRECTIONS

We have obtained high- and low-resolution IR spectra of 294 objects in the LMC. The sources were chosen from a catalog of IR-luminous objects developed by GC09 based on their massive YSO-like photometric and SED qualities. The *Spitzer* IRS spectra cover the wavelength range of 5–37 μm and show a variety of spectral features including absorption from silicates and ices and emission lines from PAHs and ionized atoms. Of the 294 sources observed, 277 spectroscopically resemble embedded massive or intermediate-mass YSOs. The remaining sources are comprised of a mixture of Herbig Ae/Be stars (with O-rich emission disks) or highly evolved stars such as RSGs, AGBs, and post-AGBs.

We classified both the YSOs and non-embedded sources into distinct groups based upon their spectral features. Embedded

objects were categorized automatically via the statistical process dubbed PCA. This allowed for an unbiased and quantitative classification scheme. Non-embedded sources—primarily evolved stars, and not the primary focus of this paper—were classified into groups by eye. The embedded sources all rise into the red end of the spectrum indicating their continuum emission is dominated by dust. Most embedded sources are identified as massive YSOs still surrounded by the dense cloud from which they formed. Our catalog largely consists of sources with luminosities consistent with YSOs of at least ~10 $M_\odot$, and therefore represent the later stages of massive star formation, after they have accreted a significant fraction of their mass. Individual spectra within the large sample of sources act as snapshots of massive star formation and collectively display an evolution of massive YSO spectra. The youngest sources (S and SE Groups) are dominated by deep silicate absorptions and ice absorptions originating from the cold outer regions of the envelope (van Dishoeck 2004). Later stages of massive YSO evolution show a gradual heating of the circumstellar envelope evidenced by the lack of ice features. Older sources also display signs of envelope depletion in the form of little or no silicate absorption. As the envelope is depleted and accretion slows or ceases, UV radiation from the young star begins to dissociate and ionize the nearby molecular material forming compact H II regions and PDRs (e.g., Hoare et al. 2007). Well known indicators of compact H II regions and PDRs are present in the spectra of later-stage massive YSOs (P, PE, and E Groups) including fine-structure emission lines and PAH emission features.

The work described here is a first step in a series of investigations using this newly formed YSO catalog into the evolution of massive YSOs and the interaction between them and their environment. Follow-up papers will address both the inter- and intra-group evolution of specific features. YSO evolution will also be tied to the state of the molecular cloud they were formed from and the nearby stellar population. Finally, work should be done to compare the spectra of these massive YSOs and their corresponding evolutionary stage with radiation transfer models to assess both the quality of the models and the accuracy of their spectral-based evolutionary placement.

REFERENCES

Beuther, H., Churchwell, E. B., McKee, C. F., & Tan, J. C. 2007, in Protostars and Planets V, ed. V. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 165

Boogert, A. C. A., Tielens, A. G. G. M., Ceccarelli, C., Boonman, A. M. S., van Dishoeck, E. F., Keane, J. V., Whittet, D. C. B., & de Graauw, T. 2000, A&A, 360, 683

Book, L. G., Chu, Y.-H., Gruendl, R. A., & Fukui, Y. 2009, AJ, 137, 3599

Boulanger, F., et al. 2000, in ISO Beyond Point Sources: Studies of Extended Infrared Emission, ed. R. J. Laureijs, K. Leech, & M. F. Kessler, ESA, SP-455, 91

Chen, C.-H. R., Chu, Y.-H., Gruendl, R. A., Gordon, K. D., & Heitsch, F. 2009, ApJ, 695, 511

Chu, Y.-H., et al. 2005, ApJ, 634, L189

Churchwell, E. B., Indebetouw, R., Benjamin, R. A., & The GLIMPSE Team 2002, BAAS, 34, 1188

Connolly, A. J., Szalay, A. S., Bershady, M. A., Kinney, A. L., & Calzetti, D. 1995, AJ, 110, 1071

Fazio, G. G., et al. 2004, ApJS, 154, 10

Feast, M.999, in IAU Symp. 190, New Views of the Magellanic Clouds, ed. Y.-H. Chu, et al. (Dordrecht: Kluwer), 542

Francis, P. J., Hewett, P. C., Foltz, C. B., & Chaffee, F. H. 1992, ApJ, 398, 476

Fukui, Y. 2002, in IAU Symp. 207, Extragalactic Star Clusters, ed. D. Geisler, E. K. Grebel, & D. Minniti (San Francisco, CA: ASP), 505

Garay, G., & Lizano, S. 1999, PASP, 111, 1049

Gardner, J. P., et al. 2006, Space Sci. Rev., 123, 485

Gerakines, P. A., et al. 1999, ApJ, 522, 357

Gibb, E. L., Gerakines, P. A., & Whittet, D. C. B. 1998, BAAS, 30, 1355

Gruendl, R. A., & Chu, Y.-H. 2009, ApJ, submitted

Gruendl, R. A., Chu, Y.-H., Seale, J. P., Matsuura, M., Speck, A. K., Sloan, G. C., & Looney, L. W. 2008, ApJ, 688, L9

Hartmann, L., Megeath, S. T., Allen, L., Luhman, K., Calvet, N., D'Alessio, P., Franco-Hernandez, R., & Fazio, G. 2005, ApJ, 629, 881

Higdon, S. J. U., et al. 2004, PASP, 116, 975

Hoare, M. G., Kurtz, S. E., Lizano, S., Keto, E., & Hofner, P. 2007, in Protostars and Planets V, ed. V. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 181

Houck, J. R., et al. 2004, ApJS, 154, 18

Indebetouw, R., Whitney, B. A., Johnson, K. E., & Wood, K. 2006, ApJ, 636, 362

Kennicutt, R. C. 2005, in IAU Symp. 227, Massive Star Birth: A Crossroads of Astrophysics, ed. R. Cesaroni, et al. (Cambridge: Cambridge Univ. Press), 3

Kessler, M. F., et al. 1996, A&A, 315, L27

Keto, E. 2007, ApJ, 666, 976

Kraemer, K. E., Sloan, G. C., Price, S. D., & Walker, H. J. 2002, ApJS, 140, 389

Krumholz, M. R., & Bonnell, I. A. 2007, arXiv:0712.0828

Krumholz, M. R., Klein, R. I., McKee, C. F., Offner, S. S. R., & Cunningham, A. J. 2009, Science, 323, 754

Kwon, W., Looney, L. W., Crutcher, R. M., & Kirk, J. M. 2006, ApJ, 653, 1358

Looney, L. W., Tobin, J. J., & Fields, B. D. 2006, ApJ, 652, 1755

Loup, C., Zijlstra, A. A., Waters, L. B. F. M., & Groenewegen, M. A. T. 1997, A&AS, 125, 419

Luhman, K. L., & Rieke, G. H. 1999, ApJ, 525, 440

McCray, R., & Kafatos, M. 1987, ApJ, 317, 190

McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565

Meixner, M., et al. 2006, AJ, 132, 2268

Murtagh, F., & Heck, A. 1987, Astrophysics and Space Science Library (Berlin: Springer), 131

Reid, N., Tinney, C., & Mould, J. 1990, ApJ, 348, 98

Rieke, G. H., et al. 2004, ApJS, 154, 25

Robitaille, T. P., Whitney, B. A., Indebetouw, R., Wood, K., & Denzmore, P. 2006, ApJS, 167, 256

Seale, J. P., & Looney, L. W. 2008, ApJ, 675, 427

Shepherd, D. S., & Churchwell, E. 1996, ApJ, 472, 225

Skrutskie, M. F., et al. 2006, AJ, 131, 1163

Smith, R. C., &The MCELS Team 1999, in IAU Symp. 190, New Views of the Magellanic Clouds, ed. Y.-H. Chu, et al. (Dordrecht: Kluwer), 28

Smith, J. D. T., et al. 2007, ApJ, 656, 770

Soifer, B. T., Russell, R. W., & Merrill, K. M. 1976, ApJ, 210, 334

Tenorio-Tagle, G., & Bodenheimer, P. 1988, ARA&A, 26, 145

Tielens, A. G. G. M., Peeters, E., Bakes, E. L. O., Spoon, H. W. W., & Hony, S. 2004, in ASP Conf. Ser. 323, Star Formation in the Interstellar Medium: In Honor of David Hollenbach, Chris McKee and Frank Shu, ed. D. Johnstone, et al. (San Francisco, CA: ASP), 135

Tobin, J. J., Looney, L. W., Mundy, L. G., Kwon, W., & Hamidouche, M. 2007, ApJ, 659, 1404

van der Bliek, N. S., et al. 2004, Proc. SPIE, 5492, 1582

van Dishoeck, E. F. 2004, ARA&A, 42, 119

van Dishoeck, E. F., & Blake, G. A. 1998, ARA&A, 36, 317

Werner, M. W., et al. 2004, ApJS, 154, 1

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

Whitney, B. A., et al. 2008, AJ, 136, 18

Wood, P. R., Whiteoak, J. B., Hughes, S. M. G., Bessell, M. S., Gardner, F. F., & Hyland, A. R. 1992, ApJ, 397, 552

Yorke, H. W., & Sonnhalter, C. 2002, ApJ, 569, 846

Zinnecker, H., & Yorke, H. W. 2007, ARA&A, 45, 481