Star Formation and Feedback in Low-metallicity Environments: From Molecular Clouds to Protostars

Theo J. O'Neill

University of Virginia
Department of Astronomy
December 2022

This thesis is submitted in partial completion of the requirements of the BS Astronomy-Physics Major.

Advisor: Prof. Rémy Indebetouw Theo J. O'Neill

Star Formation and Feedback in Low-metallicity Environments: From Molecular Clouds to Protostars

Abstract

In this thesis, we explore the relationship between molecular clouds and star formation in the Magellanic Clouds – two nearby, low-metallicity companion galaxies to the Milky Way. We first examine variations in the behavior of common tracers of molecular gas, and derive physically motivated corrections for observational biases affecting Carbon Monoxide based metrics of molecular cloud properties in low-metallicity environments. We then study molecular cloud dynamics and the progression of star formation in the young Small Magellanic Cloud star-forming region NGC 602. Finally, we explore machine-learning based techniques to identify pre-main-sequence stars in the active star-forming region N159 in the Large Magellanic Cloud.

Contents

List of figures

List of Tables

Chapter 2 appeared as:

O'Neill, T. J.; Indebetouw, R.; Bolatto, A. D; Madden, S. C; Wong, T. *Effects of CO-dark Gas on Measurements of Molecular Cloud Stability and the Size-Linewidth Relationship*. 2022, ApJ, 933, 179.

Chapter 3 appeared as:

O'Neill, T. J.; Indebetouw, R.; Sandstrom, K. M.; Bolatto, A. D.; Jameson, K. E.; Carlson, L. R.; Finn, M. K.; Meixner, M.; Sabbi, E; Sewilo, M. *Sequential Star Formation in the Young SMC Region NGC 602: Insights from ALMA*. 2022, ApJ, 938, 82.

Acknowledgments

Thank you to my advisors and mentors, Rémy Indebetouw and Kelsey Johnson, for their unwavering support during my time as an undergraduate.

Thank you to Molly Finn, Allie Costa, and the UVA Astronomy Department.

And finally, thank you to the my fellow fourth years – Cole Faggert, Ethan McKeever, Jee-Ho Kim, Jillian Maxson, Louis Seyfritz, Mint Tositrakul, and Sarah Shriner – and miscellaneous other undergrads – including AnnMarie Kelly, Alex Rosenthal, and Kayleigh Childress – without whom this thesis could not have been written.

1

Introduction

The interstellar medium (ISM) is an intricate, complex ecosystem, populated by dense gas, diffuse dust, hyperactive star-forming regions, and mysterious magnetic fields. Stars and the multiphase ISM perform a delicate dance, with each contributing to the other in a perpetual and dynamic cycle. Massive stars affect the formation of younger generations of stars through feedback in the form of stellar winds and supernovae, before recycling their components back into the ISM. In the process, the ISM drives the formation and evolution of galaxies; turbulence and compression resulting from collisions of individual molecular clouds and giant gas components hundreds of parsecs in size trigger new star formation and reshape galactic structures. Making sense of the effects of feedback is critical in understanding the evolution of stellar and galactic populations.

The ISM is deeply sensitive to the chemical makeup and enrichment of its component dust and gas. In low-metallicity (low-Z) environments, reduced dust-to-gas ratios and stronger radiation fields lead to dramatic variations in the behavior of the ISM and progression of star formation. The more deeply one wishes to probe the physics of star formation, the more detailed must ones understanding be of the physical conditions in star-forming molecular gas (temperature, density, gravitational stability, and more) across a variety of environments.

A significant fraction of the total mass of the ISM is concentrated in cold ($T \sim 10$ – 20 K), dense ($n \sim 10^2$ – 10^6 cm^{-3}) clouds of molecular gas. If these clouds fail to win their struggle against external compressing forces and their own gravitational potential, they collapse in on themselves to form dense cores and eventually young stars. Deviations from empirical scaling laws between molecular gas and star formation (e.g., the Kennicutt-Schmidt and associated relations) observed in low-Z regions can reveal important changes in physics. Understanding star formation physics in the context of these observed relationships requires being able to accurately translate observed quantities into assessments of molecular gas.

Molecular hydrogen H_2 is the most common molecule in the ISM, but is challenging to observe directly. As a result, carbon monoxide (CO) frequently takes on the role of a tracer of molecular gas. In low-Z environments, though, the fraction of H_2 gas mass that is not traced by CO is expected and observed to increase; H_2 is more effective at self-shielding itself from incident radiation than CO, which causes the $C^+/C^0/CO$ transition to retreat farther from clump edges than the HI to H_2 transition. This leaves behind a central CO-traced surface surrounded by an extended diffuse envelope of "CO-dark" gas. The varying effectiveness of CO as a H_2 tracer makes comparisons of common metrics of molecular cloud properties and star formation across environments challenging. Assessing and interpreting distant observations of the ISM in the earliest low-Z galaxies requires a comprehensive understanding the relationship between CO, H_2, and star formation in local environments.

The Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC) are two of the Milky Way's nearest and most massive companion dwarf galaxies (at distances of 50 kpc and 60 kpc, respectively). As low-metallicity (1/2 and 1/5 Solar) environments that are actively forming stars and have been significantly disrupted by interactions with each other and potentially the Milky Way, the Magellanic Clouds (MCs) present a valuable opportunity to study variations in the ISM at high spatial resolution.

In this thesis, I explore the effects of environment on the physics of molecular clouds and the progression of star formation through the lens of the SMC and LMC. In Chapter 2, I examine how metrics of molecular cloud properties and stability are affected by CO-dark gas fractions. I develop physically motivated models of the radial distribution of CO-bright vs CO-dark gas in individual molecular clumps, and show that this missing emission can significantly skew interpretations of cloud stability and evolution. I then derive analytic corrections that can be easily applied to CO observations to correct for the bias. These corrections are expected to be most important in low-metallicity and high-radiation regions like the SMC.

In Chapter 3, I apply this work to molecular clumps and the progression of star formation in the young SMC star cluster NGC 602. I make use of Atacama Large Millimeter/submillimeter Array (ALMA) observations of CO in the region to analyze the structure and dynamics of star-forming molecular clouds. I combine these data with analysis of archival photometric observations of young stars to gain insight into likely triggers for star formation in the region.

In Chapter 4, I turn to examining the development of populations of young stars in the LMC. I use machine learning methods to identify probable pre-main-sequence (PMS) stars using Hubble Space Telescope photometry, and infer spatial variations in dust extinction. I use the active star-forming region N159 as a trial group to experiment with the effectiveness of these methods.

I conclude in Chapter 5 with a comparison of the results of these studies, and extend our conclusions to differences in the process of star formation in the Milky Way vs. Magellanic Clouds.

2

Effects of CO-dark Gas on Measurements of Molecular Clouds

Draft version December 15, 2022
Typeset using LaTeX **twocolumn** style in AASTeX631

Effects of CO-dark Gas on Measurements of Molecular Cloud Stability and the Size-Linewidth Relationship

Theo J. O'Neill [1] Rémy Indebetouw [1,2] Alberto D. Bolatto [3] Suzanne C. Madden [4] and Tony Wong [5]

[1] Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA
[2] National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
[3] Department of Astronomy, University of Maryland, College Park, MD 20742, USA
[4] AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, 91191, Gif-sur-Yvette, France
[5] Department of Astronomy, University of Illinois, Urbana, IL 61801, USA

ABSTRACT

Stars form within molecular clouds, so characterizing the physical states of molecular clouds is key in understanding the process of star formation. Cloud structure and stability is frequently assessed using metrics including the virial parameter and Larson (1981) scaling relationships between cloud radius, velocity dispersion, and surface density. Departures from the typical Galactic relationships between these quantities have been observed in low metallicity environments. The amount of H_2 gas in cloud envelopes without corresponding CO emission is expected to be high under these conditions; therefore, this "CO-dark" gas could plausibly be responsible for the observed variations in cloud properties. We derive simple corrections that can be applied to empirical clump properties (mass, radius, velocity dispersion, surface density, and virial parameter) to account for CO-dark gas in clumps following power-law and Plummer mass density profiles. We find that CO-dark gas is not likely to be the cause of departures from Larson's relationships in low-metallicity regions, but that virial parameters may be systematically overestimated. We demonstrate that correcting for CO-dark gas is critical for accurately comparing the dynamical state and evolution of molecular clouds across diverse environments.

1. INTRODUCTION

Star formation is strongly correlated with tracers of molecular gas over kpc-scales (e.g., Kennicutt et al. 2007; Leroy et al. 2008; Bigiel et al. 2011), suggesting a causal relationship between the two. Since molecular clouds are the sites of star formation, understanding their dynamical states is necessary in accurately predicting star formation both in individual clouds as well as across larger populations.

Molecular hydrogen H_2 is the most abundant molecule in the interstellar medium (ISM) and is therefore closely tied to understanding the stability of molecular clouds and process of star formation. H_2 is a symmetric, homonuclear molecule with widely spaced rotational energy levels and no permanent dipole moment; as a consequence of this, it radiates very weakly and is difficult to observe directly under conditions typical of molecular clouds ($T \sim$ 10–20 K). It is therefore necessary to use more accessible molecules as tracers of H_2 to fully understand the conditions under which stars form.

CO is one of the next most abundant molecules in the ISM and can be excited easily at low temperatures, making it a popular tracer of H_2. Using CO as a tracer, the amount and spatial distribution of molecular gas in a region is often used to infer the process of star formation; however, this use of CO as a proxy for H_2 relies on the assumption that it faithfully traces the full spatial extent of H_2. It is well known that some portion of the H_2 in molecular clouds is not traced by CO: since CO is less efficient at shielding itself from FUV radiation than H_2 is, the transition from C^+ to CO occurs closer to the center of clouds than the transition from H I to H_2, resulting in a central CO-traceable region surrounded by an extended diffuse envelope of "CO-dark" H_2.

Recent studies have simulated the formation of H_2 and CO in the ISM to evaluate the expected amount of CO-dark H_2 in a variety of environments (e.g, Glover et al. 2010; Glover & Mac Low 2011; Li et al. 2018; Gong et al. 2018). Wolfire et al. (2010, hereafter W10) modeled photodissociation regions (PDRs) of individual spherical clouds and defined the fraction of molecular H_2 mass not traced by CO, or "dark gas fraction," as

$$f_{DG} = 1 - \frac{M(R_{CO})}{M(R_{H_2})}, \tag{1}$$

where $M(r)$ represents the mass contained within a radius r, R_{CO} is the radius of the CO-traceable material at which the optical depth, τ, equals 1 in the $J = 1$–0 transition, and R_{H_2} is the radius at which half of the hydrogen in the envelope surrounding the CO clump is molecular and half is atomic. This model is shown in Figure 1. Assuming standard Galactic conditions, W10 derived $f_{DG} \sim 0.3$, a result which they found to be relatively insensitive to cloud and environmental properties. Other studies both of individual cloud envelopes and at galactic scales have derived $f_{DG} \sim 0.4$, but observed a stronger dependence on environmental properties (e.g., Smith et al. 2014; Szűcs et al. 2016).

Similar values of f_{DG} have been found through observational work. In studies of individual Galactic clouds, f_{DG} has been found to be $\gtrsim$0.3 (e.g., Grenier et al. 2005; Abdo et al. 2010; Velusamy et al. 2010; Lee et al. 2012; Langer et al. 2014; Xu et al. 2016), and on galactic scales CO-dark gas has been inferred to be 0.2–0.3x as massive as the total atomic mass of the Milky Way and 1.2–1.6x as massive as its total CO-traced molecular mass (Planck Collaboration et al. 2011; Paradis et al. 2012).

The amount of CO-dark gas is expected to increase in high-radiation environments, with the $C^+/C^0/CO$ transition shifting even further into the cloud to reach higher overall column densities. Similarly, the dark-gas fraction is expected to increase in low-metallicity (low-Z) environments, where decreasing dust-to-gas ratios combine with typically stronger radiation fields to increase the efficiency of CO destruction (Madden et al. 2006; Gordon et al. 2011; Madden et al. 2020). H_2 can additionally be photodissociated via Lyman-Werner band photons, but since it can be optically thick under some A_V conditions it is able to remain self-shielded while CO is photodissociated. These effects have been supported observationally in the metal-poor outskirts of the Galaxy and in the Large and Small Magellanic Clouds (LMC and SMC, respectively, with $Z \sim 1/2\ Z_\odot$ and $Z \sim 1/5\ Z_\odot$) where $f_{DG} \gtrsim 0.8$ (Pineda et al. 2013; Jameson et al. 2018; Chevance et al. 2020).

Although much work has gone into quantifying the cause and amount of CO-dark gas in a variety of environments, the practical impact of this gas on interpretations of metrics of clump stability and evolution has not been explored in as much depth. Assessing the gravitational stability of clouds as measured by the virial parameter $\alpha_{\rm vir}$ (Bertoldi & McKee 1992) or if clouds conform to "Larson's relationships" between cloud radius, velocity dispersion, and surface density (Larson 1981) is ubiquitous in both theoretical and observational studies. In low-Z environments, departures from the typical values and relationships between these quantities for CO clouds under Galactic conditions have been observed (e.g., Bolatto et al. 2008; Hughes et al. 2013; Rubele et al. 2015; Ochsendorf et al. 2017; Kalari et al. 2020). CO-dark gas could plausibly be responsible for these variations, since f_{DG} is known to be high in these regions and cloud properties inferred from CO-traced material are not guaranteed to be representative of the overall state of the structures. Correcting for CO-dark gas may then be an essential step in evaluating the dynamical states and likely futures of molecular clouds across a range of environments.

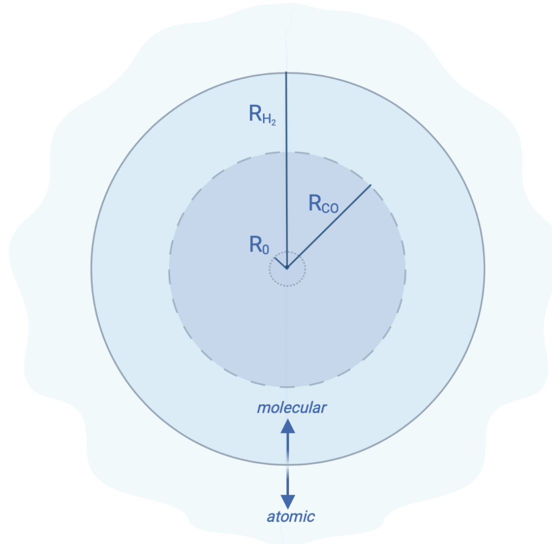

Figure 1. Clump toy model adapted from Wolfire et al. (2010). R_{H_2} is the radius at which the densities of atomic and molecular hydrogen are equal. Gas within R_{H_2} is mostly molecular and gas outside of R_{H_2} is mostly atomic. R_{CO} is the radius at which CO-traced material has $\tau = 1$ in the J=1–0 transition and is a function of f_{DG}, with higher f_{DG} yielding smaller R_{CO}. R_0 is the normalizing radius and is typically typically $\ll R_{CO}$. For a pure power-law density profile (§2.1) R_0 is arbitrary, while for a power-law profile with a core (§2.2) or Plummer profile (§2.3) it represents the radius of the flat central core.

Here we present explicitly the variation in cloud properties from what would be inferred using CO-traceable material to the "true" state of clouds including CO-dark gas. In §2 we summarize mass density profiles that clouds may follow, derive corrections for empirical clump properties to account for CO-dark gas, and explore the behavior of $\alpha_{\rm vir}$ as f_{DG} increases. We demonstrate the biases CO-dark gas creates in interpretations of size-linewidth-surface density scaling relationships in §3. We discuss the implications of our results and the effects of CO-dark gas on star formation in §4 before concluding in §5.

2. CLUMP DENSITY PROFILES

Analyzing the stability of molecular clouds ($R \gtrsim 10$ pc), clumps ($R \sim 1$ pc), and cores ($R \lesssim 0.1$ pc) is of great interest to studies of their likely evolutionary futures. To this end, Bertoldi & McKee (1992) defined the virial parameter,

$$\alpha_{\rm vir} = \frac{2\Omega_K}{|\Omega_G|} = \frac{M_{\rm vir}}{M}, \tag{2}$$

as a measure of stability, where Ω_K is the kinetic energy, Ω_G is the gravitational potential energy, M is the structure's mass, and $M_{\rm vir}$ is its virial mass. $\alpha_{\rm vir} <1$ suggests that the structure is gravitationally dominated and rapidly collapsing, $\alpha_{\rm vir} \sim 1$ indicates that a structure is gravitationally stable, and $\alpha_{\rm vir} \gg 1$ suggests that a structure is sub-critical and will likely expand unless confined by external pressure.

Variations from the expected equilibrium values of $\alpha_{\rm vir}$ have been observed in environments where f_{DG} is known to be high. In Galactic environments, $\alpha_{\rm vir}$ is frequently $\lesssim$2 (see Kauffmann et al. 2013, for a review) in clumps and clouds. In nearby low-Z dwarf galaxies and low density, low-pressure environments, $\alpha_{\rm vir}$ is frequently observed to be much larger and can reach measured values of 4–10 or more (e.g., Schruba et al. 2017, 2019). Since these environments are rich in CO-dark gas, it is possible the measured $\alpha_{\rm vir}$ could be unrepresentative of the states of full clumps, and that this additional molecular reservoir is responsible for the variations in measured $\alpha_{\rm vir}$. Alternatively, these differences could also be explained by measurement errors in σ_v and R stemming from large distance uncertainties, low velocity resolutions, or varying definitions of cloud radius.

For all clump density profiles that we will consider, we assume a one-dimensional radial velocity dispersion profile $\sigma_v(r)$ of

$$\sigma_v(r) = \left(\frac{r}{R_0}\right)^{\beta} \sigma_v(R_0), \tag{3}$$

where R_0 is a normalizing radius as shown in Figure 1. When considered in combination with a non-constant density profile $\rho(r)$, and if one considers turbulence to act as pressure support, our adopted Equation 3 leads to a gradient in energy density $\sim \rho\sigma_v^2$; we address the implications of this effect for cloud stability in §4. Additionally, we recognize that at very small scales ($\sim$0.1 pc) the effective pressure profile changes from thermal to non-thermal dominated support. Since the bulk of this work considers the effects of CO-dark gas on parsec-scales, this behavior should not impact our conclusions.

We analyze how the observationally-derived $\alpha_{\rm vir}$ depends on observed CO radius for clouds following a single power-law (§2.1), a power-law with a constant density core (§2.2), and a Plummer profile (§2.3). We derive corrections for empirical clump properties at a given dark-gas fraction f_{DG}. Finally, in §2.4 we compare the behaviors of the profiles considered and discuss the impact of density profile on the effects of CO-dark gas.

2.1. *Power-law Profile*

Clouds are very frequently modeled as having a density profile $\rho(r)$ following a simple power-law,

$$\rho(r) = \rho_c \ x^{-k}, \tag{4}$$

where ρ_c is the central density, $x = r/R_0$, and R_0 is an arbitrary radius at which ρ is normalized. Figure 2 shows $\rho(r)$ and mass $M(r)$ as a function of r for a clump with properties [$R_0 = 0.1$ pc, $R_{H_2} = 1$ pc, $M(R_{H_2}) = 300\ M_\odot$, and $\sigma_v(R_{H_2}) = 0.6$ km s^{-1}] following $k = 1$ and $k = 2$.

In Appendix A.1, we derive the virial parameter for a clump following a power-law profile,

$$\alpha_{\rm vir}(r) = \frac{3\sigma_v^2(r)}{\pi\rho_c G R_0^2}\frac{T_1(r)}{T_2(r)}, \tag{5}$$

where $T_1(r) = \left[\frac{4x^{(3-k)}}{(3-k)}\right]$ and $T_2(r) = \left[\frac{16x^{(5-2k)}}{(5-2k)(3-k)}\right]$. Figure 3 shows the variation of $\alpha_{\rm vir}$ with r for the $k = 1$ and $k = 2$ profiles of the clump shown in Figure 2. We observe a large range of outcomes as r increases depending on the velocity and power-law indices adopted. For [$k = 1,\ \beta = 0$], the cloud has decreasing $\alpha_{\rm vir}$ value as radius increases; while for [$k = 1,\ \beta = 0.5$], $\alpha_{\rm vir}$ is constant. Similarly for [$k = 2,\ \beta = 0$], $\alpha_{\rm vir}$ is constant, and for [$k = 2,\ \beta = 0.5$], $\alpha_{\rm vir}$ increases with radius.

We then cast these equations in terms of W10's f_{DG} for more insight and to derive corrections to observed molecular cloud properties for CO-dark gas. For a cloud following a power-law profile with $k < 3$, W10 defined

$$f_{DG} = 1 - \left(\frac{R_{CO}}{R_{H_2}}\right)^{3-k}. \tag{6}$$

We derive the variation in clump properties as a function of f_{DG}. Using the definition of f_{DG} in Equation 1, the total molecular mass within R_{H_2} can be found as

$$M(R_{H_2}) = \frac{M(R_{CO})}{1 - f_{DG}}. \tag{7}$$

From Equation 6 the relationship between R_{CO} and R_{H_2} is dependent on the adopted k,

$$R_{H_2} = (1 - f_{DG})^{1/(k-3)}\ R_{CO}, \tag{8}$$

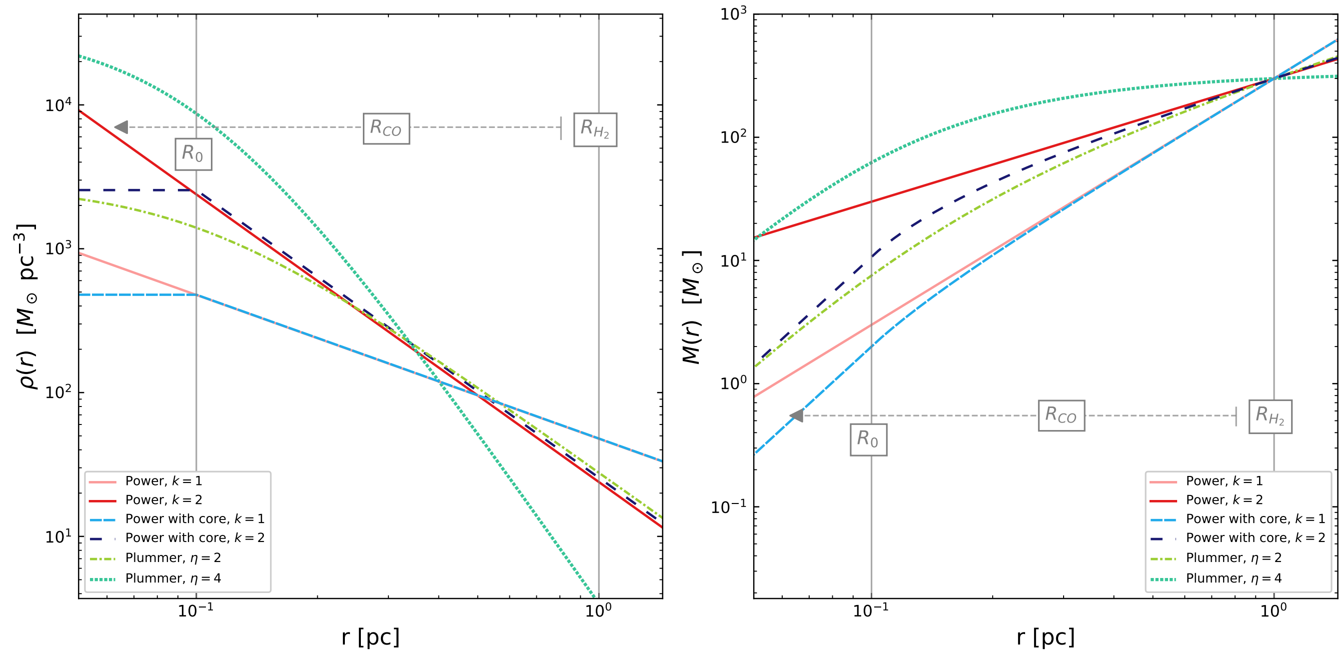

Figure 2. *Left:* Density as a function of radius for a clump with [$R_0 = 0.1$ pc, $R_{H_2} = 1$ pc, $M(R_{H_2}) = 300\ M_\odot$, $\sigma_v(R_{H_2}) = 0.6$ km s^{-1}]. R_0 and R_{H_2} are marked by the solid grey vertical lines, and the range of possible R_{CO} is shown by the dashed grey arrow. The solid pink and red curves are power-law profiles with $k = 1$ and $k = 2$, respectively (§2.1). The light blue densely-dashed and dark blue loosely-dashed curves are $k = 1$ and $k = 2$ power-laws with constant density cores, respectively (§2.2). The dash-dotted yellow-green and dotted light green curves are Plummer density profiles with $\eta = 2$ and $\eta = 4$, respectively (§2.3). *Right:* Total mass within r as a function of r for the fiducial clump, with the same line colors and styles as on the left.

and with Equation 3 evaluated at $R_0 = R_{CO}$, $\sigma_v(R_{H_2})$ can be found as

$$\sigma_v(R_{H_2}) = (1 - f_{DG})^{\beta/(k-3)}\ \sigma_v(R_{CO}). \tag{9}$$

Using Equations 6 and 7, surface density $\Sigma(r) = M(r)/\pi r^2$ at R_{H_2} becomes

$$\Sigma(R_{H_2}) = (1 - f_{DG})^{(1-k)/(k-3)}\ \Sigma(R_{CO}). \tag{10}$$

The virial mass can be expressed in terms of f_{DG} as

$$M_{\rm vir}(R_{H_2}) = (1 - f_{DG})^{(2\beta+1)/(k-3)}\ M_{\rm vir}(R_{CO}). \tag{11}$$

Finally, the CO-dark-corrected virial parameter is

$$\alpha_{\rm vir}(R_{H_2}) = \alpha_{\rm vir}(R_{CO})\ (1 - f_{DG})^{(2\beta+k-2)/(k-3)}. \tag{12}$$

In Figure 3 we show the values of R_{CO} and $\alpha_{\rm vir}(R_{CO})$ as a function of f_{DG} for the fixed $R(H_2)$=1 pc clump: for $f_{DG} = 0.3$, $R_{CO} \simeq$ [0.85 pc for $k = 1$, 0.7 pc for $k = 2$], while for f_{DG} =0.5, $R_{CO} \simeq$ [0.7 pc for $k = 1$, 0.5 pc for $k = 2$], and for f_{DG} =0.8, $R_{CO} \simeq$ [0.45 pc for $k = 1$, 0.2 pc for $k = 2$].

We consider internal pressure for the power-law profile. Under a polytropic model, turbulent pressure within a cloud is described by $P \sim \rho\sigma_v^2$. By Equations 3 and 4, the pressure gradient for this profile then follows $\mathrm{d}P/\mathrm{d}x \sim (2\beta - k)x^{2\beta-k-1}$. Thus, if $2\beta - k < 0$ an outward pressure gradient conducive to stability will be present throughout the clump. Additionally, from Equation 12 we see that while $2\beta + k < 2$, $\alpha_{\rm vir}(R_{H_2}) < \alpha_{\rm vir}(R_{CO})$, i.e., the empirical $\alpha_{\rm vir}$ from the CO-traced clump would overestimate the "true" $\alpha_{\rm vir}$ of the full cloud including CO-dark gas. In this case, relying on the CO-derived measurement alone would lead to the incorrect conclusion that the cloud is dominated by kinetic energy, and either unbound or confined by high levels of external pressure.

2.2. *Power-law Profile with a Constant Density Core*

We also examine a cloud profile that follows a power-law at large r but has a small, constant density core of radius R_0 at its center,

$$\rho(r) = \begin{cases} \rho_c \text{ for r} < \mathrm{R_0} \\ \rho_c\ x^{-k} \text{ for r} \geq \mathrm{R_0}, \end{cases} \tag{13}$$

where ρ_c is the central density and $x = r/R_0$. This has frequently been supported observationally, with $R_0 \lesssim 0.1$ pc (e.g., Girichidis et al. 2011; Juvela et al. 2018; Tang et al. 2018). We note that for this profile, unlike for the full power-law profile of §2.1, R_0 has a definite

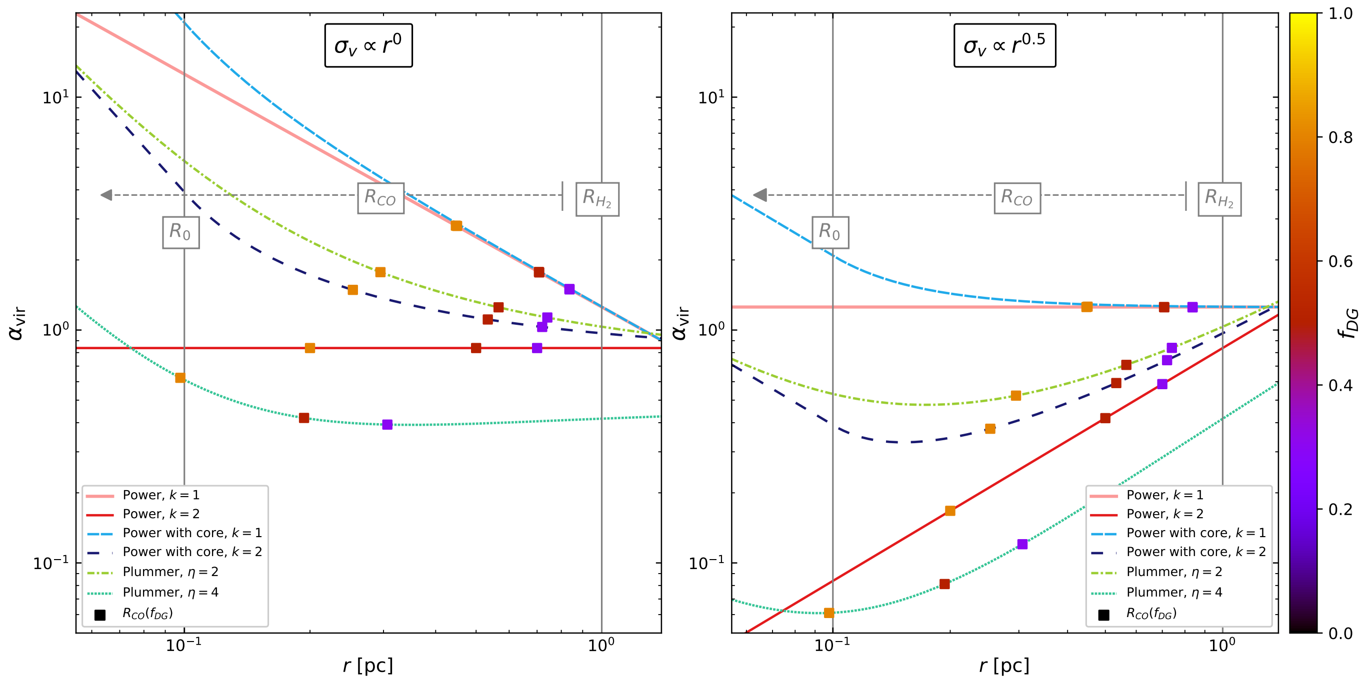

Figure 3. *Left:* The value of $\alpha_{\rm vir}$ vs r for $\sigma_v \propto r^0$ [$\beta = 0$] for the clump shown in Figure 2 with [$R_0 = 0.1$ pc, $R_{H_2} = 1$ pc, $M(R_{H_2}) = 300\ M_\odot$, $\sigma_v(R_{H_2}) = 0.6$ km s^{-1}]. The colors and styles of the profile curves, the R_0 and R_{H_2} lines, and the R_{CO} arrow are the same as in Figure 2. The square points along each curve mark the location of R_{CO} for that profile by value of f_{DG}. The purple points mark $f_{DG} = 0.3$, the red points mark $f_{DG} = 0.5$, and the orange points mark $f_{DG} = 0.8$. *Right:* Same as the left, but for $\sigma_v \propto r^{0.5}$ [$\beta = 0.5$].

physical meaning, and that R_0 is typically $\ll R_{CO}$. In Figure 2, $\rho(r)$ and $M(r)$ are shown for $k = 1$ and $k = 2$ for a clump with an identical set of properties at R_{H_2} to the clump considered in §2.1 [$R_0 = 0.1$ pc, $R_{H_2} = 1$ pc, $M(R_{H_2}) = 300\ M_\odot$, $\sigma_v(R_{H_2}) = 0.6$ km s^{-1}]. The densities of this profile and of the full power-law profile described in §2.3 are roughly in agreement at about 0.5 pc ; this is a consequence of the choice of R_0 and R_{H_2}, and changing their values changes this radius of agreement. In Appendix A.2, we follow the process outlined in §2.1 to derive the virial parameter for this profile,

$$\alpha_{\rm vir}(r) = \begin{cases} \frac{15\sigma_v^2(r)}{4\pi\rho_c G R_0^2}\frac{1}{x^2} \text{ for r} < \text{R}_0 \\ \frac{3\sigma_v^2(r)}{\pi\rho_c G R_0^2}\frac{\Pi_1(r)}{\Pi_2(r)} \text{ for r} \geq \text{R}_0. \end{cases} \tag{14}$$

where $\Pi_1(r) = \left[\frac{4}{3-k}\left(x^{3-k} - \frac{k}{3}\right)\right]$ and

$$\Pi_2(r) = \begin{cases} \frac{16}{3-k}\left(\frac{x^{5-2k}-1}{5-2k} + \frac{k(1-x^{2-k})}{6-3k} + \frac{3-k}{15}\right) \text{ for k} \neq 2 \\ \frac{16}{3-k}\left(\frac{x^{5-2k}-1}{5-2k} - \frac{k\ln(x)}{3} + \frac{3-k}{15}\right) \text{ for k} = 2. \end{cases} \tag{15}$$

The variation of $\alpha_{\rm vir}$ with r is shown in Figure 3 for $k = 1$ and $k = 2$. We observe a wide variety of behaviors as the area considered outside of the central core R_0 increases depending on the assumed density and velocity profiles. For both [k=1, β=0] and [$k = 2$, β=0], $\alpha_{\rm vir}$ decreases rapidly with increasing x. For [$k = 1$, $\beta = 0.5$], $\alpha_{\rm vir}$ plateaus marginally below the value of $\alpha_{\rm vir}$ at $x = 1$, and for [$k = 2$, $\beta = 0.5$], it increases rapidly. Any conclusions as to whether the virial parameter of the CO-traceable material accurately represents the entire cloud, including CO-dark gas, are then extremely dependent on the assumptions made.

In Appendix A.2, we also derive

$$f_{DG} = \begin{cases} 1 - \left(\frac{3-k}{3}\frac{(R_{CO}/R_0)^3}{(R_{H_2}/R_0)^{(3-k)} - \frac{k}{3}}\right) \text{ for R}_{\rm CO} < \text{R}_0 \\ 1 - \left(\frac{(R_{CO}/R_0)^{(3-k)} - \frac{k}{3}}{(R_{H_2}/R_0)^{(3-k)} - \frac{k}{3}}\right) \text{ for R}_{\rm CO} \geq \text{R}_0, \end{cases} \tag{16}$$

assuming $R_{H_2} > R_0$. We show the value of R_{CO} as a function of f_{DG} in Figure 3 for $f_{DG} = 0.3$, 0.5, and 0.8. For $f_{DG} = 0.3$, $R_{CO} \simeq$ [0.85 pc for $k = 1$, 0.7 pc for $k = 2$], while for f_{DG} =0.5, $R_{CO} \simeq$ [0.7 pc for $k = 1$, 0.5 pc for $k = 2$], and for f_{DG} =0.8, $R_{CO} \simeq$ [0.45 pc for $k = 1$, 0.25 pc for $k = 2$].

We consider the limit of a clump with a very large central core, such that R_0 approaches R_{H_2} and k is effectively zero throughout the clump. In this case, we can derive a simplified kinetic term, $\Omega_k \propto x^{2\beta+3}$ and gravitational term $\Omega_G \propto x^5$, leading to

$$\alpha_{\rm vir} \propto x^{2\beta-2}. \tag{17}$$

Therefore, while $\beta < 1$ the virial parameter will decrease as the radius r at which clump properties are evaluated increases. Most measurements of β on $\gtrsim 0.1$ pc scales range between 0.2–0.5 (e.g., Heyer & Brunt 2004; Caselli & Myers 1995; Lin et al. 2021), so this condition appears easily met. We then expect that, if this condition were met, a full clump including CO-dark gas would have a lower $\alpha_{\rm vir}$ than the result derived only from the CO-traced material, i.e., it would be more gravitationally dominated that could be inferred from CO alone.

2.3. *Plummer Profile*

The Plummer density profile (Plummer 1911) is frequently applied to molecular clouds and yields a small, flat inner core that transitions to a power-law profile at large radii. The Plummer profile follows

$$\rho(r) = \rho_c \left(\frac{1}{\sqrt{x^2+1}}\right)^{\eta}, \tag{18}$$

where ρ_c is the central density, R_0 is the radius of the central core, $x \equiv r/R_0$, and η is the index of the power-law at large radii. Pattle (2016) modeled the evolution of pressure-confined cores following Plummer-like density profiles in order to evaluate whether the cores were likely to collapse or to reach virial equilibrium as a function of radius. Here we extend this work in the context of CO-dark gas.

We derive corrections for CO-dark gas for two values of η: $\eta = 2$ as consistent with recent observational results ranging between $\eta = 1.5$–2.5 (e.g., Arzoumanian et al. 2011; Palmeirim et al. 2013; Zucker et al. 2021), and η=4 following Whitworth & Ward-Thompson (2001) and Pattle (2016). We adopt an internal cloud velocity dispersion profile following Equation 3. $\rho(r)$ and $M(r)$ are shown in Figure 2 for the fiducial clump with properties $[R_0 = 0.1$ pc, $R_{H_2} = 1$ pc, $M(R_{H_2}) = 300\ M_\odot$ $\sigma_v(R_{H_2}) = 0.6$ km s^{-1}].

In Appendix A.3, we derive the virial parameter for this profile as

$$\alpha_{\rm vir}(r) = \frac{3\sigma_v^2(r)}{\pi \rho_c G R_0^2}\frac{P_1(r)}{P_2(r)}. \tag{19}$$

where

$$P_1(r) = \begin{cases} 4(x - \arctan(x)) \text{ for } \eta = 2 \\ 2(\arctan(x) - \frac{x}{x^2+1}) \text{ for } \eta = 4, \end{cases} \tag{20}$$

and

$$P_2(r) = \begin{cases} 4P_1(r) - 16\displaystyle\int_0^x \frac{x' \arctan(x')}{x'^2+1}{\rm d}x' \text{ for } \eta = 2 \\ \arctan(x) + \frac{x - 4\arctan(x)}{x^2+1} + \frac{2x}{(x^2+1)^2} \text{ for } \eta = 4. \end{cases} \tag{21}$$

We similarly derive

$$f_{DG} = \begin{cases} 1 - \left(\dfrac{\frac{R_{CO}}{R_0} - \arctan(R_{CO}/R_0)}{\frac{R_{H_2}}{R_0} - \arctan(R_{H_2}/R_0)}\right) \text{ for } \eta = 2 \\ 1 - \left(\dfrac{\arctan(R_{CO}/R_0) - \frac{(R_{CO}/R_0)}{(R_{CO}/R_0)^2+1}}{\arctan(R_{H_2}/R_0) - \frac{(R_{H_2}/R_0)}{(R_{H_2}/R_0)^2+1}}\right) \text{ for } \eta = 4. \end{cases} \tag{22}$$

In Figure 3, we present the behavior of $\alpha_{\rm vir}$ as a function of r for this profile. We also numerically solve for and plot the expected R_{CO} using Equation 22 for f_{DG} = 0.3, 0.5, and 0.8. For f_{DG} =0.3, $R_{CO} \simeq$ [0.75 pc for $\eta = 2$, 0.3 pc for $\eta = 4$]. For f_{DG} =0.5, $R_{CO} \simeq$ [0.55 pc for $\eta = 2$, 0.2 pc for $\eta = 4$], and for f_{DG} =0.8, $R_{CO} \simeq$ [0.3 pc for $\eta = 2$, 0.1 pc for $\eta = 4$].

As in the other profiles considered, the behavior of $\alpha_{\rm vir}$ for the Plummer profile is highly variable and is dependent upon the density and velocity assumptions made. For a cloud with $\beta = 0$, $\alpha_{\rm vir}$ is roughly constant for $x > 1$ and only marginally below the value of $\alpha_{\rm vir}$ at $x = 1$. This indicates that the stability that would be inferred from just the CO-traced mass is a fairly accurate representation of the stability of the entire cloud. In contrast, for $\beta = 0.5$, $\alpha_{\rm vir}$ increases rapidly above R_0, suggesting that the CO-traced cloud would appear more gravitationally bound than the full cloud at R_{H_2}.

2.4. *Comparison of Density Profiles*

We consider the effect of the density profile on $\alpha_{\rm vir}$ and the amount by which CO-dark gas changes observed clump properties. The impact of k/η on the overall value of $\alpha_{\rm vir}$ is similar between all profiles considered, with smaller k leading to higher $\alpha_{\rm vir}$ below R_{H_2}. In particular, $\alpha_{\rm vir}$ is typically ~2x larger for the power-law profiles of §2.1 and §2.2 than in the $\alpha_{\rm vir}$ of the $\eta = 4$ Plummer profile. This is the result of the $\eta = 4$ Plummer profile having a much higher proportion of mass centrally concentrated at small r than the $k = 1$ and $k = 2$ power-law based profiles and $\eta = 2$ Plummer profile (see Figure 2).

Since $\Omega_G \sim GM^2/r$, concentrating a fixed amount of mass within a smaller area increases the object's gravitational potential and decreases the virial mass. In contrast, $\Omega_K \sim \sigma_v^2 M$ is not as dependent on the volume in which M is contained so it is unsurprising that $\alpha_{\rm vir}$ is significantly reduced for the steeper profiles. Very sub-virial clumps are expected to rapidly collapse, and so to offset this effect and move closer to stability at R_{H_2}, the Plummer profile would need to have a much higher $\sigma_v(R_{H_2})$. The assumed radial velocity dispersion index β also has a large impact on clump dynamical state and the $\alpha_{\rm vir}$ that would be inferred after correcting for CO-dark gas. For all profiles considered in this section, the choice of β generally corresponds to the "direction" of the behavior of $\alpha_{\rm vir}$ with r, whether increasing, decreasing, or constant.

Throughout this work, we use f_{DG} for a given clump as a set parameter, without attempting to tie its specific value to the underlying physics that determine the value of f_{DG}. In reality, f_{DG} is a function of the properties of the clump and the environment in which it is immersed. Since we aim to derive corrections that may be applied by observers using a specific assumed or measured value of f_{DG} to estimate clump properties, accounting for the nuances of the physical drivers of f_{DG} is beyond the scope of this work. However, we do expect that clumps with steeper density profiles will have lower f_{DG} than clumps with shallower profiles occupying the same environment under identical conditions.

We can intuitively consider that, in a given environment, a specific A_V/density threshold must be reached for CO to effectively self-shield (determined by radiation field strength, dust-to-gas ratio, etc). Since steeper profiles are more centrally concentrated in mass, they would contain a larger fraction of the total clump mass at this density floor where CO begins to be destroyed (R_{CO}); this would decrease f_{DG} despite being in an identical environment to a shallower clump.

Finally, we evaluate the overall effect of the internal density profile assumed on inferred CO-dark-corrected clump properties. For the fixed R_{H_2} we consider, the derived R_{CO} for a given f_{DG} decreases with increasing k (or η) because of our assumption that $R_0 \ll R_{CO}$. The values of R_{CO} and $\alpha_{\rm vir}(R_{CO})$ for a power-law with core profile with $k = 1$ are functionally identical to those for a full power-law, and for a $k = 2$ power-law with core profile depart only slightly from the values derived from a full $k = 2$ power-law. Under most scenarios where $R_0 \ll R_{CO}$, the corrected properties (R_{H_2}, $\alpha_{\rm vir}(R_{H_2})$, etc) derived for the power-law with core profile vary by a small amount (generally $\lesssim$10% difference) from the corrections for a full power-law. The difference between corrected properties between profiles increases with increasing f_{DG}. The difference in corrected values between the steep $\eta = 4$ Plummer and power-law profiles is larger, but this is likely more of an effect of the variation in assumed η vs. k than of profile itself; the $\eta = 2$ profile also typically differs by $\lesssim$10% from the power-law based profiles.

Therefore, we conclude that the relative steepness by which density decreases with r has a larger impact on the effects of CO-dark gas on observed clump properties than the exact form of the radial density profiles we consider. For the remainder of this work, we focus our analysis on the behavior of clumps following single power-law profiles for simplicity.

3. CO-DARK GAS AND SIZE–LINEWIDTH–SURFACE DENSITY RELATIONSHIPS

3.1. *Larson's Scaling Relationships*

Larson (1981) observed correlations between the size R, velocity dispersion σ_v, and mass surface density Σ of Galactic molecular clouds that have been confirmed and refined by later studies. The first of these relationships is a power-law relationship between the size of a cloud R and σ_v, where

$$\sigma_v \simeq C \left(\frac{R}{1\ \rm pc}\right)^{\Gamma} \rm km\ s^{-1}. \tag{23}$$

Larson (1981) originally derived $\Gamma = 0.38$ and $C = 1.1$ km s^{-1}, an estimate that Solomon et al. (1987) and Heyer et al. (2009) (hereafter SRBY and H09, respectively) later refined to

$$\sigma_v \simeq 0.72 \left(\frac{R}{1\ \rm pc}\right)^{0.5} \rm km\ s^{-1}. \tag{24}$$

Larson's second relationship is derived from observed correlations between σ_v and cloud mass M,

$$\frac{2\sigma_v^2 R}{GM} \simeq 1, \tag{25}$$

which is usually interpreted as meaning that most clouds are roughly in virial equilibrium. Alternatively, it has been suggested that this is a signature of global hierarchical collapse at all scales within clouds (Ballesteros-Paredes et al. 2011; Vázquez-Semadeni et al. 2019). Finally, cloud density and size are observed to be inversely related, $n \propto R^{-1.1}$, suggesting that surface density is independent of size and should be roughly constant for clouds under conditions similar to the Milky Way, although observations have suggested that Σ does vary over several orders of magnitude with environment (e.g., H09; Sun et al. 2018; Traficante et al. 2018; Dessauges-Zavadsky et al. 2019; Chevance et al. 2020).

As noted by H09, a natural extension of these relationships is an association between surface density Σ and the size-linewidth parameter σ_v^2/R. A virialized sphere following a power-law density distribution should follow

$$\frac{\sigma_v^2}{R} = \frac{(3-k)}{3(5-2k)}\pi G\Sigma. \tag{26}$$

In Figures 4 and 5, we compare the relationships between R, σ_v, and Σ for structures observed using CO as a tracer across a variety of environments:

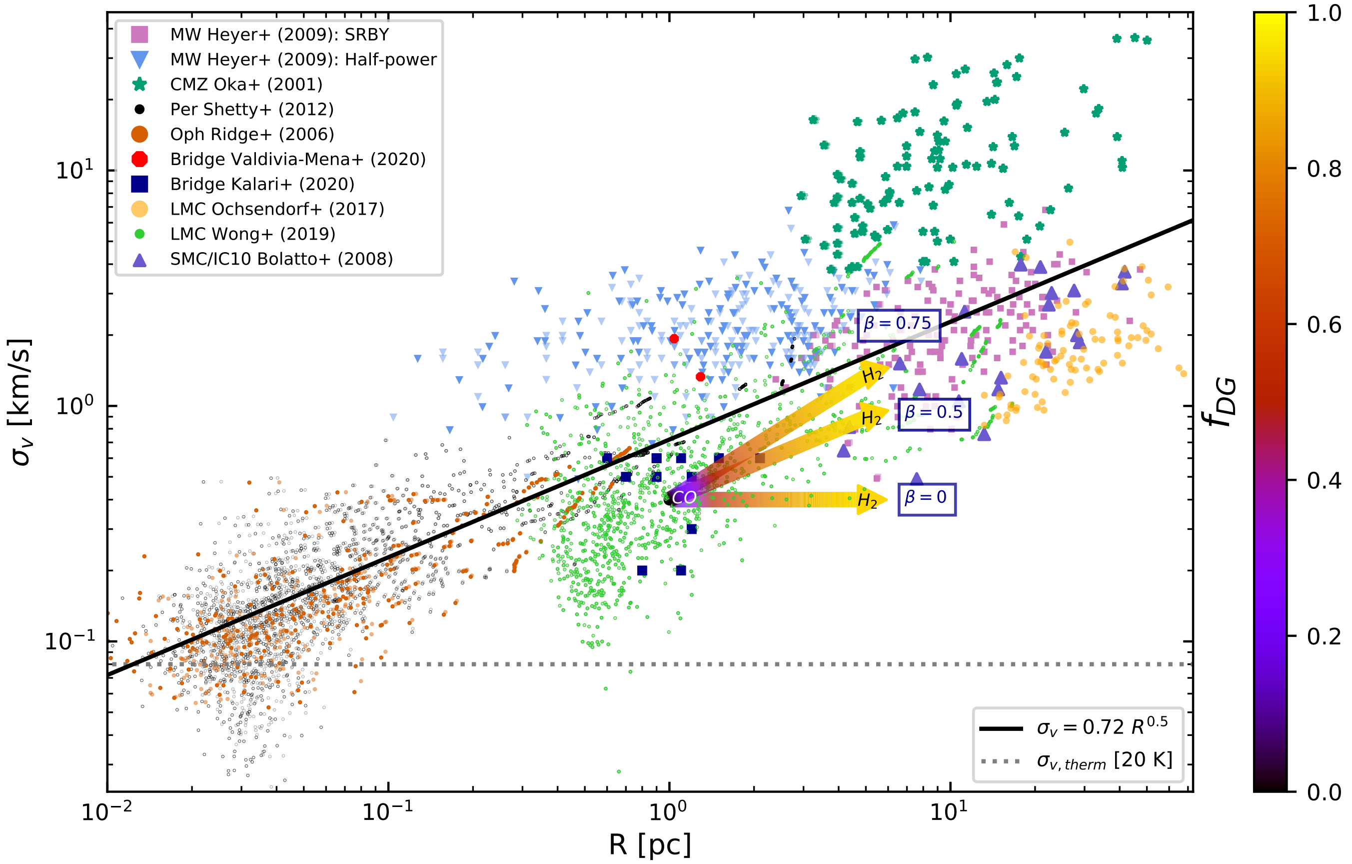

Figure 4. Velocity dispersion σ_v compared to radius R of CO-traced structures described in §3. The black line follows the relationship $\sigma_v = 0.72R^{0.5}$ (Equation 24) and the grey dotted line shows the expected contribution from thermal motion to σ_v at T=20 K. The arrows show the direction in which one would correct observed CO-traced clump properties for CO-dark gas to recover the properties of the full H_2 clump. The arrows start at physical properties typical of parsec-scale CO-traced clumps, and move towards the inferred properties of the H_2 clump. Each arrow is labeled with the power-law index k and velocity dispersion index β assumed to generate its path, and the color gradient along the arrows shows the corrected H_2 properties as a function of f_{DG}.

1. Galactic giant molecular clouds (GMCs), with sub-samples with areas defined from ^{12}CO by SRBY and from the ^{13}CO half-power contours of their central cores (H09);
2. clouds in the Galactic central molecular zone (Oka et al. 2001);
3. cores observed in the Ophiuchus molecular cloud (Ridge et al. 2006);
4. cores in the Perseus molecular cloud (Shetty et al. 2012);
5. Clumps in the Magellanic Bridge (Kalari et al. 2020; Valdivia-Mena et al. 2020);
6. Clumps in the LMC regions 30 Doradus, A439, GMC 104, GMC 1, PCC, and N59C (Wong et al. 2019);
7. GMCs in ∼150 star-forming regions throughout the LMC (Ochsendorf et al. 2017);
8. Clouds in the SMC and dwarf galaxy IC 10 (Bolatto et al. 2008),

(where the choice of terminology core/clump/cloud corresponds to commonly used size scales of ∼0.1/1/10 pc, without any differences in relevant physics implied).

In Figure 4 where R and σ_v are compared, the usual size-linewidth relationship of Equation 24 is displayed. Σ and σ_v^2/R are compared in Figure 5 for a subset of the sources listed above that have cloud mass estimates derived without assuming virial equilibrium (1, 6, 7, 8 in the list above). Equation 26 is shown as the straight black line for $k = 0$. Additionally, in pressure-bounded virial equilibrium, Σ and σ_v^2/R are related as (Field et al. 2011)

$$\frac{\sigma_v^2}{R} = \frac{1}{3}\left(\pi\gamma G\Sigma + \frac{4P_e}{\Sigma}\right), \tag{27}$$

which is shown in Figure 5 by the V-shaped curves, with $\gamma = 0.6$ for a cloud with $k = 0$.

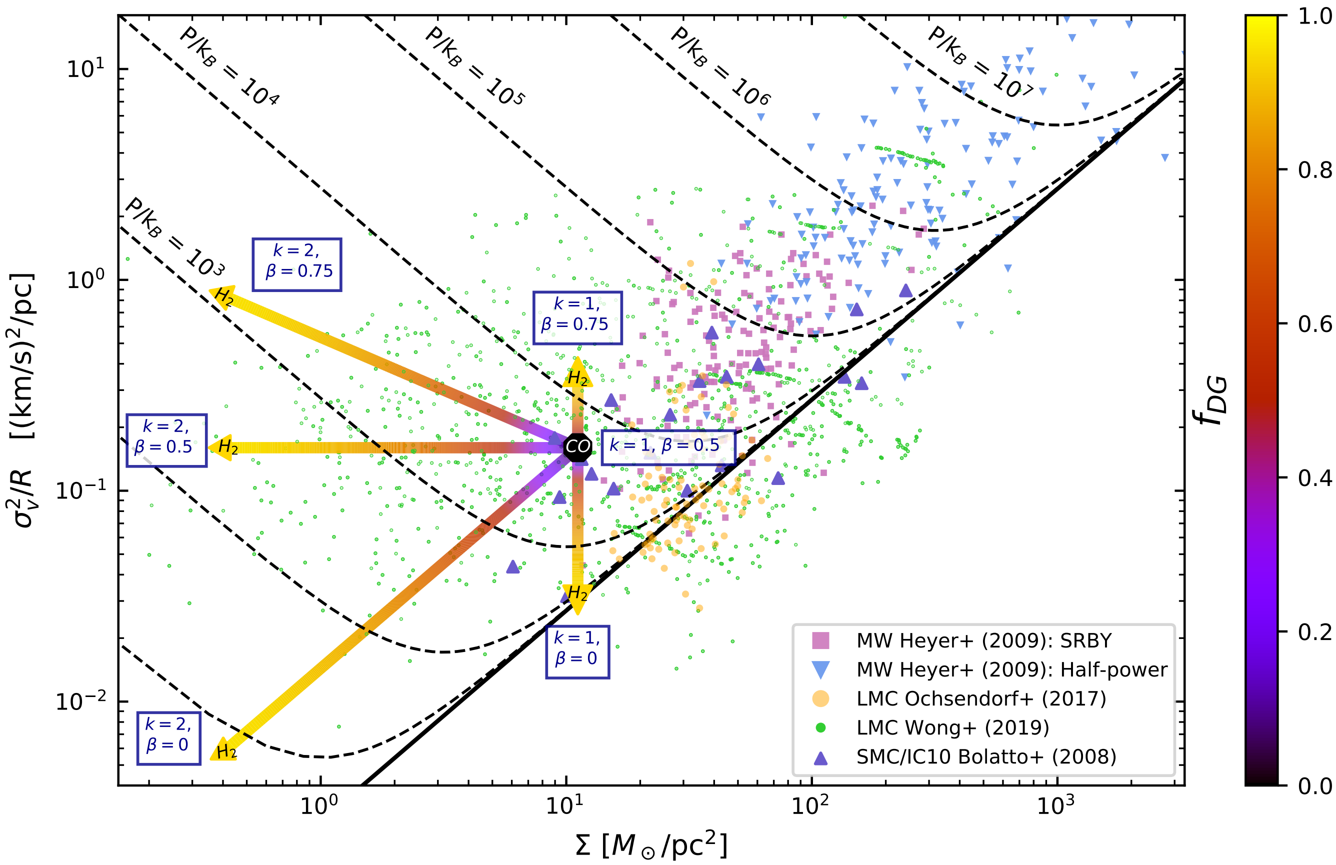

Figure 5. Size-linewidth parameter σ_v^2/R compared to surface density Σ for CO-traced structures described in §3. The black line corresponds to virial equilibrium without external pressure (Equation 26) and the dashed black curves correspond to virial equilibrium under external pressure with units for P/k_B labels in K cm^{-3} (Equation 27). The arrows are as in Figure 4 and show the direction in which one would correct observed CO-traced clump properties to recover the properties of the full clump including CO-dark H_2.

The majority of the Wong et al. (2019) and Ochsendorf et al. (2017) LMC GMCs, Kalari et al. (2020) Bridge clumps, and Bolatto et al. (2008) SMC and IC 10 clouds have **smaller** σ_v **for a given** R **than expected from Galactic clouds**, falling well under the relationship described in Equation 24. This has been observed in a variety of other low-Z environments as well, e.g. by Rubio et al. (2015) in the $Z \simeq 0.13\ Z_\odot$ dwarf galaxy Wolf–Lundmark–Melotte, and Hughes et al. (2013) in the LMC. Many of these samples also have lower Σ for a given σ_v^2/R than expected based on Equation 26, suggesting that the structures are either unbound and transient or must be confined by external pressure to remain stable, as position in this space is directly related to $\alpha_{\rm vir}$.

As part of the PHANGS-ALMA collaboration ("Physics at High Angular-resolution in Nearby GalaxieS with ALMA," Leroy et al. 2021), Sun et al. (2020) analyzed the dynamical states of molecular gas in 28 nearby disk galaxies. They derived typical midplane pressures over 1 kpc scales ranging from $P/k_B = 10^3$–10^6 K cm^{-3}, and found that the average internal turbulent pressure of clouds was typically very similar to the required cloud-scale equilibrium pressure, which they concluded indicated that most gas was in dynamical equilibrium. Wong et al. (2009) derived an average midplane hydrostatic pressure in the central regions of the LMC of $P/k_B \sim 10^4$ K cm^{-3} using HI and CO(1–0) observations, which could be sufficient to confine a large fraction of the Wong et al. (2019) LMC clumps observed to have high σ_v^2/R, as well as the majority of the Ochsendorf et al. (2017) LMC GMCs.

3.2. *Effects of CO-dark Gas on Observed Relationships*

The effects of our derived corrections for CO-dark gas in a power-law density profile clump (§2.1: Equations 8, 9, and 10) are shown by the arrows in Figures 4 and 5. The arrows start at the properties of a clump as observed solely in CO and move towards the "true" characteristics of the full clump including CO-dark gas, with color gradients along the arrow corresponding to f_{DG}. The initial conditions for the corrections displayed are [$R_{CO} = 1$ pc, $\sigma_v(R_{CO}) = 0.4$ km s^{-1}, $M(R_{CO}) = 35\ M_\odot$]; these values correspond to the medians of these quantities for roughly pc-scale CO clumps in the Wong et al. (2019) sample. Changing the arrow's origin does not impact the direction of the arrow. In Figure 4a, we only show corrections for $k = 1$, with arrows for $\beta = 0$, $\beta = 0.5$, and $\beta = 0.75$; this is because the corrections for $k = 1$ vs $k = 2$ overlap in this space and differ only in the extent to which their arrows extend. In Figure 5 we also show corrections for $k = 2$.

For $\beta = 0$, correcting for CO-dark gas causes clumps to have even lower σ_v relative to the increased R and thus drives the clumps further from following Equation 24. We note that a velocity profile this "flat" is unlikely, as turbulence within the ISM is mainly driven at large scales (Brunt et al. 2009), but we display it to demonstrate the limits of this effect. For $\beta = 0.5$, the corrections have no effect on the position of the clump in size-linewidth space relative to the expected Equation 24 relationship; this is because the standard inter-clump relationship Equation 24 and the displayed intra-clump profile share the same $\beta = \Gamma = 0.5$. By the same logic, $\beta = 0.75$ unsurprisingly brings clumps closer to agreement with Equation 24 because $\beta > \Gamma$. Corrections in Σ vs. σ_v^2/R space have a similarly variable effect. The distance of any given clump from the virial line in Figure 5 is directly proportional to the stability of the clump as measured by $\alpha_{\rm vir}$ and we interpret the corrections for a power-law profile in this context as follows.

- We again see that the assumed k and β have a large impact on the inferred corrected state: for $k = 1$, Σ is constant; while for $k > 2$, the corrected Σ is significantly reduced.
- We observe that clumps decrease in $\alpha_{\rm vir}$ and move towards $\alpha_{\rm vir} \sim 1$ in all cases where $k + 2\beta < 2$. This suggests that, if these profile conditions are met, the apparently high $\alpha_{\rm vir}$ structures traced by CO in low-Z, high f_{DG} environments may be closer to stability than expected.
- In most cases, the updated clump positions suggest that a lower level of external pressure would be required to maintain stability than would be inferred from CO-traced material alone.

The assumed density and velocity profiles then almost entirely determine the "direction" of these biases. This highlights the importance of studies of the spatial dependence of density and linewidth on the scale of individual clouds in addition to SRBY/H09-type studies comparing these quantities between cloud populations.

4. DISCUSSION

4.1. *Can CO-dark Gas Explain Departures from Larson's Relationships?*

From the clump property corrections derived in §2 and described in §3, it is clear that neglecting CO-dark gas could significantly bias the assessment of cloud placement in Larson's relationships and gravitational stability. We now examine if this effect is sufficient to explain the observed high $\alpha_{\rm vir}$ and departures from Larson's relationships in low-Z environments.

Under the corrections for a power-law profile that we have derived, low-σ_v clumps must follow an internal velocity profile with $\beta > 0.5$ (i.e., have large motions at large scales) to reconcile with the typical size-linewidth relationship described by Equation 24; however, large βs also yield increased $\alpha_{\rm vir}$ that imply the full structure is gravitationally unbound. If instead one assumes that clouds are close to virialized without external pressure, then the dark gas correction required to move observed points closer to virialization (i.e., to decrease $\alpha_{\rm vir}$) requires that clumps follow a shallow density profile and have $\beta < 0.5$ — but, shallow βs increase the amount by which these clouds fall "under" the R–σ_v relationship of Equation 24.

This contradiction is most problematic in structures with high f_{DG} as expected in low-Z or high-radiation environments, and can be resolved if clouds in these areas are: (1) overwhelmingly gravitationally unstable and dispersing rapidly as a result; or (2) require much higher levels of external pressure to remain stable than clouds in more typical environments; or (3) possess a global σ_v/R trend shifted to lower values of σ_v than the classical Equation 24 relationship (i.e., a smaller scaling coefficient C in Equation. 23) and have shallow internal density and σ_v profiles (0$\leq \beta <$0.5).

(1) is unlikely statistically simply because of the number of clouds that are observed, and a physical cause for (2) is hard to imagine since the typical ISM pressure in low-Z galaxies is ∼1–2 orders of magnitude smaller than in typical large spiral galaxies (de los Reyes & Kennicutt 2019). There are also nontrivial direct relationships between metallicity and ISM pressure in these areas because of reduced cooling and thermal balance, but predictions as a function of metallicity are generally only possible in the context of a self-regulated star formation model and thus the specifics depend on the details of that model. Additionally, the direct effects of metallicity via the cooling rate on pressure are less important than the galaxy type to the properties of molecular clouds

(3) is then the most compelling, and would be the simplest way to account for observed low-σ_v and high

α_{vir} structures in low-Z areas. Shallow density profiles of $1.5 < k < 2$ are typical on the pc-scales where the simplified isolated spherical PDR model that we consider here holds (Caselli et al. 2002; Pirogov 2009; Arzoumanian et al. 2011; Schneider et al. 2013), and even shallower profiles ($k \sim 1$) have been found in young, low-density cores and clumps (Chen et al. 2019, 2020; Lin et al. 2021). Small values of C and steep Γ relative to SRBY's $C = 0.72$ and $\Gamma = 0.5$ have been derived from CO observations for structures in the SMC, LMC, and other local dwarf galaxies where low-σ_v/ high α_{vir} structures are found (with C $\sim 0.2 - 0.6$ and $\Gamma \sim 0.55$–0.85) (Bolatto et al. 2008; Hughes et al. 2010, 2013; Wong et al. 2019). In CO-dark regions, HI can also be used as a probe of turbulence. For a sample of HI clouds in the LMC, Kim et al. (2007) derived a mean $\Gamma \simeq 0.5$.

In the pioneering Larson (1981) study, a shallow $\Gamma = 0.38$ was derived, which is similar to the Kolmogorov index for turbulent cascade in an incompressible medium $\beta \sim 1/3$. More recently, $\beta \simeq 1/2$ has frequently been found for GMCs both observationally and through simulations (e.g., Heyer & Brunt 2004; Dobbs 2015); this aligns the expectation for Burgers turbulence (Passot et al. 1988) i.e., in an isotropic system dominated by shocks, and is in accordance with SRBY's $\Gamma = 0.5$. On very small scales ($\lesssim 0.05$ pc) a break in the internal size-linewidth relationship has been observed with β approaching zero (Goodman et al. 1998; Caselli et al. 2002; Volgenau et al. 2006; Pineda et al. 2010); however, it seems unlikely that the W10 scenario of PDRs of isolated individual spherical clouds surrounded by envelopes of dark gas would hold on these sizes because cores are typically embedded within larger structures.

Shallow values of β ($\beta \sim 0.2$–0.3) have also been derived in high mass star-forming regions (Caselli & Myers 1995) and in prestellar cores and young clumps (Tatematsu et al. 2004; Lee et al. 2015; Lin et al. 2021). Bertram et al. (2015) analyzed turbulence within simulated molecular clouds using the Δ-variance method, from which they compared the values of β within the full cloud, within H_2 gas, and within CO-traced material. For initial densities ranging between $30 - 100$ cm^{-3}, the derived β ranged between $\sim$0.3–0.6 as derived from the resulting H_2 density maps, and $\sim$0.15–0.4 as traced by CO density, a difference which they attributed to the compact nature of the CO structures as compared to the more extended H_2.

We emphasize that the inter-clump size-linewidth relationship with exponent Γ is obtained by comparing populations of clumps, while the intra-clump size-linewidth relationship with exponent β is obtained by studying individual structures. The latter relationship is much more challenging to measure in the typically distant low-Z environments due to the required high angular resolutions and has only recently become possible, but is key for assessing if the implied shallow β is realistic. Overall, the measurements of β that have been obtained locally generally resemble observed values of Γ.

This observed correspondence of $\beta \sim \Gamma \simeq 0.5$ has been interpreted as reflecting the uniformity of velocity structure functions between individual clouds, so that Γ is largely set by β (Heyer & Brunt 2004). The implication from (3) that β is shallower than the observed Γ in low-Z environments creates some tension with this conclusion. One explanation for this difference could be a correlation between f_{DG} and cloud size. In their sample of LMC GMCs, Ochsendorf et al. (2017) observed a decrease in the ratio of CO-traced mass to dust-traced mass as dust-traced mass increased. Since the dust-traced mass likely includes the diffuse CO-dark gas, this suggests that a correlation between f_{DG} and cloud size exists with larger clouds having higher f_{DG}. Larger clouds would then systematically have larger relative changes between their true properties including CO-dark gas to their observed properties than smaller clouds do.

The "true" Γ relating the full clouds including CO-dark gas could then be shallower than the observed, CO-derived Γ, and instead approach (and possibly be determined by) the expected shallow β. This would explain the general steepness of CO-traced Γ in low-Z environments, as well as resolve the implied difference between β and Γ in low-Z environments. It is of course also possible that clumps in these low-Z environments do truly have different physical properties and scaling relationships than clumps under Galactic conditions.

4.2. *CO-dark Gas and Star Formation*

4.2.1. *Star Formation Efficiency Considering CO-dark Gas*

On kpc-scales, low-Z galaxies have been found to depart from the Kennicutt-Schmidt relationship, possessing higher star formation rate densities at a given molecular gas surface density as assessed by CO than found in more typical environments (e.g. Galametz et al. 2009; Schruba et al. 2012). Star formation efficiency (SFE) is frequently assessed by comparing the star formation rate (SFR) to gas mass ($\epsilon' = \mathrm{SFR} \,/\, M_{\mathrm{cloud}}$), so this departure suggests that the SFE is also much higher than under Galactic conditions.

Madden et al. (2020) showed that CO-dark gas is sufficient to cause the apparent variation from the Kennicutt-Schmidt relationship on galactic scales, and that when corrected for the missing mass star formation in these environments is not significantly more efficient. It has also been suggested that H_2 gas is not a requirement for star formation but is usually present as a consequence of the necessary shielding for stars to form (Glover & Clark 2012; Krumholz 2012). Star formation could then in principle proceed in atomic gas without the presence of molecular gas (although this would be rare), and may explain the lack of CO detections and corresponding high implied SFEs in some low-Z star-forming galaxies.

While CO-dark gas appears to be responsible for increased SFEs on large scales because surface densities averaged over large scales are increased by the addition of CO-dark gas mass, it is unclear how it impacts star formation in individual clumps. SFE is also frequently evaluated by simply comparing the total stellar mass to total molecular mass ($\epsilon = M_*/M_{\rm cloud}$), or as a function of free-fall time (ϵ_{ff}=$\tau_{ff} \times \epsilon'$). A simple but perhaps naive correction to the SFE of an individual CO-traced clump for missing H_2 mass would be $\epsilon_{H_2} = (1-f_{DG})\epsilon_{CO}$ by Equation 7, with a similar correction for ϵ'. We have shown that CO-based observations are likely to overestimate mean clump density. This would lead to underestimates of free-fall time τ_{ff} and also, depending on density profile, potentially to underestimates of ϵ_{ff} as well.

However, the ϵ-based metrics are generally derived over larger scales, which helps offset the unknown variation from the original total gas mass for any given star-forming clump to its present day mass by averaging over clumps and cores at a variety of stages in the star formation process. Using the present-day gas mass of a single clump to try to derive a by-clump efficiency loses this advantage, and so we only suggest the use of the proposed corrected ϵ_{H_2} and related quantities on larger scales and even then with caution.

It is still not well understood if the actual way and timescale over which clouds collapse in low-Z environments is different than under conditions similar to the solar neighborhood, and, if so, how this departure influences the SFR/SFE. Parmentier (2020) and Parmentier & Pasquali (2020) derived a relationship between clump radial density profile and SFR and found that steeper profiles correspond to higher initial SFRs: star formation proceeds most rapidly in the densest areas of clumps, and the centers of clumps with very steep density profiles are denser than shallower clumps of the same mass.

Since radiation fields are known to be enhanced in the interclump medium due to the decreased dust-to-gas ratios, it is plausible that the typical radial profiles of clumps could be different than in higher-Z environments. We have shown that shallow density profiles are required for the properties of low-Z clumps to approximate those of Galactic clumps, so it follows that in this scenario the low-Z clump-scale SFR could be slower than in higher-Z environments. To reconcile with the observed high SFR averaged over kpc-scales, relatively more clumps would need to exist to achieve these values.

Measurements of the total gas mass and SFR over large scales is clearly critical for these observations of SFE, and it is well understood that underestimating total mass can skew SFE estimates. An additional factor is the mechanics of how these clouds collapse to form stars at clump and core scales and the fraction of gas at these scales that actively contributes to star formation. The extent to which the diffuse CO-dark envelopes participate in star formation is unclear and is one of several contributing factors that sets SFE.

The scaling relationships between molecular gas and SFR observed over large scales can be validated by understanding the fraction of gas at clump-scales involved in star formation and the factors that affect the stability of individual cores and clumps. Detailed studies of the distribution and state of clumpy molecular gas is key in fully explaining the SFR/SFE and origin of scaling relationships at kpc-scales. Our models show the importance of CO-dark gas fraction and density and velocity dispersion profiles in influencing these properties. This work is then relevant to large scale measurements of SFEs in contextualizing interpretations of these measurements.

4.2.2. *Relationship between f_{DG} and the CO-to-H_2 Conversion Factor*

The corrections for CO-dark gas we have derived are dependent on having an estimate of the total molecular gas within R_{CO}, which could be derived through e.g., assuming local thermal equilibrium (LTE) with the use of multiple CO transitions, applying the non-LTE RADEX modeling, or similar methods. The widely-used CO-to-H_2 conversion factor

$$\begin{aligned} X_{CO} &= \frac{\mathrm{N(H_2)}}{I_{CO}}\ [\mathrm{cm^{-2}\ (K\ km\ s^{-1})^{-1}}] \\ \alpha_{CO} &= \frac{M_{\rm vir}(R_{CO})}{L_{CO}}\ [M_\odot\ (\mathrm{K\ km\ s^{-1}\ pc^2})^{-1}], \end{aligned} \tag{28}$$

is also designed to account for the untraced H_2 gas that we correct for in this work using f_{DG}. Some degree of correspondence between the two is then expected, as shown in previous works simulating the relationship between X_{CO} and environmental conditions (Shetty et al. 2011a,b; Clark & Glover 2015; Szűcs et al. 2016; Gong et al. 2018, 2020).

To demonstrate this expected relationship in the context of this work, we define a crude mass ratio factor

Y_{DG} where

$$Y_{DG} = \frac{M(R_{H_2})}{M(R_{CO})} = \frac{1}{1 - f_{DG}}, \quad (29)$$

by Equation 1. To compare the value of Y_{DG} across different environments, we define $Y_{DG,MW}$ as the typical value of Y_{DG} in the $f_{DG,MW}$ Milky Way. The expected Y_{DG} in a given environment can then be compared to $Y_{DG,MW}$ through the ratio of their respective f_{DG},

$$Y_{DG} = \left(\frac{1 - f_{DG,MW}}{1 - f_{DG}}\right) Y_{DG,MW}. \quad (30)$$

For a typical $Z \sim 1\ Z_\odot$ Galactic environment with $f_{DG,MW} = 0.3$ (W10), $Y_{DG,MW} \simeq 1.4$. From this value, $Y_{DG} \simeq 3.5\ Y_{DG,MW}$ would be expected in an environment like the $Z = 0.2\ Z_\odot$ SMC with $f_{DG} \sim 0.8$ (Jameson et al. 2018). This corresponds very well to the observed ratio between the usual Galactic $X_{CO,MW}$ and X_{CO} derived in SMC clumps: for sub-pc clumps in the SMC Wing Muraoka et al. (2017) derived $X_{\rm CO} \sim 4\ X_{\rm CO,MW}$, and for pc-scale clumps in the Magellanic Bridge Kalari et al. (2020) derived $X_{\rm CO}$ ∼3 $X_{\rm CO,MW}$ and Valdivia-Mena et al. (2020) found $X_{\rm CO}$ ∼1.5–3.5 $X_{\rm CO,MW}$. This suggests that Y_{DG} (and f_{DG}) could be used as a check of measured X_{CO} in clumps, or vice versa. In contrast, measurements of X_{CO} over cloud scales and larger ($\gtrsim$10 pc) in the SMC have ranged between 20–50 $X_{\rm CO,MW}$ (Leroy et al. 2009; Bolatto et al. 2011; Jameson et al. 2016), significantly exceeding the ratio between Y_{DG} and $Y_{DG,MW}$ that we have derived here.

This variation between X_{CO} and Y_{DG} estimates is likely caused by the well-known limits of X_{CO} at small scales (see Bolatto et al. 2013, for a review) and the limits of f_{DG} as formulated for isolated spherical PDRs by W10/in this work at large scales. When at low metallicities, X_{CO} is expected and observed to increase rapidly. On cloud and global scales, X_{CO} measurements are averaged over many clouds and so include both diffuse and dense molecular gas. For individual low-Z/high-f_{DG} clumps, though, only dense gas is reflected.

The scale (and resolution) at which clumps are measured is negatively associated with derived X_{CO}: in the LMC, for example, Fukui et al. (2008) derived $X_{CO} \sim 4\ X_{CO,MW}$ from clouds observed at ∼40 pc resolution by NANTEN, while Hughes et al. (2010) derived $X_{CO} \sim\ 2\ X_{CO,MW}$ from structures observed at ∼10 pc resolution by the Magellanic Mopra Assessment (MAGMA) survey. Lower-resolution observations run the risk of small clumps being diluted by large beam sizes, artificially inflating $\alpha_{\rm vir}$ and X_{CO} and also increasing the likelihood of such clumps not being identified at all. Resolved observations of individual pc-scale clumps in distant low-Z environments have only recently become possible and typically yield smaller conversion factors (Muraoka et al. 2017; Schruba et al. 2017; Saldaño et al. 2018; Kalari et al. 2020; Valdivia-Mena et al. 2020), approaching $X_{CO,MW}$ and in alignment with our expectations for clump f_{DG}.

4.3. *Guidance for Interpreting Observations*

We present the case of a "typical" observed CO clump with high $\alpha_{\rm vir}$, and discuss the properties that would be inferred by an observer using our derived corrections for CO-dark gas. For a clump following a typical $k = 1.5$ and $\beta = 0.5$ with [$R_{CO} = 1$ pc, $\sigma_v(R_{CO}) = 0.4$ km s^{-1}, $M(R_{CO}) = 35\ M_\odot$, $\alpha_{\rm vir}(R_{CO}) = 4.3$], a moderate Galactic $f_{DG} \sim 0.3$ (W10) yields a relatively small difference in clump properties: [$R_{H_2} = 1.3$ pc, $\sigma_v(R_{H_2}) = 0.45$ km s^{-1}, $M(R_{H_2}) = 50\ M_\odot$, $\alpha_{\rm vir}(R_{H_2}) = 4.8$]. In contrast, an extreme $f_{DG} \sim 0.9$ as occasionally derived in low-Z environments (Jameson et al. 2018) would lead to a significantly different set of inferred properties: [$R_{H_2} = 4.6$ pc, $\sigma_v(R_{H_2}) = 0.86$ km s^{-1}, $M(R_{H_2}) = 350\ M_\odot$, $\alpha_{\rm vir}(R_{H_2}) = 9.2$].

We then see that under typical assumed clump density and velocity profiles, correcting for CO-dark gas does *not* resolve the apparent instability of the structure — it actually exacerbates the issue. We emphasize again that the changes in clump properties post-correction are highly dependent upon the choice of k and β; if the same clump followed shallower profiles, a reduction in $\alpha_{\rm vir}$ could just as easily be indicated. At the same time, the magnitude of this shift makes clear that correcting for CO-dark gas is essential for an accurate assessment of clump properties in high f_{DG} environments.

5. CONCLUSIONS

We have derived easily-applied corrections to CO-derived clump properties to account for the effects of CO-dark gas. Our main conclusions are as follows:

1. For molecular clouds following power-law or Plummer density profiles, CO-derived measurements will systematically underestimate cloud mass and size. If clumps have shallow mass density and radial velocity dispersion profiles, the virial parameter $\alpha_{\rm vir}$ will be overestimated (§2).

2. In order to interpret CO observations as accurately as possible, cloud properties (e.g., size, mass, surface density, velocity dispersion, virial parameter) should be corrected using the prescriptions outlined in §2 as demonstrated in §4.3.

3. CO-derived measurements are most suspect in low-Z, high f_{DG} regions; however, CO-dark gas

is unlikely to simultaneously be the cause of observed clumps with high $\alpha_{\rm vir}$ and low σ_v relative to Larson's relationships in low-Z environments (§3, §4.1).

Understanding what other processes might drive departures from Larson's relationships and from inferred virial equilibrium should be of high priority. Attempts to correct for all of the above effects are reliant on accurate assessment of intra-cloud density and velocity profiles, and so this too should continue to be prioritized, especially on clump scales.

It is clear that assessing how star formation proceeds within clumps in low-Z regions is dependent on understanding the impact of CO-dark gas. Accounting for CO-dark gas both on local and global scales is then key in evaluating the evolutionary history and likely future of specific regions and in placing star formation in low-Z environments into its correct context. The corrections we have presented here are one tool to better leverage CO observations to estimate clump behavior after accounting for the effects of environment.

We thank Kelsey Johnson, Allison Costa, and Molly Finn for useful comments and discussion. T.J.O. and R.I. were supported during this work by NSF award 2009624. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00683.S, ADS/JAO.ALMA#2015.1.01013.S, ADS/JAO.ALMA #2016.1.00193.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

Software: Astropy (Astropy Collaboration et al. 2013, 2018); BioRender (https://biorender.com); Matplotlib (Hunter 2007); Numpy (Harris et al. 2020)

APPENDIX

A. DERIVATION OF CLUMP DENSITY PROFILES

For a clump following a given density profile $\rho(r)$, the relevant terms for calculating the virial parameter and corrections for CO-dark gas that we derive are as follows. Mass within a radius r can be found as

$$M(r) = 4\pi \int_0^r r'^2 \rho(r') dr'. \tag{A1}$$

This leads to gravitational potential energy

$$\Omega_G(r) = -4\pi G \int_0^r r' \; \rho(r') M(r') dr', \tag{A2}$$

where G is the fundamental gravitational constant. We assume a one-dimensional radial velocity dispersion profile $\sigma_v(r)$ following Equation 3. The total kinetic energy can be found as

$$\Omega_K(r) = \frac{3}{2} M(r) \sigma_v^2(r), \tag{A3}$$

and the virial mass $M_{\rm vir}(r)$ follows from requiring $2\Omega_K(r) = -\Omega_G(r)$.

A.1. *Power-law Profile*

We consider a clump following a power-law profile,

$$\rho(r) = \rho_c \; x^{-k}, \tag{A4}$$

where ρ_c is the central density, $x = r/R_0$, and R_0 is an arbitrary radius at which ρ is normalized. From Equation A1, the mass of such a structure is

$$M(r) = \pi \rho_c R_0^3 T_1(r), \tag{A5}$$

where $T_1(r) = \left[\frac{4x^{(3-k)}}{(3-k)}\right]$. From Equation A2, the gravitational term is

$$\Omega_G(r) = -\pi^2 \rho_c^2 G R_0^5 T_2(r) = -\frac{GM(r)^2}{R_0} \frac{T_2(r)}{T_1(r)^2}, \tag{A6}$$

where $T_2(r) = \left[\frac{16x^{(5-2k)}}{(5-2k)(3-k)}\right]$. (The use of $T_1(r)$ and $T_2(r)$ in these expressions will make the parallels with subsequent radial density profiles clearer.) From Equation A3, the kinetic term is

$$\Omega_K(r) = \frac{3}{2} \pi \rho_c R_0^3 T_1(r) \sigma_v^2(r). \tag{A7}$$

Requiring $2\Omega_K = -\Omega_G$, we derive the virial mass

$$M_{\rm vir}(r) = \frac{3\sigma_v^2(r) R_0}{G} \frac{T_1(r)^2}{T_2(r)}, \tag{A8}$$

which is equivalent to the classical virial mass definition (Solomon et al. 1987; MacLaren et al. 1988)

$$M_{\rm vir}(r) = \frac{3(5-2k)}{(3-k)} \frac{r\ \sigma_v^2(r)}{G}. \tag{A9}$$

From Equation 2, the virial parameter is then

$$\alpha_{\rm vir}(r) = \frac{3\sigma_v^2(r)}{\pi \rho_c G R_0^2} \frac{T_1(r)}{T_2(r)}, \tag{A10}$$

which is equivalent to

$$\alpha_{\rm vir}(r) = \frac{3(5-2k)}{4\pi\rho_c} \frac{\sigma_v^2(R_0)}{G} x^{2\beta+k-2}, \tag{A11}$$

under the scaling of the velocity dispersion profile of Equation 3.

In their Appendix A, W10 derived the dark-gas fraction for a power-law density profile with $k < 3$ as

$$f_{DG} = 1 - \left(\frac{R_{CO}}{R_{H_2}}\right)^{3-k}. \tag{A12}$$

R_{H_2} can then be solved for analytically, with $\Sigma(R_{H_2})$, $\alpha_{\rm vir}(R_{H_2})$, and related properties following as demonstrated in §2.1.

A.2. *Power-law Profile with Constant Density Core*

We consider a clump following a power-law density profile with a uniform core of radius R_0,

$$\rho(r) = \begin{cases} \rho_c \text{ for r} < \text{R}_0 \\ \rho_c\ x^{-k} \text{ for r} \geq \text{R}_0, \end{cases} \tag{A13}$$

where ρ_c is the central density and $x = r/R_0$. From Equation A1, the mass within r is

$$M(r) = \begin{cases} \pi \rho_c R_0^3 \left[\frac{4}{3} x^3\right] \text{ for r} < \text{R}_0 \\ \pi \rho_c R_0^3 \Pi_1(r) \text{ for r} \geq \text{R}_0, \end{cases} \tag{A14}$$

where $\Pi_1(r) = \left[\frac{4}{3-k}\left(x^{3-k} - \frac{k}{3}\right)\right]$. From Equation A2, the gravitational term is

$$\Omega_G(r) \equiv \begin{cases} -\frac{16}{15}\pi^2 G\rho_c^2 R_0^5 x^5 \text{ for r} < \text{R}_0 \\ -\pi^2 G\rho_c^2 R_0^5 \Pi_2(r) \text{ for r} \geq \text{R}_0, \end{cases} \tag{A15}$$

where

$$\Pi_2(r) = \begin{cases} \frac{16}{3-k}\left(\frac{x^{5-2k}-1}{5-2k} + \frac{k(1-x^{2-k})}{6-3k} + \frac{3-k}{15}\right) \text{ for k} \neq 2 \\ \frac{16}{3-k}\left(\frac{x^{5-2k}-1}{5-2k} - \frac{k\ln(x)}{3} + \frac{3-k}{15}\right) \text{ for k} = 2. \end{cases} \tag{A16}$$

From Equation A3, the kinetic term is

$$\Omega_K(r) = \begin{cases} 2\pi\rho_c R_0^3 x^3 \sigma_v^2(r) \text{ for r} < \text{R}_0 \\ \frac{3}{2}\pi\rho_c R_0^3 \Pi_1(r)\sigma_v^2(r) \text{ for r} \geq \text{R}_0. \end{cases} \tag{A17}$$

Requiring $2\Omega_K = -\Omega_G$, we then define the virial mass

$$M_{\text{vir}}(r) = \begin{cases} \frac{5R_0\sigma_v^2(r)}{G}x \text{ for r} < \text{R}_0 \\ \frac{3R_0\sigma_v^2(r)}{G}\frac{\Pi_1(r)^2}{\Pi_2(r)} \text{ for r} \geq \text{R}_0, \end{cases} \tag{A18}$$

and finally the virial parameter as

$$\alpha_{\text{vir}}(r) = \begin{cases} \frac{15\sigma_v^2(r)}{4\pi\rho_c G R_0^2}\frac{1}{x^2} \text{ for r} < \text{R}_0 \\ \frac{3\sigma_v^2(r)}{\pi\rho_c G R_0^2}\frac{\Pi_1(r)}{\Pi_2(r)} \text{ for r} \geq \text{R}_0. \end{cases} \tag{A19}$$

We cast this profile in terms of f_{DG} as defined in Equation 1, such that

$$f_{DG} = \begin{cases} 1 - \left(\frac{3-k}{3}\frac{(R_{CO}/R_0)^3}{(R_{H_2}/R_0)^{(3-k)} - \frac{k}{3}}\right) \text{ for R}_{\text{CO}} < \text{R}_0 \\ 1 - \left(\frac{(R_{CO}/R_0)^{(3-k)} - \frac{k}{3}}{(R_{H_2}/R_0)^{(3-k)} - \frac{k}{3}}\right) \text{ for R}_{\text{CO}} \geq \text{R}_0, \end{cases} \tag{A20}$$

assuming $R_{H_2} > R_0$. R_{H_2} can then be solved for analytically, with $\Sigma(R_{H_2})$, $\alpha_{\text{vir}}(R_{H_2})$, and related properties following.

A.3. *Plummer Density Profile*

We consider a clump following a Plummer density profile,

$$\rho(r) = \rho_c\left(\frac{1}{\sqrt{x^2+1}}\right)^\eta, \tag{A21}$$

where ρ_c is the central density, $x = r/R_0$, R_0 is the radius of the central core, and η is the index of the power-law at large radii. From Equation A1, the mass within r is

$$M(r) = \pi\rho_c R_0^3 P_1(r), \tag{A22}$$

where

$$P_1(r) = \begin{cases} 4(x - \arctan(x)) \text{ for } \eta = 2 \\ 2(\arctan(x) - \frac{x}{x^2+1}) \text{ for } \eta = 4. \end{cases} \tag{A23}$$

From Equation A2, the gravitational term can be written as

$$\Omega_G(r) = -\pi^2 G\rho_c^2 R_0^5 P_2(r) = -\frac{GM(r)^2}{R_0}\frac{P_2(r)}{P_1(r)^2}, \tag{A24}$$

where

$$P_2(r) = \begin{cases} 16\left(x - \arctan(x) - \int_0^x \frac{x' \arctan(x')}{x'^2+1} \mathrm{dx'}\right) & \text{for } \eta = 2 \\ \arctan(x) + \frac{x - 4\arctan(x)}{x^2+1} + \frac{2x}{(x^2+1)^2} & \text{for } \eta = 4. \end{cases} \tag{A25}$$

From Equation A3, the kinetic term is

$$\Omega_K(r) = \frac{3}{2}\pi\rho_c R_0^3 \sigma_v^2(r) P_1(r). \tag{A26}$$

We can then derive the virial mass through requiring $2\Omega_K = -\Omega_G$,

$$M_{\rm vir}(r) = \frac{3R_0\sigma_v^2(r)}{G}\frac{P_1(r)^2}{P_2(r)}, \tag{A27}$$

and finally the virial parameter

$$\alpha_{\rm vir}(r) = \frac{3\sigma_v^2(r)}{\pi\rho_c G R_0^2}\frac{P_1(r)}{P_2(r)}. \tag{A28}$$

We adapt this profile into f_{DG} as defined in Equation 1, such that

$$f_{DG} = \begin{cases} 1 - \left(\frac{\frac{R_{CO}}{R_0} - \arctan(R_{CO}/R_0)}{\frac{R_{H_2}}{R_0} - \arctan(R_{H_2}/R_0)}\right) & \text{for } \eta = 2 \\ 1 - \left(\frac{\arctan(R_{CO}/R_0) - \frac{(R_{CO}/R_0)}{(R_{CO}/R_0)^2+1}}{\arctan(R_{H_2}/R_0) - \frac{(R_{H_2}/R_0)}{(R_{H_2}/R_0)^2+1}}\right) & \text{for } \eta = 4. \end{cases} \tag{A29}$$

R_{H_2} can then be solved for numerically to obtain estimates of $\Sigma(R_{H_2})$, $\alpha_{\rm vir}(R_{H_2})$, and related properties.

REFERENCES

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 710, 133

Arzoumanian, D., André, P., Didelon, P., et al. 2011, A&A, 529, L6

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123

Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F., & Zamora-Avilés, M. A. 2011, MNRAS, 411, 65

Bertoldi, F., & McKee, C. F. 1992, ApJ, 395, 140

Bertram, E., Klessen, R. S., & Glover, S. C. O. 2015, MNRAS, 451, 196

Bigiel, F., Leroy, A. K., Walter, F., et al. 2011, ApJL, 730, L13

Bolatto, A. D., Leroy, A. K., Rosolowsky, E., Walter, F., & Blitz, L. 2008, ApJ, 686, 948

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207

Bolatto, A. D., Leroy, A. K., Jameson, K., et al. 2011, ApJ, 741, 12

Brunt, C. M., Heyer, M. H., & Mac Low, M. M. 2009, A&A, 504, 883

Caselli, P., Benson, P. J., Myers, P. C., & Tafalla, M. 2002, ApJ, 572, 238

Caselli, P., & Myers, P. C. 1995, ApJ, 446, 665

Chen, H. H.-H., Offner, S. S. R., Pineda, J. E., et al. 2020, arXiv e-prints, arXiv:2006.07325

Chen, H. H.-H., Pineda, J. E., Goodman, A. A., et al. 2019, ApJ, 877, 93

Chevance, M., Madden, S. C., Fischer, C., et al. 2020, MNRAS, 494, 5279

Chevance, M., Kruijssen, J. M. D., Vazquez-Semadeni, E., et al. 2020, SSRv, 216, 50

Clark, P. C., & Glover, S. C. O. 2015, MNRAS, 452, 2057

de los Reyes, M. A. C., & Kennicutt, Robert C., J. 2019, ApJ, 872, 16

Dessauges-Zavadsky, M., Richard, J., Combes, F., et al. 2019, NatAs, 3, 1115

Dobbs, C. L. 2015, MNRAS, 447, 3390

Field, G. B., Blackman, E. G., & Keto, E. R. 2011, MNRAS, 416, 710

Fukui, Y., Kawamura, A., Minamidani, T., et al. 2008, ApJS, 178, 56

Galametz, M., Madden, S., Galliano, F., et al. 2009, A&A, 508, 645
Girichidis, P., Federrath, C., Banerjee, R., & Klessen, R. S. 2011, MNRAS, 413, 2741
Glover, S. C. O., & Clark, P. C. 2012, MNRAS, 421, 9
Glover, S. C. O., Federrath, C., Mac Low, M. M., & Klessen, R. S. 2010, MNRAS, 404, 2
Glover, S. C. O., & Mac Low, M.-M. 2011, MNRAS, 412, 337
Gong, M., Ostriker, E. C., & Kim, C.-G. 2018, ApJ, 858, 16
Gong, M., Ostriker, E. C., Kim, C.-G., & Kim, J.-G. 2020, ApJ, 903, 142
Goodman, A. A., Barranco, J. A., Wilner, D. J., & Heyer, M. H. 1998, ApJ, 504, 223
Gordon, K. D., Meixner, M., Meade, M., et al. 2011, AJ, 142, 102
Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, Sci, 307, 1292
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Heyer, M., Krawczyk, C., Duval, J., & Jackson, J. M. 2009, ApJ, 699, 1092
Heyer, M. H., & Brunt, C. M. 2004, ApJL, 615, L45
Hughes, A., Wong, T., Ott, J., et al. 2010, MNRAS, 406, 2065
Hughes, A., Meidt, S. E., Colombo, D., et al. 2013, ApJ, 779, 46
Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
Jameson, K. E., Bolatto, A. D., Leroy, A. K., et al. 2016, ApJ, 825, 12
Jameson, K. E., Bolatto, A. D., Wolfire, M., et al. 2018, ApJ, 853, 111
Juvela, M., Malinen, J., Montillaud, J., et al. 2018, A&A, 614, A83
Kalari, V. M., Rubio, M., Saldaño, H. P., & Bolatto, A. D. 2020, MNRAS, 499, 2534
Kauffmann, J., Pillai, T., & Goldsmith, P. F. 2013, ApJ, 779, 185
Kennicutt, Robert C., J., Calzetti, D., Walter, F., et al. 2007, ApJ, 671, 333
Kim, S., Rosolowsky, E., Lee, Y., et al. 2007, ApJS, 171, 419
Krumholz, M. R. 2012, ApJ, 759, 9
Langer, W. D., Velusamy, T., Pineda, J. L., Willacy, K., & Goldsmith, P. F. 2014, A&A, 561, A122
Larson, R. B. 1981, MNRAS, 194, 809
Lee, E. J., Chang, P., & Murray, N. 2015, ApJ, 800, 49
Lee, M.-Y., Stanimirović, S., Douglas, K. A., et al. 2012, ApJ, 748, 75
Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782
Leroy, A. K., Bolatto, A., Bot, C., et al. 2009, ApJ, 702, 352
Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021, ApJS, 257, 43
Li, Q., Narayanan, D., Davè, R., & Krumholz, M. R. 2018, ApJ, 869, 73
Lin, Y., Wyrowski, F., Liu, H. B., et al. 2021, arXiv e-prints, arXiv:2112.01115
MacLaren, I., Richardson, K. M., & Wolfendale, A. W. 1988, ApJ, 333, 821
Madden, S. C., Galliano, F., Jones, A. P., & Sauvage, M. 2006, A&A, 446, 877
Madden, S. C., Cormier, D., Hony, S., et al. 2020, A&A, 643, A141
Muraoka, K., Homma, A., Onishi, T., et al. 2017, ApJ, 844, 98
Ochsendorf, B. B., Meixner, M., Roman-Duval, J., Rahman, M., & Evans, N. J. 2017, ApJ, 841, 109
Oka, T., Hasegawa, T., Sato, F., et al. 2001, ApJ, 562, 348
Palmeirim, P., André, P., Kirk, J., et al. 2013, A&A, 550, A38
Paradis, D., Dobashi, K., Shimoikura, T., et al. 2012, A&A, 543, A103
Parmentier, G. 2020, ApJ, 893, 32
Parmentier, G., & Pasquali, A. 2020, ApJ, 903, 56
Passot, T., Pouquet, A., & Woodward, P. 1988, A&A, 197, 228
Pattle, K. 2016, MNRAS, 459, 2651
Pineda, J. E., Goodman, A. A., Arce, H. G., et al. 2010, ApJL, 712, L116
Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, A&A, 554, A103
Pirogov, L. E. 2009, ARep, 53, 1127
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2011, A&A, 536, A19
Plummer, H. C. 1911, MNRAS, 71, 460
Ridge, N. A., Francesco, J. D., Kirk, H., et al. 2006, AJ, 131, 2921
Rubele, S., Girardi, L., Kerber, L., et al. 2015, MNRAS, 449, 639
Rubio, M., Elmegreen, B. G., Hunter, D. A., et al. 2015, Nature, 525, 218
Saldaño, H. P., Rubio, M., Jameson, K., & Bolatto, A. D. 2018, BAAA, 60, 192
Schneider, N., André, P., Könyves, V., et al. 2013, ApJL, 766, L17
Schruba, A., Kruijssen, J. M. D., & Leroy, A. K. 2019, ApJ, 883, 2

Schruba, A., Leroy, A. K., Walter, F., et al. 2012, AJ, 143, 138

Schruba, A., Leroy, A. K., Kruijssen, J. M. D., et al. 2017, ApJ, 835, 278

Shetty, R., Beaumont, C. N., Burton, M. G., Kelly, B. C., & Klessen, R. S. 2012, MNRAS, 425, 720

Shetty, R., Glover, S. C., Dullemond, C. P., & Klessen, R. S. 2011a, MNRAS, 412, 1686

Shetty, R., Glover, S. C., Dullemond, C. P., et al. 2011b, MNRAS, 415, 3253

Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, MNRAS, 441, 1628

Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, ApJ, 319, 730

Sun, J., Leroy, A. K., Schruba, A., et al. 2018, ApJ, 860, 172

Sun, J., Leroy, A. K., Ostriker, E. C., et al. 2020, ApJ, 892, 148

Szűcs, L., Glover, S. C. O., & Klessen, R. S. 2016, MNRAS, 460, 82

Tang, M., Liu, T., Qin, S.-L., et al. 2018, ApJ, 856, 141

Tatematsu, K., Umemoto, T., Kandori, R., & Sekimoto, Y. 2004, ApJ, 606, 333

Traficante, A., Duarte-Cabral, A., Elia, D., et al. 2018, MNRAS, 477, 2220

Valdivia-Mena, M. T., Rubio, M., Bolatto, A. D., Saldaño, H. P., & Verdugo, C. 2020, A&A, 641, A97

Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061

Velusamy, T., Langer, W. D., Pineda, J. L., et al. 2010, A&A, 521, L18

Volgenau, N. H., Mundy, L. G., Looney, L. W., & Welch, W. J. 2006, ApJ, 651, 301

Whitworth, A. P., & Ward-Thompson, D. 2001, ApJ, 547, 317

Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191

Wong, T., Hughes, A., Fukui, Y., et al. 2009, ApJ, 696, 370

Wong, T., Hughes, A., Tokuda, K., et al. 2019, ApJ, 885, 50

Xu, D., Li, D., Yue, N., & Goldsmith, P. F. 2016, ApJ, 819, 22

Zucker, C., Goodman, A., Alves, J., et al. 2021, ApJ, 919, 35

3

Sequential Star Formation in the Young SMC Region NGC 602

Draft version December 15, 2022
Typeset using LaTeX **twocolumn** style in AASTeX631

Sequential Star Formation in the Young SMC Region NGC 602: Insights from ALMA

Theo J. O'Neill [1] Rémy Indebetouw [1, 2] Karin Sandstrom [3] Alberto D. Bolatto [4]
Katherine E. Jameson [5] Lynn R. Carlson,[6] Molly K. Finn [1] Margaret Meixner [7] Elena Sabbi [8] and
Marta Sewiło [9, 4, 10]

[1] *Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA*
[2] *National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA*
[3] *Center for Astrophysics and Space Sciences, Department of Physics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA*
[4] *Department of Astronomy, University of Maryland, College Park, MD 20742, USA*
[5] *CSIRO Space and Astronomy, ATNF, PO Box 1130, Bentley, WA 6102, Australia*
[6] *Experimental College, Tufts University, Medford, MA 02155*
[7] *SOFIA Science Mission Operations/USRA, NASA Ames Research Center, Bldg. N232, M/S 232-12, P.O. Box 1, Moffett Field, CA 94035-0001*
[8] *Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[9] *Exoplanets and Stellar Astrophysics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*
[10] *Center for Research and Exploration in Space Science and Technology, NASA Goddard Space Flight Center, Greenbelt, MD 20771*

ABSTRACT

NGC 602 is a young, low-metallicity star cluster in the "Wing" of the Small Magellanic Cloud. We reveal the recent evolutionary past of the cluster through analysis of high-resolution (∼0.4 pc) Atacama Large Millimeter/submillimeter Array observations of molecular gas in the associated Hii region N90. We identify 110 molecular clumps ($R < 0.8$ pc) traced by CO emission, and study the relationship between the clumps and associated young stellar objects (YSOs) and pre-main-sequence (PMS) stars. The clumps have high virial parameters (typical $\alpha_{\rm vir} = 4-11$) and may retain signatures of a collision in the last $\lesssim$8 Myr between Hi components of the adjacent supergiant shell SMC-SGS 1. We obtain a CO-bright-to-H_2 gas conversion factor of $X_{CO,B} = (3.4 \pm 0.2) \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$, and correct observed clump properties for CO-dark H_2 gas to derive a total molecular gas mass in N90 of $16,600 \pm 2,400$ $M_\odot$. We derive a recent ($\lesssim$ 1 Myr) star formation rate of 130 ± 30 $M_\odot$ Myr^{-1} with an efficiency of $8 \pm 3\%$ assessed through comparing total YSO mass to total molecular gas mass. Very few significant radial trends exist between clump properties or PMS star ages and distance from NGC 602. We do not find evidence for a triggered star formation scenario among the youngest ($\lesssim$2 Myr) stellar generations, and instead conclude that a sequential star formation process in which NGC 602 did not directly cause recent star formation in the region is likely.

Keywords: Hii regions (694), Interstellar medium (847), Molecular clouds (1072), Small Magellanic Cloud (1468), Star Formation (1569)

1. INTRODUCTION

The young open cluster NGC 602 and associated Hii region N90 are cradled in the "Wing" of the Small Magellanic Cloud (SMC) by the supergiant shell SMC-SGS 1 (hereafter SGS 1, Meaburn 1980, shown in Figure 1a). The SMC is an ideal location to study the effects of environment on the progression of star formation: it is a low metallicity (Z ∼ 1/5 $Z_\odot$, Russell & Dopita 1992; Lee et al. 2005) and low gas surface density environment (N ∼ $(2-8) \times 10^{21}$ cm^{-2}, Leroy et al. 2007; Welty et al. 2012), and is located at a distance of only ∼60.6 kpc (Hilditch et al. 2005). Under these conditions, the relationship between star-forming gas and commonly used observables like CO line emission is expected to depart from behaviors found at higher metallicities and densities. Molecular clouds experience significant photodissociation in gas with little dust (Gordon et al. 2011) and enhanced interstellar radiation fields (Madden et al. 2006; Gordon et al. 2008; Sandstrom et al. 2010); the fraction of H_2 that is "CO-dark" also increases with decreasing metallicity (Wolfire et al. 2010; Glover & Mac Low 2011; Szűcs et al. 2016).

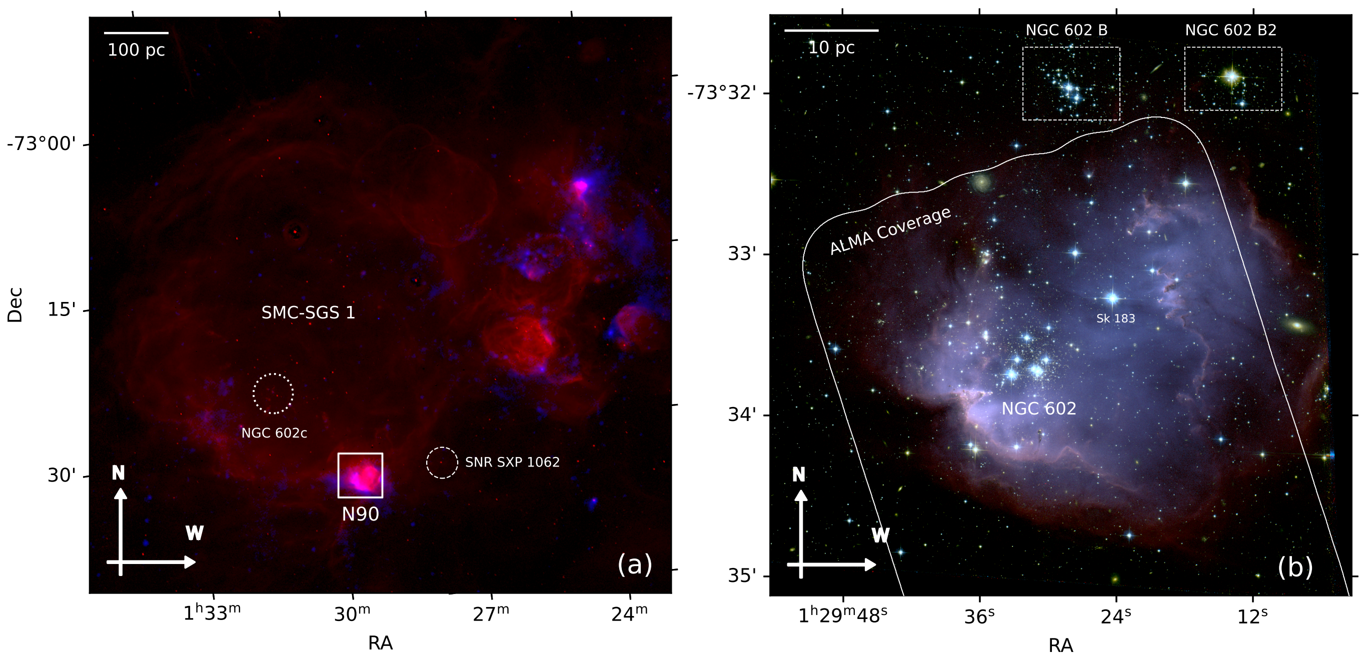

Figure 1. *(a:)* Image of the supergiant shell SMC-SGS 1 in the SMC Wing. [R, B] = [Magellanic Cloud Emission-Line Survey (MCELS) Hα image, Herschel SPIRE 250 μm image]. The extent of the Hubble Space Telescope (HST) image of N90 shown in (b) is outlined with a white square. SNR SXP 1062 is outlined with a white dashed circle and the cluster NGC 602c is marked with a white dotted circle. *(b:)* HST ACS image of N90. [R, G, B] = [Hα + F814W, F555W + F814W, F555W]. The central cluster NGC 602 is labeled, and the extent of the ALMA 12m + 7m coverage is outlined in white. The clusters NGC 602 B and NGC 602 B2 are marked with white dashed rectangles, and the massive O3 star Sk 183 is labeled.

The Wing marks the transition between the comparatively molecule-rich inner area of the SMC and the HI-dominated outer region leading to the Magellanic Bridge, which possesses even lower surface densities and metallicities than the main body of the SMC (Rolleston et al. 1999; Lehner et al. 2008; Gordon et al. 2009; Welty et al. 2012). NGC 602 and N90 have been extensively studied historically (e.g., Henize 1956; Westerlund 1964; Hodge 1983; Hutchings et al. 1991), and renewed interest in recent years has resulted in the region being remarkably well-characterized on a variety of spatial scales through a large range of wavelength regimes. When considered with its isolated location within the diffuse Wing, NGC 602/N90 presents a valuable opportunity for tests of star formation theory under dramatically different conditions from the Solar neighborhood.

The rate of star formation in the Wing has been increasing over the last 0.2 Gyr (Rubele et al. 2015, 2018), especially in the area surrounding N90 and adjacent ~500 pc diameter SGS 1 shell (Figure 1a). Ramachandran et al. (2019) performed a spectroscopic investigation of OB stars within SGS 1 and suggested that massive star formation has been ongoing in the past 100 Myr, including an extended star-formation event between 30 - 40 Myr ago. Fulmer et al. (2020) extended this work using near-UV and optical photometry and observed no radial gradient in stellar ages across SGS 1, concluding that star formation in this section of the Wing has resulted from a combination of stochastic star formation mixed with some star formation stimulated by the expansion of SGS 1.

There is a supernova remnant SNR SXP 1062 a projected ~120 pc to the west of N90 centered around a Be/X-ray pulsar binary, with age estimates ranging between (2–4) $\times$ 10^4 years (Hénault-Brunet et al. 2012) and (1–2.5) $\times$ 10^4 years (Haberl et al. 2012); however, this remnant (see Figure 1a, diameter ~25 pc) has not yet reached N90 and is unlikely to be associated with star formation in the region. The cluster NGC 602c (Westerlund 1964) and associated small, very faint HII region are located ~190 pc to the northeast of N90 (Figure 1a), hosting the massive WO-type star Sk 188 and several other young, massive stars (Ramachandran et al. 2019). Given its distance, though, it too is unlikely to be directly related to recent star formation in N90.

The stellar population of N90 itself consists of a mixture of young stars concentrated around the central OB association NGC 602 and a scattered group of much older stars that are likely related to the general SMC field population. Figure 1b presents a closer view of

NGC 602/N90, with two adjacent stellar concentrations to the north, NGC 602 B and NGC 602 B2, also identified; these clusters have estimated ages of up to 50–80 Myr and 47–160 Myr, respectively (Schmalzl et al. 2008; De Marchi et al. 2013). Through analysis of Hubble Space Telescope (HST) photometry, De Marchi et al. (2013) found that one-third of pre-main-sequence (PMS) stars in N90 itself are likely $\gtrsim$30 Myr old and one-half likely younger than 5 Myr.

The cause of the formation of the central cluster NGC 602 and more recent star formation event has been a subject of debate. Cignoni et al. (2009) found that the star formation rate in N90 began to increase $\sim$10 Myr ago and has peaked in the last $\sim$2.5 Myr. Using velocity maps derived from a survey of neutral hydrogen (Staveley-Smith et al. 1997) and optical and mid-IR HST data, Nigra et al. (2008) suggested that compression and turbulence from the interactions of expanding HI shells $\sim$7 Myr ago is responsible for the formation of NGC 602 (with the Northern shell corresponding to SGS 1). Alternatively, Fukui et al. (2020) proposed that compression resulting from a collision of two 500–600 pc radii HI clouds $\sim$8 Myr ago triggered the formation of NGC 602, and that SGS 1 is the disturbed region evacuated by this cloud collision.

Through analysis of HST optical and Spitzer Space Telescope (Spitzer) IR photometry, Carlson et al. (2007) and Carlson et al. (2011) concluded that NGC 602 formed $\sim$4 Myr ago, with a population of low-mass PMS stars forming $\sim$0.9 Myr later. They also identified 45 candidate young stellar objects (YSOs) and proposed that star formation has propagated outwards from NGC 602 to the "rim" of the HII region, with the youngest YSOs in N90 forming in the last $\sim$1 Myr. Gouliermis et al. (2007) and Gouliermis et al. (2012) analyzed the clustered spatial distribution of YSOs and PMS stars across N90 and suggested the formation of NGC 602 triggered progressive, ongoing star formation in the last 2.5 Myr in sub-clusters of PMS stars along the rim. Alternatively, De Marchi et al. (2013) suggested that a sequential star formation process, in which the formation of the earliest generations of young stars in the region did not significantly influence the formation of younger stellar generations, was more likely to have occurred.

New, high-resolution (1.3" or 0.4 pc) Atacama Large Millimeter/submillimeter Array (ALMA) data presented here clarify the amount and nature of dense gas in N90, and the history of the region's evolution. In §2, we describe the observations and analysis methods used. In §3, we analyze the structure of small molecular clumps $\sim$2–23 pc from the central cluster, as well as their association with the populations of PMS stars and YSOs. We discuss if the clumps can reveal the formation history of the region, exploring signatures of large-scale HI collisions in SGS 1 as well as evidence for feedback from NGC 602 triggering ongoing star formation along the N90 rim. In §4, we examine common metrics of molecular cloud stability and star formation efficiency on the scales of both individual clumps and the entire region. We conclude in §5 with a discussion of the implications of our results for the evolution of N90, and compare star formation progression in N90 to solar-metallicity, higher-density environments.

2. OBSERVATIONS AND ANALYSIS

2.1. *ALMA Data*

The NGC602/N90 region was observed by ALMA project 2016.1.00360.S. A 150 point mosaic was observed with a 48 antenna compact configuration of the 12m array (MOUS uid://A001/X88f/X2a2), for 49 minutes on source on December 30, 2016 with a mean PWV of 1.3mm. J0635-7516 (599 mJy at 230 GHz), J0334-4008 (432 mJy), and J0102-7546 (184 mJy) were used as bandpass, amplitude, and phase calibrator, respectively. A slightly larger region of the sky was observed in a 60 point mosaic using the 7m ACA eight times between October 12 and 24, 2016, for a total of 416 minutes on source, mostly at a PWV of $\sim$ 0.5 mm. J0006-0623 (3.6 Jy at 230 GHz) or J0522-3627 (2.8 Jy) were used as bandpass calibrator, Uranus as amplitude calibrator, and J0450-8101 (1.6 Jy) as phase calibrator. The spectral setup contains 3 spectral windows with 122.07 kHz channels each, centered on the ^{12}CO, ^{13}CO, and ^{12}C^{18}O J=2–1 transitions. These windows have 1920, 2048 channels and 234.4 MHz, 250 MHz for the 12m and 7m arrays, respectively. Additionally observed was a 2 GHz wide, 128 channel spectral window centered at 232.86 GHz. The native spectral channel spacing is 0.17 km/s, with a resolution for ^{12}CO of 0.184 km/s.

The data were processed with the ALMA Pipeline-CASA56-P1-B v42866 released with CASA 5.6.1-8 (McMullin et al. 2007), using the default recipes. Standard flagging resulted in 27% of the 12m data and 35% of the 7m data being flagged. The imaging stages of the ALMA pipeline correctly detected strong line emission and subtracted from the visibilities a linear fit to the continuum, excluding the line spectral ranges.

Calibrated visibilities from both arrays were imaged together using CASA::tclean, Total Power data were also observed with ALMA, and processed with Pipeline-Cycle4-R2-B packaged with CASA 4.7.0. The Total Power data were feathered with the interferometer data, but the interferometers alone recovered >95±5% of the ^{12}CO 2–1 flux and 100±10% of the ^{13}CO 2–1 flux in the

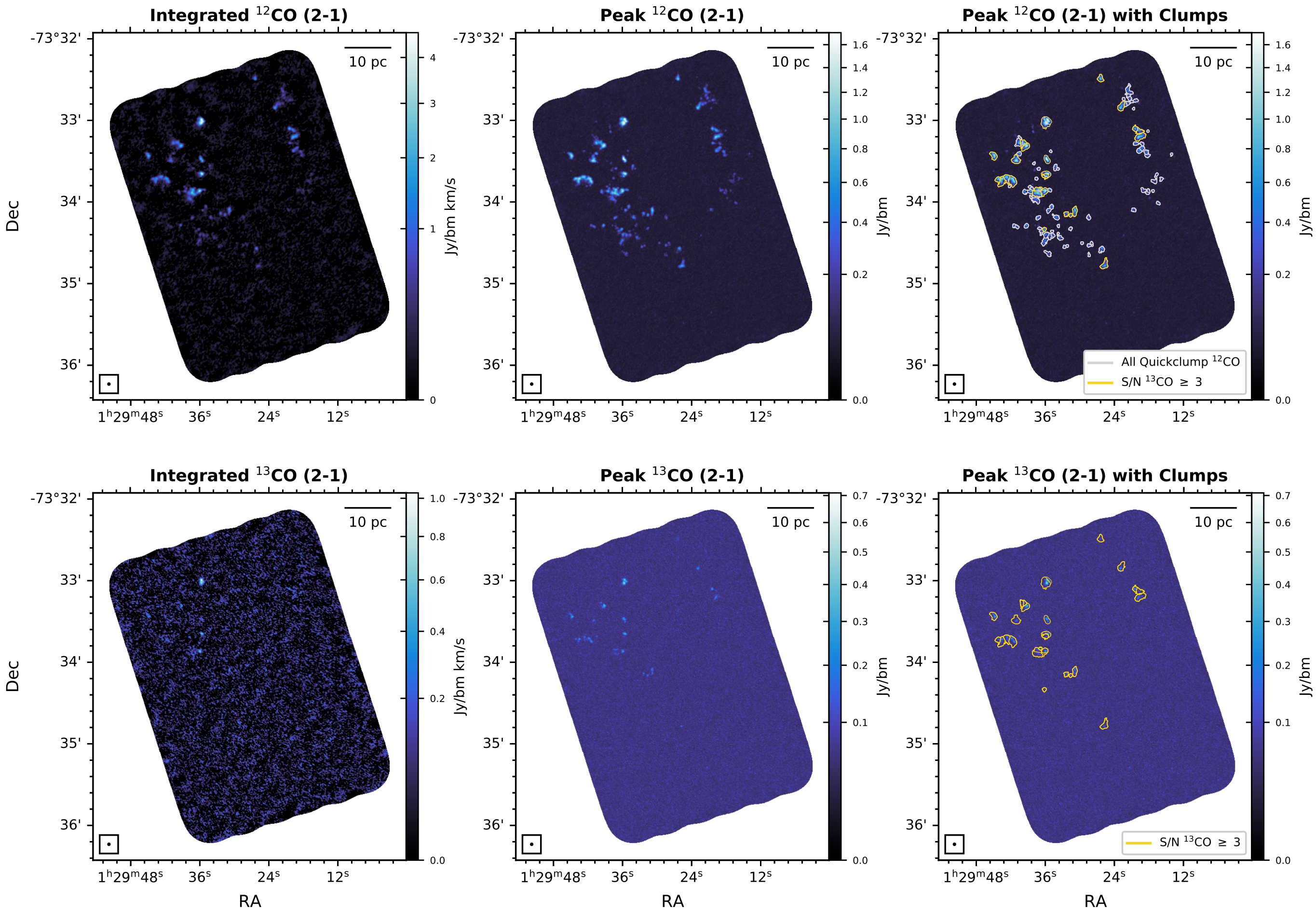

Figure 2. Images of combined 12m + 7m ALMA ^{12}CO (2–1) and ^{13}CO (2–1) data. *Top left:* Integrated ^{12}CO (2–1) intensity. *Top center:* Peak ^{12}CO (2–1) intensity. *Top right:* Peak ^{12}CO (2–1) intensity with contours of clumps identified by the `quickclump` algorithm in grey. Contours of the subset of these clumps with ^{13}CO (2–1) S/N $\geq$ 3 are in gold. *Bottom left:* Integrated ^{13}CO (2–1) intensity. *Bottom center:* Peak ^{13}CO (2–1) intensity. *Bottom right:* Peak ^{13}CO (2–1) intensity with contours of ^{13}CO (2–1) S/N $\geq$ 3 clumps in gold. In all panels, the synthesized beam size is shown in the lower left corner.

Total Power image (i.e., within the absolute calibration uncertainty), and the feathered image has higher noise on large angular scales, so we use the interferometer-only images for the clump analysis presented here. The combined 12m + 7m integrated intensity images in ^{12}CO 2–1 and ^{13}CO 2–1 are shown in Figure 2.

2.1.1. *Clump Extraction*

Molecular cloud emission is often decomposed into discrete "clumps" (R $\sim$ 1 pc) or "cores" (R $\lesssim$ 0.1 pc) to enable analysis of the characteristics of the complex structures within the clouds. Non-hierarchical clump identification methods like `clumpfind` (Williams et al. 1994) or `quickclump` (Sidorin 2017) segment position-position-velocity (PPV) cubes by identifying local maxima and assigning adjacent pixels above a minimum intensity I_{min} with a minimum intensity difference δI between the local maximum and the highest adjacent saddle point to discrete clumps. Diffuse molecular gas is typically not visible in observable CO emission in the SMC, as the low metallicity, low dust-to-gas ratio, and higher interstellar radiation fields of the SMC result in observed CO emission being segmented into more discrete, "clumpier" structures (e.g., Muraoka et al. 2017; Jameson et al. 2018) than in the Milky Way or Large Magellanic Cloud (LMC) where filamentary structures are more common (e.g., Saigo et al. 2017; Indebetouw et al. 2020). Non-hierarchical methods like `clumpfind` are then more suited to environments like the SMC due to this tendency towards discrete CO structures.

We used a version of the `quickclump` python implementation[1] modified to include a parameter defining a required minimum peak intensity of a clump, I_{minpk} (Indebetouw et al. 2020)[2]. The addition of this parameter ensures that clumps with relatively high signal to noise ratio (S/N) peaks are able to have emission in their envelopes assigned down to the level of the noise, which allows for the emission of the clump to be captured more completely, while avoiding introducing many additional clumps with low S/N peaks.

We applied the modified `quickclump` algorithm to the ^{12}CO observations, with a minimum intensity $I_{min} = 4\sigma$, minimum change in intensity between leaves $\delta I = 3\sigma$, minimum peak intensity $I_{min,pk} = 8\sigma$, and minimum number of pixels $n_{min,pix} \simeq 2$ beams. Finally, we required that each clump have ^{12}CO velocity dispersions (as calculated in §2.1.4) greater than the spectral resolution for that line ($\sigma_v \gtrsim 0.18$ km s^{-1}). This yielded a total of 110 clumps.

2.1.2. *Molecular Column Density*

We assumed local thermal equilibrium (LTE) conditions to calculate column densities for clumps with significant ^{13}CO emission. Since many of the clumps identified in ^{12}CO do not appear to have strong corresponding emission in ^{13}CO, we required that clumps have a ^{13}CO signal-to-noise (S/N) ratio within the clump's ^{12}CO boundaries of S/N $\geq$3 above 4σ to apply the LTE method. Our noise estimate was derived from the RMS noise in the Southern half of the cube ($\delta \lesssim -73°35'$) where no strong CO emission was detected in either line.

Only 29 of the 110 clumps (26%) fulfill this requirement, while the remaining 81 clumps (74%) do not possess any strong ^{13}CO emission. To overcome this obstacle, our mass estimation method takes place in two parts (described in this and the following subsection §2.1.3) and is similar to the approach taken by Wong et al. (2022), who in their study of the LMC's 30 Doradus region found that 53% of CO-detected clumps were only traced by ^{12}CO and did not have corresponding ^{13}CO emission. The majority of the clumps in N90 with correponding ^{13}CO emission are located in the NE rim and non-rim sections near NGC 602, with several others on the NW rim near the massive O3 star Sk 183 (see definitions of these subregions in Figure 7a).

We assumed that ^{12}CO is optically thick and that its excitation temperature T_{ex} is a function of brightness temperature,

$$T_{ex} = \frac{11.1\mathrm{K}}{\ln(\frac{11.1}{I_{12}+0.19}+1)}, \tag{1}$$

where I_{12} is the ^{12}CO(2–1) intensity in K. Since this result only relies on ^{12}CO emission, we derived this quantity for all clumps. Calculated T_{ex} range from 6.5–28 K (Figure 7c).

We assumed that ^{13}CO is optically thin, that ^{13}CO and ^{12}CO share the same T_{ex}, and that their relative abundance is constant. We found the ^{13}CO (2–1) optical depth of each PPV pixel for the 29 ^{13}CO-traced clumps as (Garden et al. 1991; Bourke et al. 1997; Indebetouw et al. 2013; Wong et al. 2017),

$$\tau_0^{13} = -\ln\left[1 - \frac{T_B^{13}}{10.6}\left\{\frac{1}{e^{10.6/T_{ex}}-1} - \frac{1}{e^{10.6/2.7}-1}\right\}^{-1}\right], \tag{2}$$

and column density N(^{13}CO) as

$$\mathrm{N}(^{13}\mathrm{CO}) = 1.2\times10^{14}\frac{(\mathrm{T_{ex}}+0.88\ \mathrm{K})\ \mathrm{e}^{5.29/\mathrm{T_{ex}}}}{1-\mathrm{e}^{-10.6/\mathrm{T_{ex}}}}\int\tau_\nu^{13}\mathrm{d}\nu. \tag{3}$$

The maximum N(^{13}CO) was 2.6×10^{16} cm^{-2}. We assumed an abundance ratio of H_2 to ^{13}CO of 1.25×10^6 following Jameson et al. (2018) for the SMC, such that N(H_2) = 1.25×10^6 N(^{13}CO). This abundance ratio is the combination of ^{12}C/H and ^{12}CO/^{13}CO abundance ratios; the former is constrained by UV absorption measurements (see references in Tchernyshyov et al. 2015) to ±40%, and the latter by NLTE modeling of CO emission lines at ∼10pc resolution (Nikolić et al. 2007), with another 40% uncertainty. The total ^{13}CO/H_2 could then be off by a factor of two, most likely in the direction that underestimates H_2. The maximum H_2 column density observed was 3.2×10^{22} cm^{-2}.

We calculated LTE masses for the clumps with significant ^{13}CO detections as

$$M_{LTE} = 1.36\ m_{H_2}\sum\mathrm{N}(\mathrm{H_2})\delta\mathrm{x}\delta\mathrm{y}, \tag{4}$$

where 1.36 is a factor derived from cosmic abundances to convert from H_2 mass to total mass including helium, δx and δy are pixel sizes, and m_{H_2} is the mass of an H_2 molecule. LTE clump masses ranged from 13 $M_\odot$ to 286 $M_\odot$, with a median mass of 64 $M_\odot$ and total mass of all ^{13}CO-traced clumps of 2435 ± 330 $M_\odot$.

Of course, it is possible that the clumps are not in LTE. The systematic effects of using the LTE approximation can be understood by analyzing many non-LTE models. From large grids of Radex (van der Tak et al. 2007) models, we find that under typical molecular cloud conditions, where ^{12}CO has optical depths of a few and

[1] https://github.com/vojtech-sidorin/quickclump/

[2] https://github.com/indebetouw/quickclump

^{13}CO between ∼0.5–2, the LTE method tends to slightly (10-20%) overestimate the ^{12}CO excitation temperature (since its optical depth is less than the infinite assumed). On the other hand, if ^{13}CO has optical depth <1, its excitation temperature can be lower than that of ^{12}CO by up to a factor of 2. For very cold clouds ($T_K \lesssim$10 K), the LTE method underestimates the true ^{13}CO column density by up to a factor of ∼2, and for very dense ($n_H \gtrsim$5000 cm^{-3}) and warm ($T_K \gtrsim$50 K) clouds, the LTE method overestimates the ^{13}CO column density by up to a factor of ∼2. However, the calculated ^{13}CO column density from the LTE method is within 25% of the true value for fairly wide ranges of parameter space: ^{12}CO column densities between 10^{16} and $3{\times}10^{18}$, and 15 K$\lesssim$T$_K$ $\lesssim$65 K. For the N90 clumps we apply the LTE method to, the average N(^{12}CO) is $5.9{\times}10^{17}$ and the average T_{ex} is 19 K, so we expect the LTE mass estimates to be reasonable.

2.1.3. *CO-to-H$_2$ Conversion Factor*

Although the LTE method described above is powerful when multiple lines are traced, it is limited and becomes less reliable for clumps with weak ^{13}CO detections. To circumvent this issue, we derived an $X_{\rm CO}$ CO-to-H$_2$ conversion factor from the clumps traced by ^{13}CO, and applied it to the 81 clumps with ^{13}CO S/N ≥ 3 to obtain estimates of their masses. We use the notation $X_{CO,B}$ to indicate that this factor is only intended to include gas that is ^{12}CO-"bright", and does not account for diffuse "CO-dark" gas in clump envelopes. We discuss the role of CO-dark gas in N90 in §4.2.

We fit $X_{CO,B}$ as

$$\mathrm{N(H_2) = X_{CO,B}\ W_{12CO}} \tag{5}$$

where N(H$_2$) is the total H$_2$ column density in units of cm^{-2} and W$_{12\rm CO}$ is the integrated ^{12}CO line intensity in units of K km s^{-1}. In the Milky Way, Bolatto et al. (2013) recommended an average value of $X_{\rm CO,\,MW}$ = 2×10^{20} cm^{-2} (K km s^{-1})$^{-1}$. We performed ordinary least squares regression with a heteroskedasticity-consistent standard error estimator ("HC3" in the python package `statsmodels`, Seabold & Perktold (2010)). The best-fit slope was $X_{\rm CO,B}$ = $(3.4 \pm 0.2) \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$, or equivalently $X_{\rm CO,B} \sim 1.7\ X_{\rm CO,\,MW}$.

This fit is shown in Figure 3. The 95% confidence interval for $X_{\rm CO,B}$ is [3.0×10^{20}, 3.75×10^{20}] cm^{-2} (K km s^{-1})$^{-1}$, and standard fit diagnostics suggest this model is adequate (coefficient of determination R^2 = 0.95, F-test statistic calculated for robust covariance of $F = 309.8$ with $p < 0.001$). The $X_{\rm CO,B} \sim 1.7\ X_{CO,MW}$ we derived is consistent with, albeit slightly lower than, values of $X_{\rm CO,B}$ found in other SMC regions observed at parsec-scales in ^{12}CO and/or ^{13}CO: In the Magellanic Bridge, Kalari et al. (2020) and Valdivia-Mena et al. (2020) derived values of $X_{\rm CO}$ of ∼2–4 $X_{\rm CO,MW}$, while Muraoka et al. (2017) found $X_{\rm CO} \sim 4\ X_{\rm CO,MW}$ in the star forming region N83C in the southeast Wing. Jameson et al. (2018) similarly found an average $X_{\rm CO} \sim 5\ X_{\rm CO,MW}$ across several star forming regions in the Southwest Bar of the SMC at <3 pc scales.

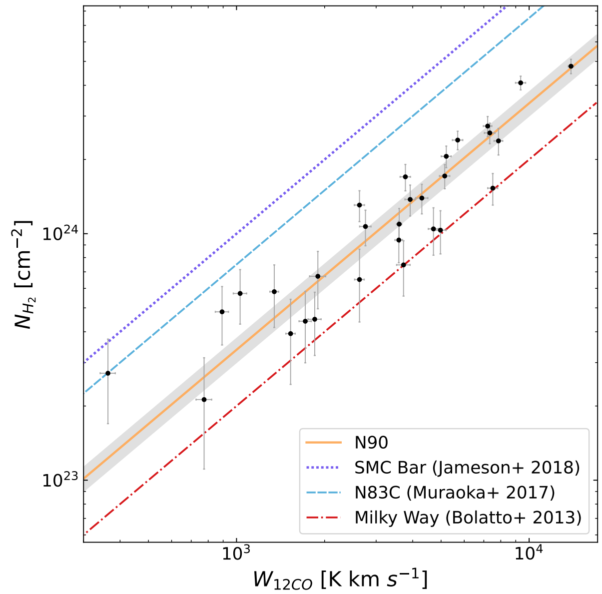

Figure 3. Total column density N(H$_2$) is compared to total integrated ^{12}CO intensity W_{12CO} for clumps with ^{13}CO S/N ≥ 3. The solid orange line shows the best fit CO-to-H$_2$ conversion factor $X_{CO,B}$ for the clumps in N90, $X_{CO,B}$ = $(3.4 \pm 0.2) \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$, and is surrounded by gray shading showing the 95% confidence interval for the fit. The dotted purple line was fit to star forming regions in the SMC Bar by Jameson et al. (2018), with $X_{CO} = 1 \times 10^{21}$ cm^{-2} (K km s^{-1})$^{-1}$. The dashed blue line was fit to clumps in N83C in the SMC Wing by Muraoka et al. (2017), with $X_{CO} = 7.5 \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$. The dot-dashed red line shows a typical $X_{CO} = 2 \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$ for the Milky Way (Bolatto et al. 2013).

We applied our derived $X_{CO,B}$ factor to the clumps without significant ^{13}CO detections and derived masses $M_{X_{CO,B}}$ ranging between 1.2 – 52 M$_\odot$, with a mean value of 10.7 M$_\odot$ and total mass of 860 M$_\odot$. We also applied this factor to the clumps traced by ^{13}CO. Going forward, we use the $M_{X_{CO,B}}$ mass estimates for all clumps for consistency. We find a total gas mass for all clumps traced by ^{12}CO in N90 of $\mathrm{M_{X_{CO,B}}} \sim 3310 \pm 250\ M_\odot$. The 26% of clumps that are traced by ^{13}CO contribute 74% of the total CO-bright clump mass.

Our mass estimate is slightly lower than the total clump mass estimate of 3800 $M_\odot$ in N90 made by Fukui et al. (2020) using only the 7m ^{12}CO ALMA observations. The difference between these results stems from variations in mass calculation and clump identification methods, not in spatial filtering, because >95% of the ^{12}CO flux in the 7m map is recovered in our 12+7m map). Fukui et al. (2020) identified 19 clumps with radii between 1.9–3 pc and used an $X_{CO,B}$= $7.5{\times}10^{20}$ cm^{-2} (K km/s)$^{-1}$ (Muraoka et al. 2017), as opposed to our sample of 110 clumps with radii between 0.2–0.8 pc and lower adopted $X_{CO,B}$.

2.1.4. *Other Clump Properties*

A full catalog of clump properties is presented in machine-readable format in Table A1. We determined the radius R of each clump by fitting an ellipse to its half-light contour and converting the FWHM values of the ellipse's major and minor axes to the standard deviation of a Gaussian profile. We then multiplied by 1.91 to calculate the "effective radius" as defined by Solomon et al. (1987). We report radii as the geometric mean of the major and minor axes, and these values range from 0.26 pc to 0.77 pc with a median of 0.40 pc.

We calculate average surface densities as $\Sigma = M_{X_{CO,B}}/(\pi R^2)$. The median Σ is 24 $M_\odot$ pc^{-2}, with a standard deviation of 40 $M_\odot$ pc^{-2}. To estimate the volume densities of clumps, we assume that the clumps follow a power-law density profile,

$$\rho(r) = \rho_c \left(\frac{r}{R_0}\right)^{-k}, . \qquad (6)$$

where R_0 is a normalizing radius that we set to be $R_0 =$ 0.1 pc for all clumps, and ρ_c is the density of the clump at R_0. Using the derivation in Appendix A.1 of O'Neill et al. (2022), this central density can be estimated as

$$\rho_c = \frac{(3-k)}{4\pi}\frac{M(r)}{R_0^k\ r^{3-k}} \qquad (7)$$

Power-law indices of k ∼1–2 have frequently been derived for clumps and cores (e.g., Caselli et al. 2002; Pirogov 2009; Chen et al. 2019, 2020; Lin et al. 2022) and we adopt $k = 1$ for the N90 clumps. Through solving Eqn. 7 with $r = R_{CO}$, densities at 0.1 pc range between 10^1–10^3 M$_\odot$ pc^{-3}, with an average of $\rho_c \simeq 190$ M$_\odot$ pc^{-3}.

We calculated velocity dispersions σ_v and peak CO velocities v_{LSRK} by fitting Gaussian distributions to intensity-weighted ^{12}CO velocity profiles. We do not correct σ_v for the expected contribution from thermal motion (∼0.08 km s^{-1} for CO at 20 K); this is discussed further in §4.1. Values of σ_v ranged from 0.23 km s^{-1} to 1.07 km s^{-1} with a median of 0.45 km s^{-1}. Peak ^{12}CO velocities ranged between v_{LSRK} =159–179 km s^{-1}, with a mean of 167 km s^{-1}.

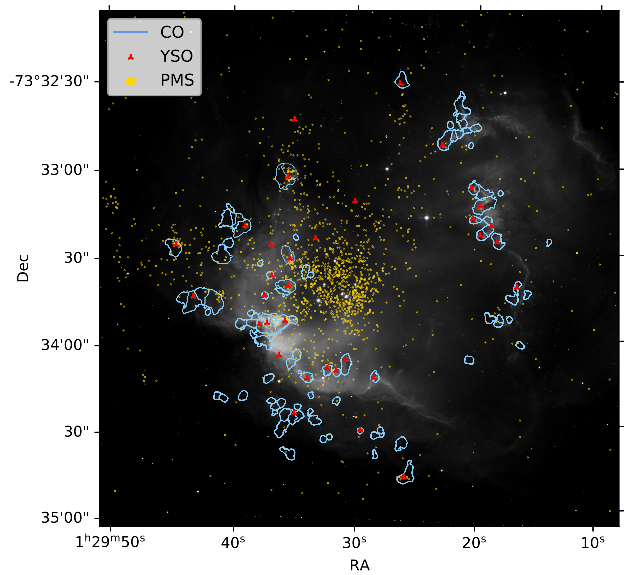

Figure 4. The HST ACS Hα image is shown with the locations of PMS stars (yellow points), YSO candidates (red triangles), and the 2D boundaries of the CO clumps identified by the `quickclump` algorithm (light blue contours).

2.2. *Archival Data: IR-identified YSO candidates*

Two types of young stellar objects have been analyzed in the N90 region: solar-mass PMS stars identified with high resolution HST optical and near-IR photometry, and YSO candidates with infrared excess emission attributed to circumstellar dust identified using Spitzer and Herschel Space Telescope (Herschel) data. We revisit these populations here in order to assess their relationship with the resolved molecular gas, but do not attempt to re-do the careful classification of previous authors.

2.2.1. *HST Data*

We used HST F555W (∼V) and F814W (∼I) observations of N90 reduced by Schmalzl et al. (2008) to study the distribution of solar-mass PMS stars in the region, as selected by Gouliermis et al. (2012) (herafter G12). The locations of PMS stars are shown in Figure 4. The data from HST GO program 10248 consist of 2156s and 2269s total integration time with ACS WFC in F555W and F814W, respectively. Photometry was performed using

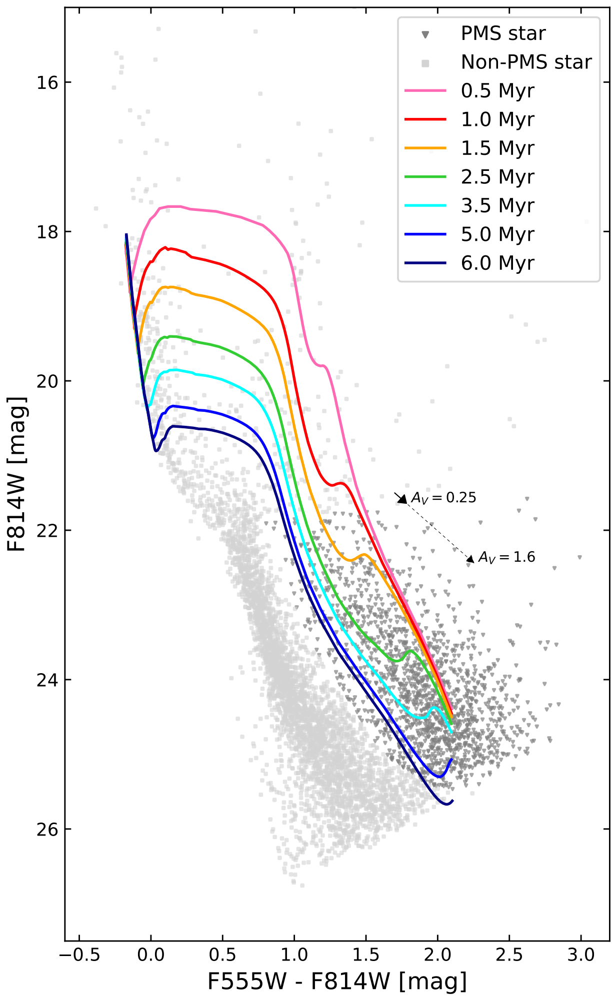

Figure 5. Locations of stars in F555W - F814W, F814W color-magnitude space. PMS stars are marked with dark gray triangles, and non-PMS stars are marked with light grey squares. PMS isochrone models for ages between 0.5 and 6 Myr are shown as colored curves. A reddening vector for $A_V = 0.25$ mag is shown with by the thick black solid arrow, and for $A_V = 1.6$ mag with the thin black dashed arrow.

DOLPHOT[3], reaching a depth of 26 mags in each filter, albeit at reduced completeness below 23 and 22.5 magnitudes in F555W and F814W, respectively (see Schmalzl et al. 2008, for details).

G12 separated faint PMS stars from lower main sequence stars by fitting pairs of Gaussians to the distribution of stars contained within bands perpendicular to the main sequence. The minimum between the Gaussian corresponding to the MS population and that corresponding to the redder PMS population is taken as the classification boundary. We select PMS stars as falling within or to the right of the region bounded by F814W $>$ 21.8 mag and F814W $<$ -1.9(F555W - F814W)^2 + 8.3(F555W - F814W) + 16.6. We generated isochrones using the Pisa PMS evolutionary models (Tognelli et al. 2011) and the IDL program TA-DA (Da Rio & Robberto 2012) to create synthetic photometry based on the Kurucz (1993) atmospheric models. The models were calculated for Z=0.003, Y=0.254, and mixing length parameter α = 1.2. We assumed A_V = 0.25 mag and E(B-V) = 0.08 mag (Carlson et al. 2007; Gouliermis et al. 2012), $R_V = 3.1$ (Schultz & Wiemer 1975; Gordon et al. 2003), and a distance modulus of $\mu = 18.91$ mag (corresponding to ∼60.6 kpc, Hilditch et al. 2005). We assigned ages to each PMS star based on the track it was closest to for ages between 0.5 and 6.5 Myr in 0.25 Myr steps.

The resulting isochrones for ages of 0.5, 1, 1.5, 2.5, 3.5, 5, and 6 Myr are shown in Figure 5 with the selected PMS stars. In §3.3.2, we estimate an average $A_V \simeq 3.1$ mag through the centers of CO clumps that contain PMS stars. For the ∼10% of PMS stars that appear contained within projected 2D clump, a typical embedded star in the center of a clump might then reasonably be affected by half of this value, $A_V \simeq 1.6$ mag. Reddening vectors for both A_V=0.25 mag and A_V=1.6 mag are plotted in Figure 5. They fall at steep angles to the isochrones and indicate that high levels of differential reddening could significantly skew age estimates. We discuss this possibility further in §3.3.2.

2.2.2. *Spitzer and Herschel Data*

We also re-analyzed Spitzer-identified intermediate- and high-mass YSO candidates in the region. Carlson et al. (2011, herafter C11) combined V, I, J, H, K, Spitzer IRAC 3.6-8.0 μm, and Spitzer MIPS 24 μm photometry, using the high-resolution optical data to remove background galaxies. C11 first fit the sources with stellar photospheres, removing sources consistent with stars, and then fit the remainder with Robitaille et al. (2006) YSO models to identify and classify all sources consistent with intermediate-mass YSOs. Starting with the C11 combined photometry catalog with galaxies removed, but including sources they classified as stars, we added aperture photometry from the Herschel HERITAGE survey (Meixner et al. 2013; Seale et al. 2014) at 100, 160, 250, and 350 μm using the Spectral and Photometric Imaging Receiver (SPIRE) and Photodetector Array Camera and Spectrometer (PACS).

[3] The ACS module of DOLPHOT is an adaptation of the photometry package HSTphot (Dolphin 2000). It can be downloaded from http://americano.dolphinsim.com/dolphot/.

We re-calculated aperture photometry of all Spitzer and 2MASS images in order to directly compare our aperture photometry code against the catalog photometry, and have a consistently calculated number for all bands in which an upper limit was required. Our script simply extracts the pixel sum in circular apertures at the source location, with radii of [1, 1, 3, 3, 3, 3, 9, 18, 8, 11, 18, 25, 37] arcseconds in filters [F555W, F814W, IRAC1, IRAC2, IRAC3, IRAC4, MIPS24, MIPS70, PACS100, PACS160, SPIRE250, SPIRE350, SPIRE500]. A background consisting of the median value in an annulus around each aperture was subtracted. The aperture photometry agrees within uncertainties with the previous C11 photometry for most sources, except those that suffer from confusion and crowding.

We visually assessed the spectral energy distribution (SED) and image cutouts in all filters using a script to assemble that information on a single page for each source, in order to evaluate which filters were contaminated by neighboring sources and/or diffuse emission; in the case of contamination, we used the aperture flux density as an upper limit in subsequent fitting. After this assessment, most of the sources had to be fit with upper limits in the new longer-wavelength bands because of the lower angular resolution of those data; however, these upper limits are still sufficient to exclude some models included in C11 that are very bright in the far-infrared.

We fit the sources with the updated Robitaille (2017) set of YSO models, which cover a wider range of parameter space more uniformly than the previous Robitaille et al. (2006) grid. However, the newer models do not have associated stellar masses as the 2006 models did, so we match each model's log L and log T to the nearest PARSEC PMS photosphere model (Bressan et al. 2012) to determine an $M_\star$. The new set of models are parameterized by envelope characteristic density ρ_0 and centrifugal radius R_C. Given that the circumstellar envelope has the particular density distribution of a rotating infalling toroid, one can uniquely calculate an "envelope accretion rate" $\dot{M}$ for a model's given ρ_0, R_C, and stellar mass. This parameter is largely a convenient way to parameterize the degree to which the source is embedded, with higher $\dot{M}/M_\star$ indicating a more embedded source as discussed in Robitaille et al. (2006). When we compare these properties to CO clumps in §3.3.1, the general properties of the YSO candidate population will be considered, but the precise fitted mass or envelope mass will not dramatically alter our conclusions.

For each source, we calculate χ^2 for each model, and use the probability-weighted mean value of each fit parameter and the full-width at half-maximum of the parameter's 1D marginalized probability density function (PDF) as the fitted parameter and its uncertainty. We examined all sources' PDFs as a function of $M_\star$ and envelope $\dot{M}$ - a minority of sources have mulitply-peaked PDFs but the adopted uncertainty range in all cases encompasses both peaks so is a reasonable measure of the data's ability to constrain the source properties.

Our re-analysis does not change the list of intermediate-mass YSO candidates relative to C11. The primary addition to C11's analysis is that the addition of longer-wavelength upper limits eliminates luminous, heavily embedded models, with high circumstellar dust columns. Addition of longer-wavelength photometry in a few cases suggests an infrared excess for some sources classified by C11 as bare stellar photospheres, but none of these additions are definitive, and higher resolution long-wavelength imaging will be required to conclusively measure any infrared excess.

We preserve C11's classification of "K" source as non-YSOs in our analysis: K049 (J012903.28-733413.2) has a tentative 100 μm detection of 35±15 mJy, but is located in filamentary diffuse emission. K194 (J012920.73-733327.1) has a marginal 100 μm measurement of 4.5±3 mJy. The most likely IR-excess candidate amongst C11's "K" sources is K456 (J012954.82-733231.5) with a marginal 100 μm measurement of 2.7±1.6 mJy, but also a 24 μm flux of 1.1±0.45 mJy in excess of a stellar photosphere. All three lie outside of the region mapped in CO, so their classification has no effect on our conclusions. C11 noted two sources that they called stars but with infrared excess emission in their analysis with longest wavelength of 24 μm. S235 and S213 are relatively brights star located within the central bubble; we confirm that both have 24 μm emission in excess of a photosphere, but neither are conclusively detected at longer wavelengths, as they would be if they had a massive circumstellar envelope, so we keep the C11 "star" classification. Neither is associated with CO emission.

The longer-wavelength data do, however, significantly change the stellar and envelope masses for the most massive sources - all of the most massive sources fit by C11 are fit with models that have bright FIR excess emission, and in all cases, we measured the FIR emission to be modest, excluding those massive models that are acceptable fits to C11's shorter wavelength range. The largest fit YSO mass is 8±1 $M_\odot$, whereas C11's best fits included one object consistent with 26 $M_\odot$ (we find 4±1 $M_\odot$) and five with masses between 10 and 12 $M_\odot$. In total, we find a higher mass estimate than C11 for only one YSO (Y227, J012937.37-733352.4, with C11 mass of 6.86 $M_\odot$ vs our estimate of 6.99 $M_\odot$); all other YSO candidates have lower masses. We derived a total YSO mass of $\sim 160 M_\odot$ in N90, which is significantly less than C11's previous estimate of a total YSO mass of ~ 300 $M_\odot$. The fit properties of all YSOs are reported in Table A2.

Our fits to the Robitaille models yield two constraints on the age of the intermediate-mass YSO candidates: the age of the PARSEC photospheric model matched with each Robitaille (2017) YSO model, and the fitted mass divided by fitted envelope accretion rate. The "photospheric age" is fairly well constrained, to $\sim 10^5$ years for the more massive YSOs >5 $M_\odot$, increasing to $\sim 10^6$yrs for the lower mass $M_\star \simeq 2$ $M_\odot$ YSOs. The fitted mass divided by accretion rate is consistent, agreeing to within the order-of-magnitude uncertainties on that ratio, but has the even larger caveat that the current fitted accretion rate is almost certainly not the accretion rate through the entire mass assembly of the YSO, and that the fitted accretion rate only has physical meaning insofar as the rotating infalling envelope density distribution used by Robitaille actually represents the accretion process - in general we are far more confident that this fitting process can constrain the total envelope dust mass than we are in interpreting that quantity as an accretion rate. Nevertheless, these two constraints suggest ongoing intermediate-mass star formation in NGC 602 over the last 1–2 Myr.

3. DYNAMICS AND EVOLUTION OF N90

3.1. *Morphology of N90*

N90 is characterized by a ring-shaped "rim" ~ 30 pc in diameter that frames a central cavity containing NGC 602 (Figure 1b). The rim is most clearly visible to the east and west, with the northeast section of the rim appearing to be more diffuse and the southeast section denser. Many background galaxies are clearly visible outside the edges of N90 due to the low surface density, generally transparent surrounding environment of the SMC Wing ($A_V \lesssim 10^{-2}$ on degree scales; Gordon et al. 2014).

N90's rim displays clear evidence of photodissociation and has many "pillar"-like features. Most of the molecular clumps in N90 are arranged along or immediately outside of the rim, and many of the pillars are closely associated with CO emission. Two representative regions of the rim are shown in Figure 6. CO emission closely traces the edges of the Hα emission, and most of the pillars point towards the center of N90; this suggests that they are the result of radiation from the cluster NGC 602 (Gritschneder et al. 2010).

Many of the YSO candidates in N90 (§3.3.1) are embedded within the pillars and clumps along the rim. In their study of ~ 200 giant molecular clouds (GMCs) in

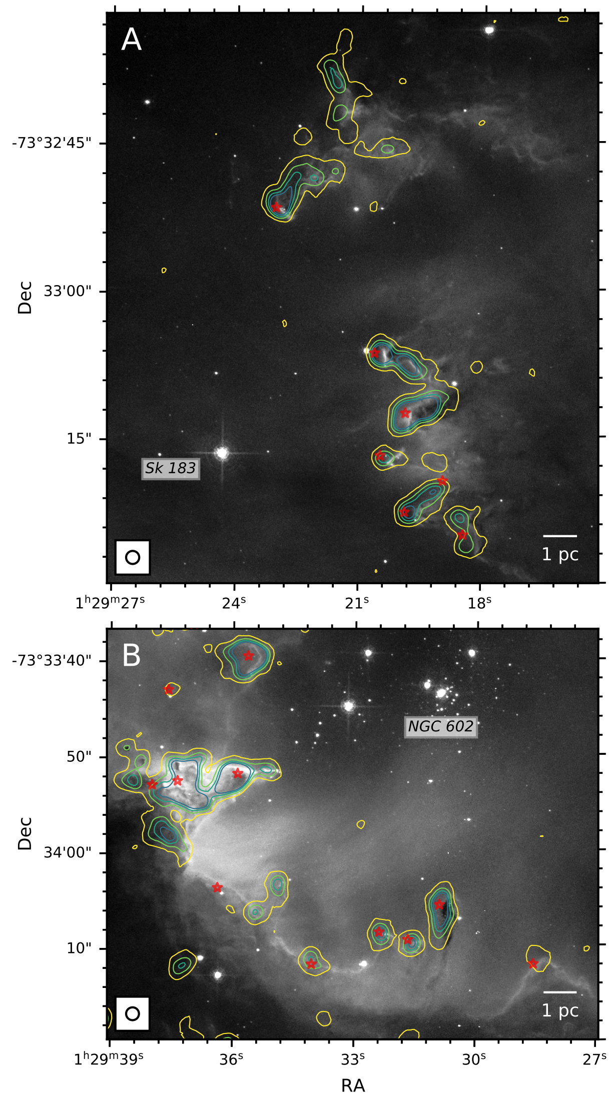

Figure 6. *Top:* Expanded image of Figure 7a's Region A, along the West rim near the massive star Sk 183 (labeled). Contours of integrated ^{12}CO intensity are shown, with levels of 0.25, 0.5, 0.75, and 1 Jy beam^{-1} km s^{-1} drawn in yellow, green, teal, and blue, respectively. The synthesized beam size for the CO observations is shown in the lower left corner. The locations of YSO candidates are shown by red stars. HST ACS Hα image is shown in gray. *Bottom:* As the left panel, but for Figure 7a's Region B, on the East rim near NGC 602 (labeled).

the LMC, Ochsendorf et al. (2016) found that massive star formation is most likely to occur at the edges of clouds nearest to young stellar clusters, implying these clusters stimulate clump and star formation. The YSO candidates in N90 shown in Figure 6 are also mostly concentrated on the edges of the rim/their host pillars that are closest to NGC 602 or Sk 183; this is especially apparent in Region A.

We assess whether the formation of the YSOs in N90 may have been triggered by NGC 602. Here we define triggered star formation as a process in which the interaction between the earliest generation of stars and their environment directly causes the formation of subsequent generations. This stands in contrast to sequential star formation scenarios where the formation of distinct stellar generations are largely unrelated.

3.2. *Spatial Variation in Clump Properties*

We assessed visually which CO clumps are on the rim (Figure 7a) to investigate if any trends between clump properties and 2D position relative to the rim exist. Through this assessment, we estimate that 64 of the clumps are on the rim, and 46 are outside of it. 76 of the clumps are located in the Eastern half of the region nearer to the NGC 602 cluster, and 34 are in the Western half nearer to the massive star Sk 183. The spatial variation of the column densities, excitation temperatures, and radial velocities of the CO clumps are shown in Figure 7b–d.

Clumps in the NE, both in the rim and non-rim sections, appear to have higher surface densities (Figure 7b) and excitation temperatures (Figure 7c, which under the LTE assumptions is indirectly equivalent to ^{12}CO brightness temperature) than clumps along the SE rim and non-rim sections. The NE clumps may thus be more directly affected by the central OB association NGC 602, as gas associated with the Hα rim could present less of a barrier to mechanical and radiative heating from the association than for clumps shielded by (and possibly located within or behind) dense gas to the south. We used 2D radial distances between the center of NGC 602 (α =01 29 32.133, δ = -73 33 38.13) and all clumps on the Eastern half of N90 as a proxy for the true 3D distances between their positions. We calculated Spearman's rank correlation coefficient ρ_s between the distance from NGC 602 and clump properties; there are weak-to-moderate negative correlations between distance and σ_v ($\rho_s = -0.40$, $p < 0.001$) and distance and v_{LSRK} ($\rho_s = -0.58$, $p < 0.001$), and only weak correlations between distance and R, $M_{X_{CO,B}}$, or T_{ex} (all $|\rho_s| < 0.3$).

Along the western rim, the warmest and densest clumps are concentrated in the NW section near the massive O3 star Sk 183 (α =01 29 24.6, δ = -73 33 16.43). Although the clumps in the northwest are relatively removed from NGC 602 itself, their proximity to Sk 183 could be responsible for their marginally higher excitation temperatures relative to the cooler and more isolated clumps in the SW rim. Evans et al. (2012) found that Sk 183 is likely responsible for the majority of hydrogen-ionizing photons in N90 and Ramachandran et al. (2019) confirmed its outsized effect, with Sk 183 is contributing up to 30% of all ionizing photon flux in their sample of ∼300 OB stars scattered throughout the entirety of SGS 1. We stacked the CO spectra within 5 pc of Sk 183 and identified no significant (>3σ) emission in either ^{12}CO or ^{13}CO. This suggests that it has successfully dissociated its immediate surroundings, and could even be responsible for the nearby northern "gap" in the Hα rim. We also explored correlations between Western clump properties and their distance from Sk 183, and found moderate negative correlations between distance and: σ_v ($\rho_s = -0.51$, $p = 0.002$); $M_{X_{CO,B}}$ ($\rho_s = -0.48$, $p = 0.004$); and T_{ex} ($\rho_s = -0.49$, $p = 0.003$).

Figure 8 compares the distributions of T_{ex}, σ_v, R, and Σ between all rim and non-rim clumps. The probability density functions of these properties are represented using Gaussian kernel density estimations (KDEs), with bandwidths chosen using "Scott's Rule" as implemented in the Scipy python package (Virtanen et al. 2020). By performing Kolmogorov-Smirnov (K-S) tests comparing the properties of rim vs non-rim clumps, we found no significant differences between the distributions of R (with a p-value of $p = 0.43$ at a significance level requirement of $\alpha = 0.05$), T_{ex} ($p = 0.26$), or Σ ($p = 0.682$) for the two groups. There was a significant difference in the distributions of σ_v between rim and non-rim clumps ($p = 0.002$), with a one-sided Mann-Whitney U test revealing that clumps along the rim on average have larger σ_v ($p = 0.003$).

The overall lack of any strong correlations and relatively few significant differences between groups suggests that stellar feedback (e.g., radiation pressure or stellar winds) has not had a significant differential impact on the physical properties of the molecular clumps, or at least that any systematic trend cannot be extracted from the random cloud property variations with the modest range of central cluster distance available in these data (∼2–23 pc). Indebetouw et al. (2013) found a similar lack of significant trends from 10 to 25 pc from the cluster R136 in the LMC region 30 Doradus.

3.2.1. *Clues for Formation of NGC 602*

We stacked the CO spectra in the half of the ALMA coverage to the south of N90 ($\delta \lesssim -73°35'$) where no clumps were identified by the `quickclump` algorithm. We found no significant ($> 3\sigma$) CO emission in this area. Assuming $T_{ex} = 10$ K, the mean N(^{12}CO) in this region is less than 4.5×10^{16} cm^{-2}. Thus, we find no

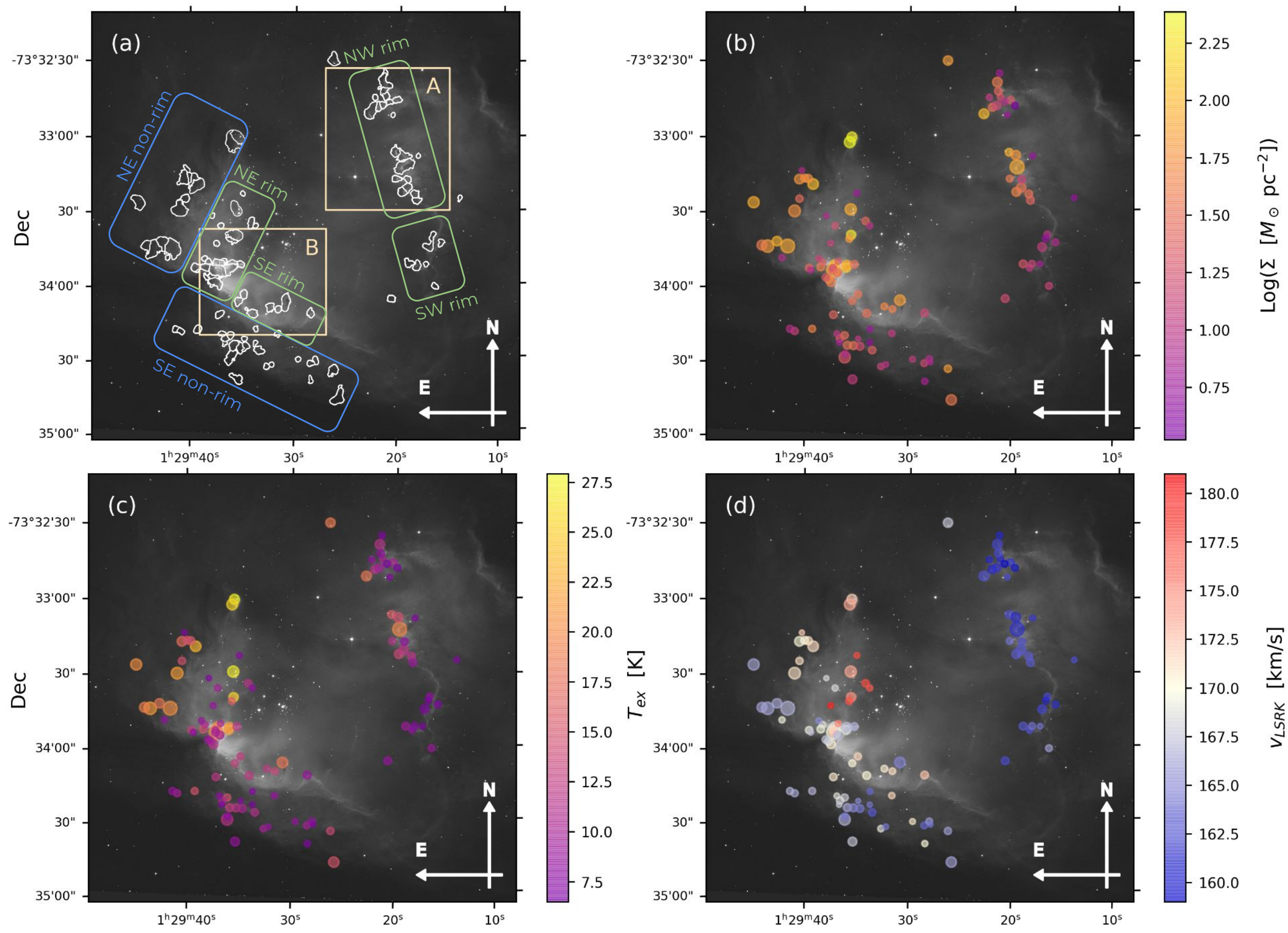

Figure 7. Spatial variation of CO clump characteristics, with HST Hα image shown in gray. *(a:)* The regions we define as part of the rim are shown in green, and the corresponding non-rim regions are marked in blue. The 2D clump boundaries identified by the `quickclump` algorithm are shown in white. Details of the two orange rectangles A and B are shown in Figure 6. *(b:)* Average surface density Σ of clumps, with point sizes proportional to clump area. *(c:)* Maximum clump excitation temperature T_{ex}. *(d:)* Peak ^{12}CO velocity of clumps, with red/blue color scale centered at the mean radial velocity of HI components surrounding N90, v_{LSRK} ~170 km s^{-1}, as derived from Fukui et al. (2020) and Nigra et al. (2008).

strong evidence for CO to the south of N90 and conclude that CO-traceable molecular gas is largely localized in the site of massive star formation to the north, which is adjacent to the southern rim of SGS 1. This correspondence suggests that turbulence and compression resulting from the southward expansion of HI component(s) within SGS 1 triggered both the formation of dense molecular gas and then stars in N90.

Fukui et al. (2020) proposed that the collision of two clouds hundreds of parsecs in diameter was responsible for the formation of NGC 602. In this scenario, the larger of their two HI clouds would have moved south from the northeast before colliding with a smaller, less-massive cloud moving north from the southwest. They suggested that SGS 1 is simply a cavity inside the more massive cloud created by this collision. In the alternative colliding shells formation scenario proposed by Nigra et al. (2008), one shell moved south from the northwest and the other moved south from the northeast, with NGC 602 forming at their intersection.

We examine if evidence of these collisions in SGS 1 7–8 Myr ago could still be preserved within the current clumps. At solar metallicity, photoelectric heating rates in molecular gas are usually sufficiently higher than energy injection rates from the decay of turbulence that a relationship between dissipation of turbulence and kinetic temperature is not expected. Mechanical heating can be traced by kinetic temperature when the heating is particularly high, as in central starbursts (e.g., Kazandjian et al. 2016; Mangum et al. 2019). At low metallicity, however, both cooling rates and photoelectic heating rates are sufficiently reduced that even more modest levels of turbulent dissipation might be able to affect the gas kinetic temperature.

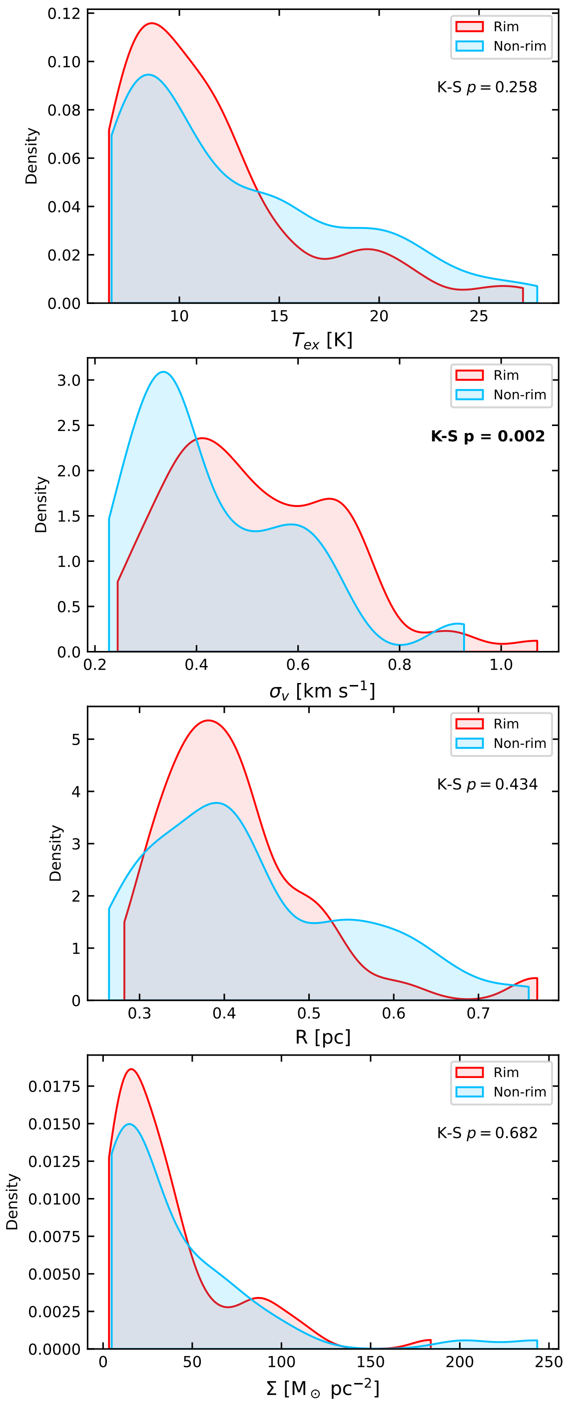

Figure 8. Kernel density estimations (KDEs) comparing physical properties of clumps on the N90 rim (red) and not on the N90 rim (blue), with both groups assigned as in Figure 7a. p-values resulting from two-sided K-S tests comparing the rim vs non-rim distributions are shown in each panel, with any $p < 0.05$ set in bold to indicate that there is a statistically significant difference between the distributions. Top: Excitation temperature T_{ex}. Second from top: Velocity dispersion σ_v. Third from top: Radius R. Bottom: Surface density Σ.

Turbulence is expected to dissipate on the order of a crossing time (Elmegreen & Scalo 2004). The HI clouds considered in Fukui et al. (2020) have diameters d$\sim$ 600 pc, velocity dispersions σ_v $\sim$10 km s^{-1}, masses $M \sim 8\times 10^6$ M$_\odot$, and so a mean volume density <n$_H$ > of 3 H cm^{-3}. The shell most closely associated with N90 by Nigra et al. (2008) has d$\sim$200 pc, σ_v $\sim$6 km s^{-1}, M $\sim$3$\times$10^5 M$_\odot$, and also <n$_H$ > = 3 H cm^{-3}. Estimating a turbulent dissipation rate as $0.5{<}\rho{>}\sigma_v^2$ / $\tau_{\rm crossing}$ with $\tau_{\rm crossing}$ $\sim$d/σ_v yields 1$\times$10^{-27} and 3$\times$10^{-27} erg s^{-1} cm^{-3} for the Nigra et al. (2008) and Fukui et al. (2021) clouds, respectively. The cooling rate for the neutral ISM dominated by C$^+$ and O^0 line emission at 1/5 $Z_\odot$ is $\lesssim$10^{-26} erg s^{-1} cm^{-3} in clouds at these low densities (Wolfire et al. 1995). These estimates are sufficiently uncertain as to preclude a definitive statement, but the cooling rate exceeding the estimated heating rate suggests that any excess kinetic energy from the cloud collision has probably been radiated away, and feedback from the central cluster and Sk 183 is still more likely the dominant energetic driver. However it is not impossible that at low metallicity, the signature of more turbulent HI gas might still be present in current properties of the molecular gas formed from that HI gas.

We compare the radial velocities (RVs) of the clumps to nearby stars in N90 and to SGS 1's proposed HI components. All velocities are reported in the LSRK frame. Clump RVs range from 161 to 179 km s^{-1}, with a mean of 168 km s^{-1}. Evans et al. (2012) derived an RV of 151 $\pm$ 1 km s^{-1} for Sk 183 and Ramachandran et al. (2019) derived a mean RV for $\sim$17 OB stars in N90 of 158$\pm$4 km s^{-1}. The cause of this offset from the RVs of the CO clumps is unknown. The Ramachandran et al. (2019) measurements have typical uncertainties of $\pm$10 km s^{-1}, making it difficult to tell how significant this shift is. In contrast, the clump RVs are in close agreement with the RVs of nearby HI structures along the "ring" of SGS 1: Nigra et al. (2008) derived an RV of 168$\pm$5 km s^{-1} for a proposed progenitor HI gas clump "curled" around the south of N90. Similarly, the larger of the two Fukui et al. (2020) clouds had a range of velocities from 163 – 183 km s^{-1}, with a peak at $\sim$173 km s^{-1}.

In Fig. 7d we show the spatial variation of the clump RVs relative to the mean of these two measurements, 170 km s^{-1}. Clumps along the eastern rim near NGC 602 appear consistent with this value (median +0.2 km s^{-1} from 170 km s^{-1}), while clumps in the northeast and southeast non-rim regions (median -3.1 km s^{-1}) and the western half (median -7.7 km s^{-1}) mostly appear blueshifted. Gvaramadze et al. (2021) derived a central RV of 177 $\pm$ 6 km s^{-1} for background Hα emission near

SNR SXP 1062, which falls just outside of the SGS 1 HI "ring". Since these measurements of HI and Hα RVs along in SGS 1 are largely consistent with the radial velocities of the clumps in N90 on the edge of the ring, a close connection between SGS 1 and the formation of N90 appears likely.

3.3. *Clump Association with YSOs and PMS Stars*

Many previous studies have analyzed the characteristics of candidate PMS stars and YSOs in N90 (e.g., Carlson et al. 2007, 2011; Cignoni et al. 2009; Schmalzl et al. 2008; Gouliermis et al. 2007, 2012; De Marchi et al. 2013). To better inform our analysis of the observed CO clumps we replicate and extend some aspects of this extensive past analysis. Hereafter, we refer to the candidate several solar-mass sources with MIR excess emission selected from Spitzer and Herschel as "YSOs", and the generally lower-mass sources selected from their location in an HST color-magnitude diagram as "PMS stars".

3.3.1. *YSOs*

Of the thirty-three YSOs identified in §2.2 that are inside the ALMA-observed area, twenty-eight are located inside the projected 2D boundaries of a CO clump (∼85%). Correspondingly, 27 of the 110 clumps (∼25%) appear to contain at least one YSO. By performing one-sided K-S tests, we found significant differences between the properties of clumps that contain vs. do not contain YSOs: clumps containing YSOs have higher R ($p < 0.001$), σ_v ($p < 0.001$), $M_{X_{CO,B}}$ ($p < 0.001$), and excitation temperatures T_{ex} ($p < 0.001$), and lower virial parameters $\alpha_{\rm vir}$ ($p = 0.005$, see §4.3) .

These findings are consistent with comparisons of YSOs and clumps in other low-Z regions. In the star-forming complex N159 in the LMC, for example, Nayak et al. (2018) found that CO clumps containing YSOs were more massive than clumps that did not contain YSOs, and that massive YSOs were typically associated with the most massive clumps. Similarly, in the LMC region N55 Naslim et al. (2018) observed that molecular cores containing YSOs possessed larger linewidths and masses than those that did not.

The majority of clumps (83 out of 110, or ∼75%) are not associated with any YSOs, but may have been in the past. We calculated the distance between each clump and its nearest YSO and found moderate negative correlations between distance to the nearest YSO and clump mass ($\rho_s = -0.48$, $p < 0.001$), and distance and σ_v ($\rho_s = -0.47$, $p < 0.001$). We estimated the distance that a clump could have plausibly travelled since formation of this generation of YSOs began 1 – 2 Myr ago (C11). The mean σ_v of all clumps is 0.48 km s^{-1} with a standard deviation of 0.17 km s^{-1}. Using this distribution of σ_v as a proxy for the speed at which clumps may be moving relative to each other, we calculated the distance traveled over a timescale of 1.5 Myr at a relative speed of 0.65 km s^{-1} to derive a potential distance traveled estimate of ∼1 pc. Twenty of the 83 clumps that do not contain YSOs have a YSO within this distance. From Mann-Whitney U-tests, these clumps have significantly larger σ_v than clumps outside of this distance ($p = 0.002$, mean σ_v of 0.52 km s^{-1} vs. 0.41 km s^{-1}).

Only 5 of the 33 YSOs (15%) are not embedded within the projected 2D boundaries of a clump, so traditional hypothesis testing to compare the properties of YSOs within clumps to YSOs outside of clumps is not appropriate. A simple comparison of the median masses of the two groups suggests that YSOs inside clumps are more massive than YSOs outside of clumps (median $M = 3.4\ M_\odot$ vs $2.3\ M_\odot$) and have higher envelope $\dot{M}$ accretion rates (median $\dot{M} = 2.1 \times 10^{-6}\ M_\odot$ yr^{-1} vs $\dot{M} = 0.76 \times 10^{-6}\ M_\odot$ yr^{-1}). This suggests that the YSOs not embedded within clumps are currently less actively accreting material, although as discussed in §2.2 this fitted value is a limited indicator of the actual historical accretion rate of the objects.

We also checked if CO emission that was not assigned to a clump by the `quickclump` algorithm was present around seemingly isolated YSOs by stacking CO spectra around each YSO within an area equal to three synthesized beams. We found no robust evidence for strong CO emission around these YSOs (Y198, Y271, Y283, Y290, Y358) with all YSOs having one or fewer 3σ detections in ^{12}CO and ^{13}CO out of the 105 channels in the stacked spectra.

Seale et al. (2012) found that a large number of massive YSOs in four LMC GMCs were not associated with any molecular clumps detected with HCO$^+$, which they suggested to be the result of the disruption of clumps on $\lesssim$ 1 Myr timescales. In the 30 Doradus region of the LMC, Nayak et al. (2016) found that massive YSOs were more likely to be associated with CO clumps than their low-mass counterparts and concluded that the less-massive YSOs not associated with clumps were likely more evolved than the embedded YSOs, as they would have had sufficient time to dissipate their natal molecular clumps through UV radiation. We draw similar conclusions that unassociated YSOs may be more evolved than embedded YSOs, and that feedback from YSOs may affecting be the molecular gas on the scale of individual clumps ($\lesssim$1 pc). We note, though, that small sample sizes involved weaken the power of these conclu-

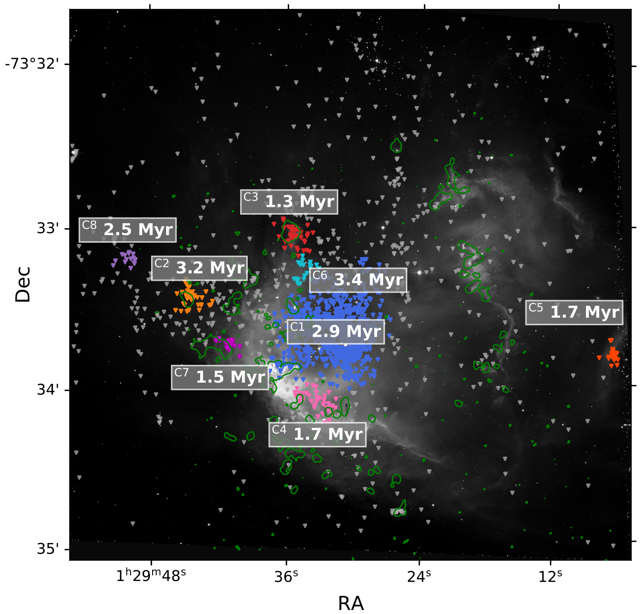

Figure 9. Locations of the PMS star clusters identified by the `DBSCAN` algorithm. PMS stars are marked with triangles. The clusters are denoted by color, and are labeled with their Table 1 ID numbers and the mean ages of their members. Non-clustered PMS stars are marked with grey triangles. Hα image is in grey and contours of integrated ^{12}CO(2–1) emission at 0.25 Jy beam^{-1} (km s^{-1})$^{-1}$ is shown in green.

Table 1. Properties of PMS Star Clusters

ID	RA (°)	Dec (°)	$N_{\rm PMS}$	Mean Age (Myr)	$\sigma_{\rm Age}$ (Myr)
C1	22.382	-73.561	779	2.9	1.6
C2	22.436	-73.557	49	3.3	1.4
C3	22.398	-73.551	44	1.9	1.9
C4	22.389	-73.568	41	1.8	1.5
C5	22.278	-73.563	27	2.0	1.9
C6	22.394	-73.554	22	3.2	1.8
C7	22.422	-73.562	21	2.0	1.9
C8	22.459	-73.553	15	3.5	1.3

Note—The reported position of each cluster is the mean position of all members of that cluster. $N_{\rm PMS}$ is the number of stars per cluster.

sions, as do the large uncertainties on the masses and accretion rates of each individual YSO.

3.3.2. *PMS Stars*

G12 analyzed the spatial distribution of PMS stars in N90 and identified 14 sub-clusters. We supplement this work using the `DBSCAN` algorithm (Density-Based Spatial Clustering of Applications with Noise, Ester et al. 1996), a non-parametric clustering method that defines clusters as regions of high density separated by regions of low density. It requires the assignment of two parameters: the minimum number of points to form a cluster, MinPts, and the maximum distance over which two points can be considered neighbors, ϵ. We set the minimum number of cluster members as MinPts = 15 stars. Following Rahmah & Sitanggang (2016), we found the optimal value of ϵ by creating a nearest-neighbors graph of the distance between the k = MinPts nearest-neighbors of points in ascending order and identifying the approximate point of maximum curvature. This yielded $\epsilon = 4.4$".

We identified 8 distinct clusters of PMS stars, the locations of which are shown in Figure 9 and properties summarized in Table 1. Like G12, we find overdensities of PMS stars in the central NGC 602 association and along the rim of the H II region. The largest cluster, Cluster 1, is centered around NGC 602 and consists of 779 members (∼50% of the total PMS sample). Oskinova et al. (2013) analyzed *Chandra* and *XMM-Newton* observations of N90 and found evidence for extended X-ray emission around this central cluster, and also in a feature on the rim directly to the north of NGC 602 where the `DBSCAN` algorithm identifies Cluster #3. They attributed these features to the effects of many low-mass PMS stars and YSOs unresolved in the X-ray data, which is consistent with the high concentration of resolved PMS stars in these locations that we find here.

Through the isochrone fitting described in §2.2 and shown in Figure 5, we find that the majority of the PMS stars in N90 are consistent with ages less than 5 Myr, with a mean age of 3 Myr. We note that the uncertainties on these estimates are large due to the simple method of isochrone matching that we adopted. Cluster 1 has a mean age of ∼3 Myr, which is consistent with the age estimated for PMS stars in NGC 602 by C11 and G12. Clusters 3, 4, and 7 are scattered along the H II rim and possess slightly lower mean ages of ∼2 Myr. This would seemingly support a star formation history in which the formation of these clusters was triggered by the formation of Cluster 1.

However, in all clusters the dispersions in member age are large (typically $\gtrsim$1.5 Myr). Additionally, there is a notable spread across V–I in the entire sample of PMS stars, with a handful of extremely young, red stars appearing to exist (m_{555} - $m_{814} \gtrsim 2.5$ mag). Visual inspection reveals that many of the reddest sources are located in extremely crowded areas (for example, Cluster 5 at the edge of the HST coverage). These sources could be genuine, or could simply be a result of confusion or indicate the presence of misclassified asymptotic giant branch stars or unresolved background galaxies.

Many of the reddest sources are also associated with bright CO emission along the rim; this suggests that significant differential reddening could be resulting from the brightest clumps. Of the 1569 PMS stars, 130 fall within the projected boundaries of a CO clump. These sources appear significantly younger ($p < .001$ from a one-sided Mann-Whitney U test, median age of 2 vs 3.5 Myr), i.e. redder, than PMS stars not inside CO clumps. Conversely, 40 of the 110 CO clumps contain at least one PMS star. Clumps that contain PMS stars have significantly higher masses ($p < 0.001$), surface densities ($p < 0.001$), and σ_v ($p = 0.003$) than clumps that do not contain PMS stars, and significantly lower virial parameters $\alpha_{\rm vir}$ ($p < 0.001$, see §4.3).

Clusters 3 and 7 along the Hα rim (with apparent mean ages of 1.9 and 2.0 Myr, respectively) are both associated with strong CO emission. De Marchi et al. (2013) observed that the youngest and reddest PMS stars in N90 were located in dense areas of the rim, and suggested that if the extinction towards these stars was significantly higher than the rest of the sample ($A_V \simeq 2.25$ mag vs. $A_V = 0.25$ mag) the ages of the outlier stars would be comparable to the ages of the PMS stars in the central NGC 602. In the N83 region of the SMC Wing, Lee et al. (2015) derived a relationship between CO intensity and extinction, $I_{CO}/A_V = 1.5$ K km s^{-1} (mag)$^{-1}$. The mean I_{CO} within clumps in N90 that contain PMS stars is 4.7 K km s^{-1}, which corresponds to $A_V \sim 3.1$ mag; a centrally embedded star might experience half of this value, $A_V \sim 1.6$ mag. Thus, reddening resulting from the CO clumps would be sufficient to cause Clusters 3 and 7 to appear much younger than they truly are. If this is the case, we would then find no strong evidence for a triggered star formation scenario.

Additionally, we find no correlation between age and radial distance from the center of NGC 602 in the full sample of PMS stars ($\rho_s = 0.08$, $p < 0.001$). The crossing time through the ∼30 pc HII region is 3 Myr, which is equal to the mean ages of the entire PMS sample and most contained clusters. Since many of the PMS stars are located on the edges of the HII region and the spatial distribution of their ages is roughly uniform, this suggests a sequential star formation scenario is more likely than a triggered event (i.e., the formation of local subclusters of PMS or OB stars would not have had time to directly influence the formation of other subclusters).

4. MOLECULAR GAS & STAR FORMATION IN LOW METALLCITY ENVIRONMENTS

4.1. *Size-Linewidth-Surface Density Relationships*

The relationship between size, linewidth, and surface density in molecular clouds has been extensively studied as a proxy for their dynamical states. Larson (1981) identified correlations between global cloud properties including size R, linewidth σ_v, and surface density Σ. The first of these correlations follows

$$\sigma_v = C\left(\frac{R}{1\ \mathrm{pc}}\right)^{\Gamma}\ \mathrm{km\ s^{-1}}, \tag{8}$$

with $\Gamma = 0.5$ and $C \simeq 0.72$ km s^{-1} pc$^{-0.5}$ (e.g., Solomon et al. 1987; Heyer et al. 2009, hereafter SRBY and H09, respectively). A virialized spherical clump described by a power-law density distribution $\rho \propto r^{-k}$ should additionally follow the relationship

$$\frac{\sigma_v^2}{R} = \frac{(3-k)}{3(5-2k)}\pi G\Sigma. \tag{9}$$

Figure 10 compares σ_v, R, and Σ for the N90 clumps. We also show cores, clumps, and GMCs from:

- Two samples of Galactic GMCs observed by H09, with the first defined from SRBY's ^{12}CO GMCs and the second from the half-power contours of the central GMCs cores observed in ^{13}CO (median radii R of 9.7 pc and 1.6 pc, respectively)
- clouds in the Galactic center studied by Oka et al. (2001) in ^{12}CO (median R of 8.7 pc)
- clouds in the Ophiuchus molecular cloud in ^{13}CO from Ridge et al. (2006) (median R of 0.07 pc)
- cores identified using the dendrogram algorithim within the Galactic molecular cloud Perseus A observed in ^{13}CO by Shetty et al. (2012) (median R of 0.07 pc)
- clumps in the Magellanic Bridge studied by Kalari et al. (2020) in CO (1–0) (mean R of 1.1 pc)
- clumps in the Magellanic Bridge studied in ^{12}CO (2–1) by Valdivia-Mena et al. (2020) (mean R of 1.1 pc)
- GMCs identified by (Ochsendorf et al. 2017) in ^{12}CO in ∼150 LMC star-forming regions (median R of 27 pc)
- Clumps identified by Wong et al. (2019) in ^{12}CO in the LMC regions 30 Doradus, A439, GMC 104, GMC 1, PCC, and N59C (median R of 1 pc)

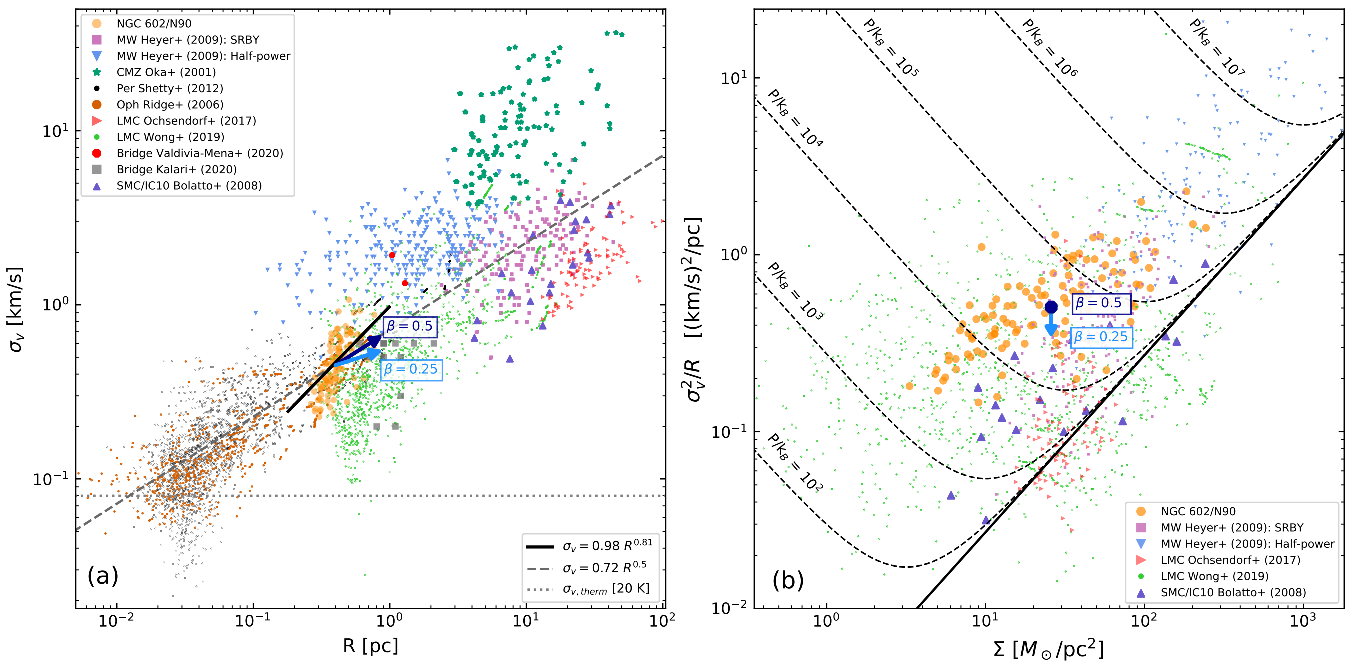

Figure 10. Size-linewidth-surface density plots, adapted from O'Neill et al. (2022) with the addition of N90 clumps. *(a:)* Velocity dispersion σ_v compared to radius R of clumps from studies described in §4.1. The arrows show the corrections needed to account for CO-dark gas on observable CO clump properties for a typical clump with [R_{CO}=0.40 pc, $M(R_{CO}) = 13\ M_\odot$, $\sigma_v(R_{CO}) = 0.45$ km s^{-1}]. The arrows move from the properties of the observed CO clump to the inferred full state of the clump including CO-dark gas. The dark blue arrow follows a velocity profile $\sigma_v \propto R^{0.5}$ [$\beta = 0.5$] and the light blue arrow follows a velocity profile $\sigma_v \propto R^{0.25}$ [$\beta = 0.25$]. Both arrows show the effect for a density profile $\rho \propto R^{-1}$ [$k = 1$] and $f_{DG} \sim 0.8$. The grey dashed line follows the relationship $\sigma_v = 0.72R^{0.5}$ derived by SRBY and the grey dotted line is the expected σ_v at T=20 K from thermal motion. The black solid line was fit to the N90 clumps and follows $\sigma_v = 0.98\ R^{0.81}$. *(b:)* Size-linewidth parameter σ_v^2/R compared to surface density Σ, with correction arrows as in (a). The black line corresponds to virial equilibrium (Equation 9 with $k = 0$), and the black curves correspond to virial equilibrium under varying degrees of external pressure (with units of P/k in K cm^{-3}, Equation 11 with $\Pi = 0.6$).

- GMCs in the SMC and dwarf galaxy IC 10 observed in CO (2–1) and (1–0) by (Bolatto et al. 2008) (median R of 15 pc)

Figure 10a compares σ_v to R. The expected contribution to linewidth by thermal motion at T=20 K, $\sigma_{v,\text{th}} \sim 0.08$ km s^{-1}, is also shown. We fit a relationship for the N90 clumps of $\Gamma = 0.81 \pm 0.10$ and $C = 0.98 \pm 0.09$ km s^{-1}. Similarly steep values of Γ compared to SRBY have been derived in other low-Z, low-density regions throughout the SMC, LMC, and other local dwarf galaxies (with $\Gamma \sim 0.55 - 0.85$ and $C \sim 0.2 - 0.6$ km s^{-1}, Bolatto et al. 2008; Hughes et al. 2010, 2013; Wong et al. 2019; Finn et al. 2022). Our fit C is significantly higher than derived in these studies, suggesting that the clumps in N90 have larger linewidths at a given size than structures in those other dwarf galaxy regions.

Kepley et al. (2016) studied 8 GMCs in the low-Z (Z $\sim$ Z$_{SMC}$) starburst dwarf galaxy II Zw 40 and derived similarly high C; they attributed this to high linewidths and surface densities stemming from a merger between dwarf galaxies that triggered the starburst, rather than high external pressure supporting the GMCs against collapse. Imara & Faesi (2019) identified ~120 GMCs in the moderately low-Z ($Z \sim 0.7Z_\odot$) starburst dwarf galaxy He 2-10, and found both higher velocity dispersions, surface densities, and C than in comparable Milky Way clouds. They also fit a very steep size-linewidth slope of $\Gamma = 1.3$, which they suggested could be the result of tidal interactions or energy and momentum injected from nearby superstar clusters.

Figure 10b compares surface density Σ to the size-linewidth parameter [σ_v^2/R]. Virialized clumps are expected to follow Equation 9, but the N90 clumps have higher σ_v^2/R's for a given Σ than would be expected based on from this trend. A likely reason that clumps in a region would deviate from expected trends in size-linewidth space, and have higher kinetic energy at a given size scale, is the OB stars in the center of NGC 602 injecting kinetic energy into the surrounding gas.

If there is significant inter-cloud thermal pressure acting on the molecular clumps, as observed by Oka et al. (2001) in the Galactic center, and the clumps are assumed to be in virial equilibrium with that pressure, turbulence can be treated as a pressure term (Field et al. 2011). This increase in internal turbulent pressure would then be reflected as higher linewidths than expected, as observed in the relationship between σ_v^2/R and Σ in clumps in N90. The external pressure P_e for a clump to remain bound (under the assumption that they are virialized, which is discussed in §4.3) can be found using (Elmegreen 1989),

$$P_e = \frac{3\Pi M \sigma_v^2}{4\pi R^3}. \tag{10}$$

The median external pressure required for clumps to remain bound is $P_e/k_B \sim (2.4 \pm 1.3) \times 10^4$ K cm^{-3}. Clumps on the rim require on average 1.5× larger external pressures to remain bound than clumps that are not on the rim, but a K-S test reveals no overall .significant difference between the distributions of the two groups ($p = 0.14$).

The surface density Σ and size-linewidth parameter σ_v^2/R for clumps in pressure-bounded virial equilibrium can be related to P_e as (Field et al. 2011),

$$\frac{\sigma^2}{R} = \frac{1}{3}\left(\pi \Pi G \Sigma + \frac{4P_e}{\Sigma}\right). \tag{11}$$

We adopt $\Pi = 0.6$ for a uniform sphere (Field et al. 2011). These relationships are represented in Figure 10b by the black V-shaped curves for P_e/k_B between 10^2–10^7 K cm^{-3}. The N90 clumps do not appear consistent with being supported by any one value of P_e, but the majority of the clumps require between $P_e = 10^3$ K cm^{-3} and 10^6 K cm^{-3}.

The clumps identified by Wong et al. (2019) in six GMCs in the LMC with ^{12}CO and ^{13}CO emission were observed at comparable angular resolution to our data (having a synthetic beam size of 3.5" [0.8 pc], vs 1.3" [0.4 pc] in the data we present here). They found higher σ_v at a given R in regions with higher infrared surface brightness (a plausible measure of star formation rate and stellar feedback). The mean 8μm brightness of ∼0.5 MJy/sr in N90 corresponds to the lower end of surface brightnesses in the Wong et al. (2019) study of LMC regions. Thus N90 is consistent with the Wong et al. (2019) findings in this parameter space, although care should be exercised since both measures may be affected by reduced metallicity. The 8μm diffuse emission at a given radiation intensity is lower in the SMC than LMC, and the CO measurements may also be affected as discussed below.

4.2. *"CO-dark" Gas: Effects of Low Metallicity on CO Diagnostics in N90*

At low metallicities, reduced dust-to-gas ratios and typically stronger radiation fields increase the efficiency of CO destruction (Madden et al. 2006; Gordon et al. 2011; Madden et al. 2020) and decrease the effectiveness of CO as a tracer for H_2. The $C^+/C^0/CO$ transition retreats further into the center of clumps, which causes the fraction of "CO-dark" H_2 gas mass not traced by CO to increase (Wolfire et al. 2010; Glover & Mac Low 2011). It is possible that the discrepancies in N90 clump properties compared to expected size-linewidth-surface density trends are the result of this increased proportion of CO-dark gas. We apply corrections derived by O'Neill et al. (2022, hereafter O22) to account for the expected contribution of CO-dark gas on observed clump properties in N90.

As defined by Wolfire et al. (2010), the fraction of H_2 gas mass that is CO-dark in a clump, f_{DG}, can be expressed as

$$f_{DG} = 1 - \frac{M(R_{CO})}{M(R_{H_2})}, \tag{12}$$

where $M(r)$ is the mass contained within a given radius r, R_{CO} is the radius of the CO-traceable material, and R_{H_2} is the radius at which half of the hydrogen in the clump's diffuse envelope is molecular and half is atomic. Under typical Galactic conditions, f_{DG} is found to be $\gtrsim$0.3 (e.g., Grenier et al. 2005; Abdo et al. 2010; Velusamy et al. 2010; Lee et al. 2012; Langer et al. 2014; Xu et al. 2016).

To apply the corrections derived by O22, we assume that clumps follow a power-law density profile with $k = 1$ (Equation 6) and internal velocity dispersion profile, $\sigma_v(r)$, of

$$\sigma_v(r) = \left(\frac{r}{R_{CO}}\right)^{\beta} \sigma_v(R_{CO}). \tag{13}$$

$\beta \sim$0.2–0.3 is common in observations and simulations of young cores and clumps (Caselli & Myers 1995; Tatematsu et al. 2004; Lee et al. 2015; Lin et al. 2022), while a steeper β ∼0.5 in alignment with the global size-linewidth relationship of SRBY with $\Gamma = 0.5$ (Equation 8) has been found to hold on larger scales (e.g., Heyer & Brunt 2004; Dobbs 2015). Based on these findings, we adopt $\beta = 0.25$ for the bulk of our analysis, but also present results for $\beta = 0.5$.

We estimate an appropriate f_{DG} for N90 based on measurements of CO-dark gas content in other low-Z, low-density environments. In the LMC, Chevance et al. (2020) derived $f_{DG} \gtrsim$0.75 in the star-forming region 30 Doradus, and in the HII region N11, Lebouteiller et al. (2019) found that the majority (40–100%) of molecular gas was CO-dark. Throughout the low-Z outskirts of the Milky Way Pineda et al. (2013) derived f_{DG} ∼0.8, while in the nearby low-metallicity dwarf galaxy NGC 4214 (Z ∼0.3–0.4 $Z_\odot$, Hermelo et al. 2013), Fahrion et al. (2017) derived $f_{DG} = 0.79$. Pineda et al. (2017) found that 77% of the total molecular gas in their sample of 18 line-of-sight pointings across the SMC was CO-dark H_2, and in the SMC regions N66, N25+N26, and N88 (located in the northern Bar, southwest Bar, and Wing, respectively). Requena-Torres et al. (2016) derived a typical fractional abundance of CO-dark gas to be 80–95%. In four star-forming regions in the nearby southwest Bar of the SMC with cloud density is comparable to N90 and the Wing ($\bar{N} \lesssim 2 \times 10^{21}$ cm^{-2}), Jameson et al. (2018) derived an average $f_{DG} \simeq 0.8$. Based on these results, we adopt $f_{DG} \sim 0.8 \pm 0.1$ for an estimated dark-gas fraction in N90 for the remainder of this work.

O22 derived corrections for clump properties under a power-law density profile, including

$$\begin{aligned} R_{H_2} &= [1 - f_{DG}]^{1/(k-3)}\ R_{CO}, \\ \sigma_v(R_{H_2}) &= [1 - f_{DG}]^{\beta/(k-3)}\ \sigma_v(R_{CO}), \\ M(R_{H_2}) &= [1 - f_{DG}]^{-1}\ M(R_{CO}), \\ \Sigma_{H_2} &= [1 - f_{DG}]^{(1-k)/(k-3)}\ \Sigma_{CO}. \end{aligned} \tag{14}$$

After applying these corrections to the $M_{X_{CO,B}}$-derived clump masses, which estimate the mass within the CO-traced regions, we find a total molecular gas mass in N90 of $M_{DG} \sim 16,600 \pm 2,400\ M_\odot$. The errors on this estimate are derived from the combined uncertainties of f_{DG} and $M_{X_{CO,B}}$.

In Figure 10a and Figure 10b, we demonstrate the effects of these corrections on a typical observed clump with properties [R_{CO}=0.40 pc, $M(R_{CO}) = 13\ M_\odot$, $\sigma_v(R_{CO}) = 0.45$ km s^{-1}, $\Sigma_{CO} = 26\ M_\odot$ pc^{-2}]. The arrows show the change in placement in size-linewidth-surface density space from the example clump's observed CO properties to its inferred "true" characteristics when including CO-dark gas for $\beta = 0.25$ and $\beta = 0.5$. Assuming $f_{DG} = 0.8$ and $k = 1$, the preferred $\beta = 0.25$ yields corrected properties of [$R_{H_2} = 0.9$ pc, $M(R_{H_2}) = 65\ M_\odot$, $\sigma_{v,H_2} = 0.55$ km s^{-1}, $\Sigma_{H_2} = 26\ M_\odot$ pc^{-2}]. In this case, the corrections bring clump properties closer to agreement with expected Σ–$[\sigma_v^2/R]$ trends. Since this change is only by a relatively small amount, though, and O22 demonstrated that the effects of these corrections vary significantly depending on the density and velocity profiles assumed, we conclude that is unlikely but not impossible that enhanced photodissociation in this low-Z environment is responsible for observed departures from size-linewidth-surface density trends in N90.

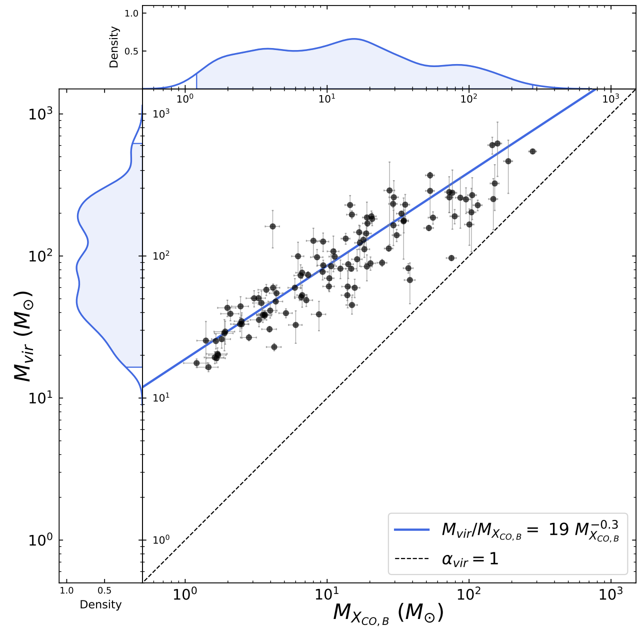

Figure 11. Clump virial mass $M_{\rm vir}$ is compared to clump $X_{CO,B}$-derived mass $M_{X_{CO,B}}$. The black dashed line shows a 1:1 relationship where $M_{\rm vir} = M_{X_{CO,B}}$, i.e., where the virial parameter $\alpha_{\rm vir} = 1$ and stability could be expected, and the blue line shows a best-fit relationship $M_{\rm vir}/M_{X_{CO,B}} = 19\ M_{X_{CO,B}}^{-0.34}$. The two mass distributions are represented individually by Gaussian kernel density estimations placed horizontally on the top axis of the scatter plot ($M_{X_{CO,B}}$) and vertically on its left axis ($M_{\rm vir}$).

4.3. *Stability of Clumps*

The evolution and stability of clumps can be also studied through assessing their virial masses, which for a clump with a power-law density profile (Equation 6) can be found as (Solomon et al. 1987; MacLaren et al. 1988)

$$M_{\rm vir}(r) = \frac{3(5-2k)}{(3-k)}\ \frac{r\ \sigma_v^2(r)}{G}, \tag{15}$$

where k is the index of the power-law. As in §4.2 we assume $k = 1$. Virial masses for clumps in N90 ranged from 16 $M_\odot$ to 620 $M_\odot$, with a median mass of 85 $M_\odot$ and total virial mass of all clumps of 14,045 ± 515 $M_\odot$.

We compare the relationship between individual $M_{\rm vir}$ and $M_{X_{CO,B}}$ in Figure 11. A relationship $M_{\rm vir}/M \propto M^{-\eta}$ has been observed to hold for dense clumps in many regions, with estimates of η typically ranging from $\eta \sim 0.3$ to 0.4 in Galactic clouds (e.g., Yonekura et al. 1997; Ikeda et al. 2009). A least-squares fit for clumps in N90 yields $M_{\rm vir}/M_{X_{CO,B}} = (18.8 \pm 1.3)\ M_{LTE}^{(-0.34\ \pm\ 0.04)}$. KDEs of the distributions of $M_{\rm vir}$ and $M_{X_{CO,B}}$ are also shown, with $M_{\rm vir}$ being centered at higher masses than $M_{X_{CO,B}}$.

We then calculated the virial parameter $\alpha_{\rm vir}$ defined by Bertoldi & McKee (1992) for all clumps as

$$\alpha_{\rm vir} = \frac{2\Omega_K}{|\Omega_G|} = \frac{M_{\rm vir}}{M_{X_{CO,B}}}, \tag{16}$$

where Ω_G is total gravitational potential energy and Ω_K is the total kinetic energy. $\alpha_{\rm vir} \sim 1$ indicates that a clump is gravitationally stable, and $\alpha_{\rm vir} \gg 1$ indicates that a clump is sub-critical and will likely expand unless confined by external pressure. There is a wide variation in $\alpha_{\rm vir}$ from clump to clump, with $\alpha_{\rm vir}$ ranging from 1.3 to 39. The median value is $\alpha_{\rm vir} = 7.95$, with lower and upper quartiles of 4.4 and 11.5, respectively.

This is significantly higher than many recent measurements of $\alpha_{\rm vir} \lesssim 2$ in molecular clouds (see Kauffmann et al. 2013 for a review), and may be an overestimate due to the effects of CO-dark gas, which we discuss below. In the 30 Doradus region of the LMC, Wong et al. (2017) found that virial clump masses were typically an order of magnitude larger than CO-derived masses, which is similar to what we observe here. Schruba et al. (2017) studied ∼150 small CO clumps (mean radius $R \simeq 2.3$ pc) in five star-forming regions in the $Z = 1/5\ Z_\odot$ dwarf galaxy NGC 6822 and found large values of α_{vir} (from ∼1 to $\gtrsim$10). Similarly high values of α_{vir} have also been found in clumps in the Galactic Central Molecular Zone (Myers et al. 2022) and in cores in the Pipe Nebula (Lada et al. 2008).

The increased degree of photodissociation in low-Z environments (see §4.2) compromises the fundamental assumption that CO emission accurately traces clump mass, and by extension interpretations of $\alpha_{\rm vir}$ values. We expect the enhanced amount of CO-dark gas in this region to cause $M_{X_{CO,B}}$ to be a significant underestimate of the total amount of molecular gas (§4.2), so correcting for the "true" values of $\alpha_{\rm vir}$ including CO-dark gas could bring the clumps more in line with expected trends. O22 derived the CO-dark-corrected virial mass and parameter as

$$\begin{gathered} M_{\rm vir}(R_{H_2}) = [1 - f_{DG}]^{(2\beta+1)/(k-3)}\ M_{\rm vir}(R_{CO}), \\ \alpha_{vir,H_2} = [1 - f_{DG}]^{(2\beta+k-2)/(k-3)}\ \alpha_{vir,CO}. \end{gathered} \tag{17}$$

We correct clump virial masses and parameters using Equations 14 and 17, with a CO-dark gas mass fraction $f_{DG} \sim 0.8$ assumed in §4.2 and velocity dispersion profiles following $\beta = 0.25$ and $\beta = 0.5$. For [$k = 1$, $\beta = 0.5$], values of α_{vir} remain unchanged from the original estimate. For the preferred [$k = 1, \beta = 0.25$], the new median value is $\alpha_{\rm vir,H_2} = 5.3$. Although this is a reduction, this is still much higher than would be expected for a virialized clump.

We note that the calculations of virial masses suffer from large uncertainties stemming from the determination of radii and assumption of a spherical clump, and that the many uncertainties in the calculations of column densities (especially those stemming from assuming equal excitation temperatures for ^{12}CO and ^{13}CO and constant abundance ratios between ^{12}CO, ^{13}CO, and H_2), $X_{CO,B}$ masses, and the dark-gas mass fraction are also significant. Still, the only marginal decrease in $\alpha_{\rm vir}$ implies that the clumps are either confined by high levels of external pressure or are not evolving near a virialized state. In any case, it is unlikely that CO-dark gas is responsible for the observed high $\alpha_{\rm vir}$.

The high values of $\alpha_{\rm vir}$ suggest that the clumps have higher internal kinetic energies than clumps in other regions and galaxies. If the clumps are long-lived, this imbalance must be addressed by external pressure or magnetic fields. The data are generally consistent with N90 being an energetic region and contributing a higher inter-clump-medium pressure than in more quiescent regions. If the central OB cluster is responsible for this energetic state, we would expect to see some variation in virial parameter and other clump properties with proximity to the cluster.

However, we found no significant difference in the distributions of $\alpha_{\rm vir}$ between the rim vs. non-rim clumps ($p = 0.095$ from a K-S test). As shown in Figure 12, there is no correlation between $\alpha_{\rm vir}$ and distance from NGC 602 for clumps on the Eastern rim ($\rho_s = 0.04$, $p = 0.7$), while on the Western rim, there is only a moderate trend for increased $\alpha_{\rm vir}$ as a function of distance from Sk 183 ($\rho_s = 0.46$, $p = 0.006$). This is similar to the recent results of Wong et al. (2022) and Finn et al. (2022) in the LMC, who both found no significant correlations between clump $\alpha_{\rm vir}$ and distance from the super star cluster R136. The overall absence of any strong trends with position in N90 suggests that the the entire region is energetic for a different reason, e.g., the aftereffects of the collision between the HI clouds/supershells.

4.4. *Efficiency of Low-Z Star Formation*

The supergiant shell SGS 1 falls on the boundary between the HI-dominated outskirts and molecule-rich center of the SMC, and the star formation rate (SFR) per unit area and mass has been observed to drop dramatically along such transitions (Krumholz 2013). The low amount of CO emission to the south of N90 compared to the concentrated CO emission to its north may reflect this transition in the Wing. When gas mass is assessed using CO emission, low-Z environments have been observed to have higher star formation rates (SFRs), and by extension higher apparent star formation efficiencies

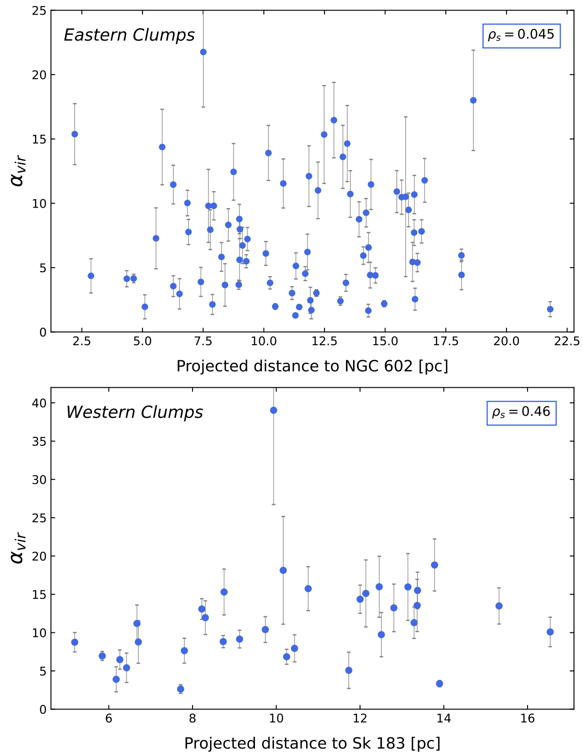

Figure 12. *Top:* Virial parameter $\alpha_{\rm vir}$ is compared to projected distance to NGC 602 for clumps in the Eastern half of N90. The value of Spearman's ρ_s rank correlation coefficient for the two variables is shown in the top right. *Bottom:* As the top, but for projected distance to the massive star Sk 183 for clumps on the Western half of N90.

(SFEs) than higher-Z regions (e.g., Galametz et al. 2009; Schruba et al. 2012; Schruba et al. 2017), although this may simply be the result of CO being a poor tracer of H_2 at low Z. Here we explore metrics of SFE in N90.

We initially examined the SFE in N90 at the scale of individual clumps ($\lesssim$1 pc). If one calculates a by-clump efficiency as the ratio of the mass of any contained YSOs to CO-dark corrected clump mass, efficiency appears to decrease with clump mass, with values ranging from 0.5% to 16.5%; however, this ratio cannot account for the unknown difference between the original clump mass that created a given YSO and the observable extant gas mass. Since the gas within a few parsecs of a YSO is quickly disrupted by the star formation process and the dynamic range of the Spitzer-identified YSO masses is small, it is difficult to conclude much from such a clump-scale efficiency.

Although the offset in time between current and original gas mass is an unavoidable limitation of SFE measurements, comparing the total YSO and CO masses is more useful in that it likely averages over a significant portion of the clump evolutionary sequence. It can then

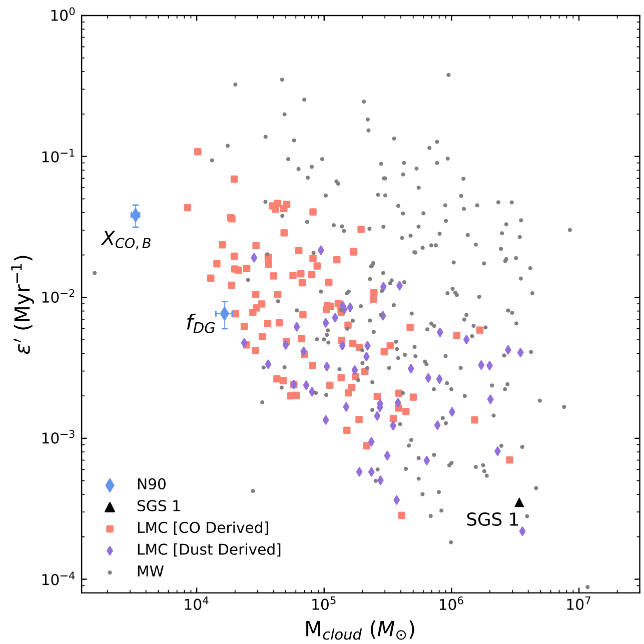

Figure 13. Star formation efficiency ϵ' = SFR/$M_{\rm cloud}$ is compared to cloud mass $M_{\rm cloud}$. The ϵ' value we derived from the CO-bright gas mass in N90 using an $X_{CO,B}$ conversion factor ($X_{CO,B} = 3.4 \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$ is shown along with the ϵ' derived through correcting for CO-dark gas mass ($f_{DG} \simeq 0.8$, §4.2) by the labeled blue points. The ϵ' we derived for SGS 1 is marked with a black diamond. Results for a sample of LMC GMCs analyzed by Ochsendorf et al. (2017) are shown: pink squares had masses derived through α_{CO} conversion, and purple diamonds had masses derived from Jameson et al. (2016)'s dust-based molecular hydrogen map of the LMC. A sample of Galactic clouds studied by Lee et al. (2016) is also shown by gray circles.

provide a more meaningful efficiency estimate than analysis of individual clumps. With this in mind, we studied SFE and the SFR throughout the entirety of N90 through analyzing the overall population of candidate YSOs.

To account for incompleteness in the YSO sample, we assumed a two-part stellar IMF of the form dN/d log M $\propto M^{-\alpha}$, where $\alpha = 0.3$ below 0.5 $M_\odot$ and $\alpha = 1.3$ above 0.5 $M_\odot$ (Kroupa 2001). We found no high mass YSOs ($\gtrsim$ 10 $M_\odot$), which suggests that the IMF for the current generation of star formation in N90 is not fully populated. We scaled this mass function to the peak of our observed YSO mass distribution at 3 $M_\odot$ and integrated over 0.08 to 50 $M_\odot$ to derive a total YSO mass in N90 of $M_* \simeq 1250 \pm 160$ $M_\odot$, which corresponds to an estimated total number of YSOs of N(YSO) $\simeq$ 260.

In comparison, C11 derived a total YSO mass of 2250 $M_\odot$ as inferred through the same method. This difference stems from our globally lower YSO mass estimates. Our list of YSO candidates is identical to C11's but (all but one of the) revised individual YSO masses are lower than the original estimates, so the peak of our stellar mass function is shifted to lower masses than C11's. Since C11 and this work both infer the total YSO mass through scaling a Kroupa (2001) IMF to the peak of the YSO mass function, it is unsurprising that we derive a significantly lower total mass estimate.

Using only the $X_{\rm CO,B}$-derived CO-bright gas mass, this total YSO mass yields a recent SFE ($\epsilon = M_*/M_{\rm gas}$) in N90 of $\epsilon_{X_{CO,B}} = 38 \pm 7$ %. After correcting for CO-dark gas, we estimate an overall $\epsilon_{DG} = (1 - f_{DG})\epsilon_{X_{CO,B}}$ (O22) of $\epsilon_{DG} = 8 \pm 3$ %. This estimate is significantly lower than the 20% formation efficiency estimate that Fukui et al. (2020) derived for N90, which is unsurprising due to our addition of CO-dark gas mass to our calculation.

We calculated the SFR in N90 as $\rm SFR_{YSO}$ = N(YSOs) $\times \bar{M}/t_*$ (Ochsendorf et al. 2017), where $\bar{M}$ is the mean mass of the fully populated IMF (here 0.5 $M_\odot$) and t_* is the typical YSO age (here 1 Myr). We derived a $\rm SFR_{YSO} = 130 \pm 30$ $M_\odot$ $\rm Myr^{-1}$. From studying the optical PMS population in N90, Cignoni et al. (2009) found that the SFR has been increasing over the last 10 Myr, with a SFR of 150 $M_\odot$ $\rm Myr^{-1}$ between 5 and 2.5 Myr ago and reaching a peak of $300 - 700$ $M_\odot$ $\rm Myr^{-1}$ in the last 2.5 Myr. C11 derived a SFR of 2200 $M_\odot$ $\rm Myr^{-1}$ over the last 1 Myr through analyzing the YSO MF; we attribute our reduced estimate to our globally lower fitted YSO masses discussed earlier in this subsection.

We then derived the SFE using the notation of Kennicutt & Evans (2012) where ϵ' = SFR / $\rm M_{gas}$. Using the $X_{CO,B}$ CO-bright mass estimate, we estimate that $\epsilon'_{X_{CO,B}} = 0.04 \pm 0.01$ $\rm Myr^{-1}$, while the CO-dark corrected mass estimate yields $\epsilon'_{DG} = 0.01 \pm 0.005$ $\rm Myr^{-1}$. If the SFRs derived by Cignoni et al. (2009) or C11 were used ϵ' would increase by a small amount, but would not change the conclusions we draw below.

In Figure 13, these results are compared to values derived for ∼150 star forming regions in the LMC through analysis of YSO MFs by Ochsendorf et al. (2017) and for a sample of ∼190 Galactic clouds analyzed by Lee et al. (2016). Ochsendorf et al. (2017) and Lee et al. (2016) both found that ϵ' increased with decreasing cloud masses. For comparison, we also derive an estimated ϵ' for SGS 1. Rubele et al. (2018) derived a SFR of $\sim 1.19 \times 10^{-3}$ $M_\odot$ $\rm yr^{-1}$ in the last 8 Myr for a 21' × 21.5' region centered in SGS 1. We assume a typical N(HI) in this area of 2 $\times 10^{21}$ $\rm cm^{-2}$ (Welty et al. 2012) and a mean atomic mass of $\bar{m} = 1.5 m_H$ (Nigra et al. 2008). This yields an estimated gas mass within the Rubele et al. (2018) region of 3.4×10^6 $M_\odot$ and an $\epsilon' = 3.5 \times 10^{-4}$ $\rm Myr^{-1}$, which is significantly lower than the values derived in N90 and consistent with decreasing ϵ' with increasing gas mass.

Although on the low end of values observed for similar LMC and Galactic clouds, the SFE we derive for the relatively low-mass N90 is consistent with this trend. We therefore conclude that star formation in N90 and this region of the Wing of the SMC is not significantly more efficient than in higher-Z environments.

5. DISCUSSION AND CONCLUSIONS

5.1. *Evolutionary History of N90*

We review the scenarios for the formation of N90 presented in previous studies in Table 2, and add the results of this work. A combination of stimulated and stochastic star formation in this region of the SMC Wing has been ongoing for at least 100 Myr, with a notable extended star formation event between 25–40 Myr ago (Ramachandran et al. 2019; Fulmer et al. 2020). Between 7–8 Myr ago a collision occurred between ∼ 500 pc components of HI within the supergiant shell now identified as SGS 1 (Nigra et al. 2008; Fukui et al. 2020). Turbulence and compression stemming from this collision triggered the formation of NGC 602 3–5 Myr ago (Carlson et al. 2011; Gouliermis et al. 2012) and subsequent creation of the HII region N90. The parsec-scale CO clumps to the north of N90 may retain signatures of the HI collision in the form of inflated excitation temperatures and column densities, but determining whether a collision between shells (Nigra et al. 2008) or clouds (Fukui et al. 2020) is more likely to be responsible is not yet possible.

Intermediate-mass YSOs have been forming along the HII rim over the last 1–2 Myr (Carlson et al. 2011), but it is unclear if this was triggered by the formation of the central cluster. There is very little variation in the ages of PMS stars with distance from NGC 602, and although some isolated clusters of young ($\lesssim$2 Myr) PMS stars appear to exist along the rim, they coincide with CO emission sufficiently strong to cause age underestimates by $\gtrsim$1 Myr. There is some evidence for YSOs disrupting their natal clumps on the parsec-scale, but on the scale of the entire region, there are very few significant correlations between clump properties and distance from NGC 602 or Sk 183. We conclude that a sequential star formation process, in which the creation of the central NGC 602 cluster did not directly cause the formation of the YSOs or PMS stars along the rim, is more likely to be present than a triggered scenario.

5.2. *Conclusions*

Table 2. History of N90

Time	Event and Derivation Method	Reference
	Older Populations Form in SGS 1	
$\gtrsim 100$ Myr	SF begins in SGS 1, from SED and HRD fitting	Ramachandran et al. (2019)
$\sim$50 Myr	Subclusters NGC 602 B and NGC 602 B2 form, from optical CMD fitting	De Marchi et al. (2013)
25 – 40 Myr	Extended SF event in SGS 1, from optical/near-UV CMD fitting	Fulmer et al. (2020)
	Formation event for NGC 602	
8 Myr	Collision between clouds, from separation in HI velocity components	Fukui et al. (2020)
7 Myr	Collision between shells, from expansion velocities of HI shells	Nigra et al. (2008)
	SF begins in central cluster NGC 602	
$\lesssim 5$ Myr	Central PMS stars form, from optical CMD fitting	Gouliermis et al. (2012)
2 – 4 Myr	Central OB and PMS stars form, from optical CMD fitting	Carlson et al. (2011)
	Subsequent SF in N90	
2.5 Myr	Maximum SF rate reached, from optical PMS population	Cignoni et al. (2009)
$\lesssim 2.5$ Myr	PMS stars form in sub-clusters along rim, from optical CMD fitting	Gouliermis et al. (2012)
2 Myr	Median apparent age of PMS stars in CO clumps, from optical CMD fitting	This work
1 – 2 Myr	Ongoing intermediate-mass SF, from SED fitting of YSOs	This work
$< 1 – 2$ Myr	YSOs form, from SED fitting of evolutionary phase	Carlson et al. (2011)

We present results from ALMA observations of molecular gas in the low-metallicity star-forming region NGC 602/N90. The main conclusions of this analysis are as follows:

1. CO emission in N90 is confined to 110 sub-parsec-scale clumps arranged around the region's rim. Only 26% of clumps are traced by both ^{12}CO and ^{13}CO, with the remaining 74% of clumps only being traced by ^{12}CO with no strong corresponding ^{13}CO emission (§2.1.2). We derive a CO-to-H_2 conversion factor of $X_{\rm CO,B} = (3.4 \pm 0.2) \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$ $\simeq$ 1.7 $X_{\rm CO,MW}$ from the clumps that do possess strong emission in both ^{12}CO and ^{13}CO. Applying this factor to all clumps yields a total CO-traced mass of $3,310 \pm 250$ $M_\odot$ (§2.1.3).

2. We estimate a total molecular gas mass in N90 of $16,600 \pm 2,400$ $M_\odot$ through CO-dark gas mass correction (§4.2).

3. Clumps in N90 do not agree with expected trends in size-linewidth-surface density space, and have larger velocity dispersions and lower surface densities than predicted by relationships derived from Galactic clouds (§4.1). Additionally, CO-derived clump masses are significantly lower than virial masses, yielding high virial parameters (typical $\alpha_{\rm vir}$ = 4–11) and implying that clumps are either dispersing or confined by high levels of external pressure (§4.3). We use models of clumps with CO-dark gas to demonstrate that it is unlikely that CO-dark gas is responsible either of these effects.

4. We refit *Spitzer* YSO candidates identified by Carlson et al. (2011) and find by including new mid-to-far IR photometry that nearly all objects are less massive than previously estimated. Analysis of the present day accretion rate of the YSO candidate reveals that intermediate-mass star formation has likely been occurring throughout N90 in the last 1–2 Myr (§3.3.1). 85% of YSO candidates within the field observed by ALMA appear to be embedded within CO clumps. We derive a recent ($\lesssim$1 Myr) SFR of 130 ± 30 $M_\odot$ Myr^{-1}, with a total YSO mass of 1250 ± 160 $M_\odot$ and CO-dark gas corrected SFE of $\epsilon \simeq 8 \pm 3\%$ (§4.4).

5. We find no strong evidence that NGC 602 has directly triggered star formation along the rim of N90. Spatial position relative to NGC 602 and the rim are poor predictor of clump properties (§3.2), as is association with YSOs or PMS stars (§3.3), and there is no correlation between the age of PMS stars and radial distance from NGC 602. Although some clusters of PMS stars along the rim appear marginally younger than PMS stars surrounding NGC 602 ($\sim$2 Myr vs 3 Myr), they are coincident with strong CO emission and thus the high extinction in these regions could cause their ages to appear younger than they truly are (§3.3.2).

Our analysis of the now-resolved sub-parsec-scale clumps in N90 has revealed the sequential star formation history of the region, and its evolution relative to the SMC Wing. This more complete census of the total molecular gas mass in the region allows for an improved estimate of star formation efficiency on both by-clump and region scales. After correction for CO-dark molecular gas content, we find that star formation in N90 is not more efficient than star formation in similarly massive solar-metallicity, higher-density environments.

We consider N90 in the context of star formation in general in metal-poor environments. Despite the low-metallicity and low-density environment of the SMC Wing, the properties of molecular clumps and SFE in N90 do not appear to dramatically differ from their Galactic counterparts. If we extend this conclusion from the SMC to other nearby, small galaxies, it is likely that although star formation is initially sporadic in such environments, once regions have developed for a sufficient period of time their behavior does not depart significantly from the process of star formation in higher-metallicity, higher-density regions.

We thank Jay Gallagher for helpful discussions. T.J.O. and R.I. were supported during this work by NSF award 2009624. The material is partially based upon work supported by NASA under award No. 80GSFC21M0002 (M.S.). M. S. acknowledges partial support from the NASA ADAP grant No. 80NSSC22K0168. M.M. was supported by NSF award 2054178. K.S. was supported by NSF award 2108081. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1842490 (M.K.F.). This paper makes use of the following ALMA data: ADS/JAO.ALMA#2016.1.00360.S, ADS/JAO.ALMA#2012.1.00683.S, ADS/JAO.ALMA #2015.1.01013.S, and ADS/JAO.ALMA#2016.1.00193.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

Software: APLpy (Robitaille & Bressert 2012; Robitaille 2019); astrodendro (Robitaille et al. 2019); Astropy (Astropy Collaboration et al. 2013, 2018); CASA (McMullin et al. 2007); DOLPHOT (Dolphin 2016); Matplotlib (Hunter 2007); Numpy (Harris et al. 2020); OpenCV (Bradski 2000); Pandas (McKinney 2010; pandas development team 2020); quickclump (Sidorin 2017); Seaborn (Waskom 2021); scikit-learn (Pedregosa et al. 2011); Scipy (Virtanen et al. 2020); spectral_cube (Ginsburg et al. 2019); statsmodels (Seabold & Perktold 2010); TA-DA (Da Rio & Robberto 2012)

Facilities: ALMA, HST, Spitzer, Herschel

APPENDIX

A. CLUMP AND YSO PROPERTIES

We present tables of the properties of CO clumps (Table A1) and YSO candidates (Table A2) in N90.

Table A1. Clump Properties

ID	RA	Dec	$^{12}CO_{pk}$	$^{13}CO_{pk}$	W_{12}	R	$v_{LSRK,12}$	$\sigma_{v,12}$	M_{LTE}	$M_{X_{CO,B}}$	Σ	ρ_c	M_{vir}	α_{vir}	Assoc. YSO	Group
	(°)	(°)	(K)	(K)	(K km s^{-1})	(pc)	(km s^{-1})	(km s^{-1})	($M_\odot$)	($M_\odot$)	($M_\odot$ pc^{-2})	($M_\odot$ pc^{-3})	($M_\odot$)			
1	22.3982	-73.5501	22.55	6.26	9355.61	0.55	171.88	0.90	245.1	188.9	201.7	1008.6	464.7	2.5		
2	22.3986	-73.5581	21.85	3.79	5146.64	0.62	173.74	0.56	102.6	103.9	86.8	434.2	203.1	2.0	Y270	r
3	22.3990	-73.5506	20.20	7.11	13901.61	0.61	173.55	0.93	286.1	280.7	243.4	1217.1	544.6	1.9	Y327	
4	22.3984	-73.5610	20.01	4.96	7215.35	0.50	174.96	1.07	163.6	145.7	183.7	918.7	603.0	4.1	Y251	r
5	22.4135	-73.5553	18.15	4.97	5699.55	0.58	171.30	0.61	143.6	115.1	109.1	545.3	227.7	2.0	Y312	
6	22.4002	-73.5647	16.54	4.24	4298.67	0.49	168.72	0.71	83.5	86.8	116.9	584.3	257.2	3.0	Y223	r
7	22.4208	-73.5583	16.41	2.66	3721.89	0.64	171.09	0.38	44.7	75.2	57.6	287.9	96.8	1.3		
8	22.4314	-73.5622	16.23	3.99	4976.71	0.68	166.36	0.48	61.9	100.5	68.4	341.8	166.7	1.7	Y287	
9	22.4001	-73.5644	16.22	3.69	3596.83	0.51	167.30	0.69	65.5	72.6	87.5	437.4	258.7	3.6		r
10	22.4369	-73.5573	15.37	4.71	5210.15	0.61	166.32	0.65	123.3	105.2	88.9	444.5	268.0	2.5	Y326	
11	22.4234	-73.5622	15.13	3.71	7338.28	0.76	166.79	0.56	153.5	148.2	81.8	409.1	251.7	1.7		
12	22.4053	-73.5649	14.72	1.86	7512.51	0.77	169.59	0.64	91.6	151.7	82.5	412.6	324.2	2.1	Y227, Y240	r
13	22.3320	-73.5535	14.42	2.79	7856.07	0.77	161.21	0.88	142.5	158.6	85.3	426.6	619.5	3.9	Y217	r
14	22.4049	-73.5643	14.19	2.26	3586.16	0.47	168.89	0.75	56.4	72.4	102.3	511.6	281.7	3.9		r
15	22.4336	-73.5621	13.88	2.41	1850.34	0.54	166.14	0.38	26.9	37.4	40.1	200.6	82.0	2.2		
16	22.3601	-73.5416	13.51	1.76	2758.22	0.51	167.72	0.59	64.0	55.7	67.9	339.5	186.1	3.3	A340	
17	22.4277	-73.5617	13.29	2.95	3921.42	0.52	166.55	0.59	82.5	79.2	93.6	467.9	190.3	2.4		
18	22.3456	-73.5475	13.23	3.63	4710.30	0.53	162.64	0.67	62.6	95.1	106.7	533.3	250.4	2.6	Y285	r
19	22.3787	-73.5684	13.07	2.92	3776.69	0.58	164.70	0.68	101.9	76.3	72.2	360.8	278.6	3.7	Y170	r
20	22.4189	-73.5547	12.43	1.48	2580.00	0.53	170.18	0.53	-	52.1	59.7	298.4	157.5	3.0		
21	22.3354	-73.5518	10.94	3.84	2631.43	0.42	162.40	0.92	78.2	53.1	95.3	476.3	369.6	7.0	Y237	r
22	22.3991	-73.5614	10.58	1.83	1345.89	0.37	177.47	0.54	34.8	27.2	64.3	321.4	112.9	4.2		r
23	22.4154	-73.5547	10.57	2.91	1717.81	0.44	171.33	0.62	26.4	34.7	56.8	283.8	178.0	5.1		
24	22.3575	-73.5798	10.07	2.68	1896.67	0.57	166.28	0.34	40.2	38.3	37.4	187.1	67.7	1.8	Y090	
25	22.4136	-73.5715	10.00	1.19	1000.70	0.38	167.03	0.47	-	20.2	44.0	220.0	88.9	4.4		
26	22.4104	-73.5646	9.92	1.35	1208.22	0.42	169.67	0.45	-	24.4	44.4	222.0	89.6	3.7		r
27	22.4007	-73.5723	9.88	1.70	774.79	0.40	168.02	0.38	12.7	15.6	31.0	155.2	59.7	3.8		
28	22.3326	-73.5521	9.77	1.85	2631.80	0.52	161.14	0.73	39.0	53.1	62.7	313.4	287.3	5.4		r
29	22.4189	-73.5569	9.74	1.19	740.74	0.40	171.20	0.33	-	15.0	29.6	148.2	45.2	3.0		
30	22.4168	-73.5546	9.73	1.78	1530.42	0.43	170.20	0.56	23.5	30.9	54.2	270.8	140.1	4.5		
31	22.3591	-73.5762	9.02	1.35	1653.76	0.40	165.93	0.69	-	33.4	66.9	334.5	198.7	6.0		
32	22.3323	-73.5563	9.02	1.25	1757.42	0.48	162.98	0.67	-	35.5	48.5	242.7	229.8	6.5	Y196	r
33	22.3817	-73.5694	8.74	2.81	892.46	0.36	171.44	0.58	28.9	18.0	43.2	216.2	130.1	7.2	Y162	r
34	22.4150	-73.5647	8.00	1.14	686.99	0.40	169.48	0.36	-	13.9	27.7	138.6	53.0	3.8		r
35	22.3977	-73.5684	7.91	1.51	732.43	0.38	170.33	0.45	-	14.8	32.3	161.7	81.3	5.5		r
36	22.3996	-73.5734	7.90	1.27	943.82	0.44	166.46	0.43	-	19.1	31.7	158.7	84.3	4.4		
37	22.3955	-73.5677	7.73	1.16	805.49	0.38	171.55	0.49	-	16.3	35.6	177.8	94.8	5.8		r

Table A1 *continued*

Table A1 *(continued)*

ID	RA	Dec	$^{12}CO_{pk}$	$^{13}CO_{pk}$	W_{12}	R	$v_{LSRK,12}$	$\sigma_{v,12}$	M_{LTE}	$M_{X_{CO,B}}$	Σ	ρ_c	M_{vir}	α_{vir}	Assoc. YSO	Group
	(°)	(°)	(K)	(K)	(K km s^{-1})	(pc)	(km s^{-1})	(km s^{-1})	($M_\odot$)	($M_\odot$)	($M_\odot$ pc^{-2})	($M_\odot$ pc^{-3})	($M_\odot$)			
38	22.3849	-73.5691	7.55	3.35	1028.77	0.37	170.68	0.68	34.3	20.8	47.5	237.4	182.3	8.8	Y171	r
39	22.3970	-73.5643	7.37	1.38	842.36	0.42	166.68	0.53	-	17.0	30.5	152.7	123.8	7.3		r
40	22.3919	-73.5699	7.30	1.04	905.78	0.43	170.20	0.50	-	18.3	31.2	156.0	111.6	6.1	Y174	r
41	22.4052	-73.5700	7.26	1.38	697.39	0.41	170.67	0.45	-	14.1	26.1	130.7	87.5	6.2		
42	22.3303	-73.5558	7.00	1.55	1461.28	0.51	162.82	0.69	-	29.5	35.6	177.8	259.0	8.8	Y197	r
43	22.3272	-73.5565	6.98	1.23	934.29	0.42	163.04	0.57	-	18.9	33.3	166.5	144.1	7.6	Y179	r
44	22.4006	-73.5748	6.92	1.39	1363.03	0.62	166.89	0.67	-	27.5	22.5	112.7	289.2	10.5		
45	22.3971	-73.5735	6.89	1.15	1016.27	0.43	165.09	0.65	-	20.5	35.6	177.8	190.1	9.3	Y148	
46	22.3926	-73.5594	6.79	1.31	691.45	0.40	177.65	0.38	-	14.0	27.4	137.2	60.9	4.4		r
47	22.4063	-73.5663	6.74	1.38	1452.12	0.48	170.65	0.57	-	29.3	40.1	200.6	164.4	5.6		r
48	22.3403	-73.5440	6.72	1.55	1711.77	0.53	162.08	0.56	-	34.6	39.6	198.2	175.5	5.1		r
49	22.3350	-73.5548	6.70	1.53	832.18	0.37	164.05	0.61	-	16.8	38.8	193.9	146.9	8.7	Y206	r
50	22.3423	-73.5468	6.55	1.28	952.40	0.43	160.74	0.61	-	19.2	32.6	163.1	169.7	8.8		r
51	22.3896	-73.5740	6.42	1.12	614.31	0.41	162.82	0.44	-	12.4	23.7	118.3	81.4	6.6		
52	22.4097	-73.5641	6.11	1.19	460.36	0.35	168.86	0.46	-	9.3	23.7	118.3	77.3	8.3		r
53	22.3946	-73.5736	5.67	0.90	509.58	0.43	164.36	0.37	-	10.3	17.6	88.1	61.1	5.9		
54	22.4073	-73.5660	5.67	1.26	1443.38	0.42	168.87	0.73	-	29.1	52.8	264.2	232.7	8.0		r
55	22.4035	-73.5640	5.67	1.81	363.58	0.33	175.09	0.46	16.3	7.3	21.2	105.8	73.6	10.0		r
56	22.4025	-73.5727	5.65	1.20	559.86	0.40	168.54	0.49	-	11.3	22.4	111.8	99.0	8.8		
57	22.3349	-73.5460	5.60	1.08	528.11	0.39	161.49	0.45	-	10.7	21.8	108.9	84.7	7.9		r
58	22.4054	-73.5650	5.52	1.19	669.50	0.40	172.79	0.56	-	13.5	26.7	133.5	132.5	9.8		r
59	22.3391	-73.5456	5.48	1.03	353.19	0.34	160.94	0.37	-	7.1	19.3	96.7	48.8	6.8		r
60	22.4206	-73.5718	5.39	0.85	297.48	0.39	166.52	0.28	-	6.0	12.7	63.6	32.7	5.4		
61	22.3861	-73.5759	5.37	0.96	253.89	0.35	169.09	0.33	-	5.1	13.5	67.7	39.6	7.7		
62	22.4055	-73.5647	5.19	1.03	330.24	0.37	171.85	0.37	-	6.7	15.8	79.1	53.0	7.9		r
63	22.4037	-73.5653	5.19	1.06	943.70	0.43	166.95	0.64	-	19.1	32.7	163.4	186.7	9.8		r
64	22.3686	-73.5756	5.12	0.91	325.28	0.37	163.52	0.37	-	6.6	15.5	77.6	51.3	7.8		
65	22.4047	-73.5600	4.97	1.30	329.91	0.34	168.44	0.46	-	6.7	18.0	89.8	76.3	11.5	Y264	r
66	22.3406	-73.5466	4.77	1.08	464.89	0.45	161.06	0.43	-	9.4	14.8	73.9	85.9	9.1		r
67	22.3292	-73.5644	4.62	1.28	548.05	0.42	162.92	0.49	-	11.1	19.5	97.6	107.8	9.7		r
68	22.3265	-73.5572	4.41	1.39	740.07	0.42	162.90	0.67	-	14.9	27.1	135.5	195.5	13.1		r
69	22.3185	-73.5669	4.39	1.15	293.45	0.37	165.99	0.39	-	5.9	13.6	68.1	59.8	10.1		r
70	22.4111	-73.5636	4.31	1.36	217.83	0.32	165.19	0.40	-	4.4	13.7	68.7	54.7	12.4		r
71	22.4083	-73.5658	4.26	0.96	513.34	0.41	167.73	0.40	-	10.4	20.0	99.8	69.7	6.7		r
72	22.3363	-73.5683	4.22	1.15	464.06	0.43	162.36	0.53	-	9.4	16.4	82.1	126.3	13.5		
73	22.4065	-73.5619	4.20	0.81	194.54	0.31	179.52	0.30	-	3.9	12.7	63.4	30.5	7.8	Y255	r
74	22.4224	-73.5715	4.08	1.02	209.90	0.39	166.81	0.24	-	4.2	9.0	45.1	22.9	5.4		
75	22.3974	-73.5773	4.04	1.05	433.77	0.49	167.69	0.28	-	8.8	11.8	59.2	38.8	4.4		
76	22.3687	-73.5699	4.03	0.97	420.55	0.43	172.38	0.47	-	8.5	15.0	74.8	98.0	11.5	Y142	r

Table A1 *continued*

Table A1 *(continued)*

ID	RA	Dec	$^{12}CO_{pk}$	$^{13}CO_{pk}$	W_{12}	R	$v_{LSRK,12}$	$\sigma_{v,12}$	M_{LTE}	$M_{X_{CO,B}}$	Σ	ρ_c	M_{vir}	α_{vir}	Assoc. YSO	Group
	(°)	(°)	(K)	(K)	(K km s^{-1})	(pc)	(km s^{-1})	(km s^{-1})	($M_\odot$)	($M_\odot$)	($M_\odot$ pc^{-2})	($M_\odot$ pc^{-3})	($M_\odot$)			
77	22.3818	-73.5721	3.99	1.10	215.29	0.35	171.62	0.36	-	4.3	11.5	57.5	47.8	11.0		
78	22.3841	-73.5757	3.98	1.65	139.30	0.30	164.30	0.29	-	2.8	9.6	48.2	26.7	9.5		
79	22.3734	-73.5751	3.93	1.35	179.92	0.32	170.87	0.34	-	3.6	11.4	56.9	38.1	10.5	Y118	
80	22.3435	-73.5457	3.82	1.29	196.62	0.33	160.71	0.34	-	4.0	11.4	57.0	41.3	10.4		r
81	22.3399	-73.5450	3.81	1.05	718.68	0.41	162.25	0.73	-	14.5	27.5	137.3	228.4	15.7		r
82	22.3666	-73.5751	3.76	1.04	180.02	0.36	165.65	0.32	-	3.6	8.9	44.4	38.8	10.7		
83	22.4257	-73.5635	3.74	1.10	151.55	0.33	169.65	0.38	-	3.1	8.7	43.7	50.4	16.5		
84	22.3195	-73.5611	3.66	0.86	206.17	0.36	162.19	0.40	-	4.2	10.1	50.3	59.8	14.4	Y143	r
85	22.4143	-73.5649	3.63	0.89	123.62	0.31	167.28	0.33	-	2.5	8.2	41.1	34.7	13.9		r
86	22.3300	-73.5549	3.56	1.24	322.38	0.40	162.55	0.42	-	6.5	12.9	64.7	72.9	11.2		r
87	22.3910	-73.5600	3.56	1.12	162.66	0.33	177.21	0.38	-	3.3	9.8	48.8	50.5	15.4		r
88	22.4038	-73.5721	3.54	0.75	164.01	0.31	168.88	0.33	-	3.3	10.8	54.1	35.5	10.7		
89	22.3265	-73.5649	3.47	1.03	184.61	0.33	162.45	0.41	-	3.7	11.1	55.3	57.8	15.5		r
90	22.4005	-73.5745	3.20	0.91	174.52	0.37	164.11	0.31	-	3.5	8.0	40.2	38.4	10.9		
91	22.3953	-73.5727	3.16	0.72	120.61	0.26	165.04	0.35	-	2.4	11.1	55.6	33.1	13.6		
92	22.3961	-73.5563	3.11	0.73	89.55	0.31	177.72	0.28	-	1.8	5.8	29.2	26.0	14.4		r
93	22.3255	-73.5644	3.07	1.26	308.87	0.38	163.82	0.50	-	6.2	13.8	68.8	99.6	16.0		r
94	22.3362	-73.5476	2.95	1.02	81.51	0.30	162.29	0.28	-	1.6	5.9	29.3	25.2	15.3		r
95	22.3087	-73.5570	2.92	0.98	72.28	0.30	162.47	0.23	-	1.5	5.0	25.1	16.5	11.3		
96	22.3258	-73.5522	2.91	1.32	80.05	0.31	161.55	0.24	-	1.6	5.4	27.1	19.3	11.9		r
97	22.3906	-73.5716	2.90	0.87	83.71	0.31	163.52	0.25	-	1.7	5.6	27.9	20.5	12.1		
98	22.3394	-73.5430	2.81	0.52	123.97	0.35	160.78	0.30	-	2.5	6.5	32.4	33.1	13.2		r
99	22.4028	-73.5731	2.76	0.85	82.34	0.28	167.84	0.25	-	1.7	6.5	32.6	19.1	11.5		
100	22.3662	-73.5755	2.74	1.04	84.38	0.26	164.22	0.27	-	1.7	7.7	38.7	20.1	11.8		
101	22.3162	-73.5620	2.69	0.83	170.67	0.37	161.04	0.35	-	3.4	8.2	41.2	46.7	13.5		r
102	22.4179	-73.5538	2.68	0.82	94.77	0.28	171.81	0.32	-	1.9	7.8	39.0	29.4	15.3		
103	22.4082	-73.5589	2.65	1.22	98.20	0.28	168.46	0.38	-	2.0	7.9	39.7	43.2	21.8		r
104	22.3223	-73.5644	2.65	0.57	103.14	0.30	164.09	0.35	-	2.1	7.2	35.8	39.2	18.8		r
105	22.3216	-73.5625	2.58	0.98	396.29	0.47	162.03	0.51	-	8.0	11.7	58.3	127.9	16.0		r
106	22.3683	-73.5777	2.56	1.13	121.61	0.34	170.15	0.35	-	2.5	6.9	34.5	44.2	18.0		
107	22.3374	-73.5462	2.56	1.27	204.92	0.37	159.52	0.64	-	4.1	9.4	47.2	161.5	39.0		r
108	22.3201	-73.5615	2.50	1.04	93.63	0.35	160.73	0.28	-	1.9	4.9	24.6	28.6	15.1		r
109	22.3906	-73.5732	2.36	0.81	59.62	0.28	164.27	0.24	-	1.2	4.8	24.1	17.6	14.6		
110	22.3332	-73.5466	2.25	0.69	69.29	0.37	160.04	0.26	-	1.4	3.3	16.6	25.4	18.1		r

Note—Clumps without LTE mass estimates have ^{13}CO S/N < 3. ρ_c is the volume density at 0.1 pc. "Assoc. YSO" gives the first four characters of the identifiers assigned by Carlson et al. (2011) and listed in Table A2 of any YSOs which fall within the projected 2D boundaries of the clump. A denotation of 'r' in Group indicates the clump is on the N90 rim. Error estimates for most quantities are available in a complete, machine-readable version of this table.

Table A2. YSO Properties

ID	Name	RA	Dec	M	$\dot{M}$
		(°)	(°)	($M_\odot$)	(10^{-6} $M_\odot$ yr^{-1})
Y090	J012925.97-733446.8	22.3582	-73.5797	3.71	35.58
Y096	J012906.41-733348.6	22.2767	-73.5635	7.78	3.50
Y118	J012929.62-733430.1	22.3734	-73.5750	2.00	0.36
Y142	J012928.56-733411.9	22.3690	-73.5700	3.54	1.93
Y143	J012916.77-733340.7	22.3199	-73.5613	2.02	0.19
Y148	J012935.11-733423.9	22.3963	-73.5733	3.64	7.48
Y149	J012856.16-733242.4	22.2340	-73.5451	2.93	0.33
Y162	J012931.68-733409.2	22.3820	-73.5692	2.91	91.79
Y163	J012859.33-733244.8	22.2472	-73.5458	2.89	0.43
Y170	J012930.90-733405.6	22.3788	-73.5682	3.33	11.20
Y171	J012932.39-733408.4	22.3850	-73.5690	3.28	6.61
Y174	J012934.05-733411.7	22.3919	-73.5699	3.58	0.94
Y179	J012918.44-733324.9	22.3269	-73.5569	2.24	4.22
Y196	J012919.88-733322.5	22.3328	-73.5563	3.76	0.75
Y197	J012918.95-733319.4	22.3290	-73.5554	3.57	1.18
Y198	J012936.38-733403.6	22.4016	-73.5677	5.94	6.05
Y206	J012920.48-733316.8	22.3353	-73.5547	2.88	1.75
Y217	J012919.87-733312.5	22.3328	-73.5535	3.88	1.40
Y223	J012935.89-733351.7	22.3996	-73.5644	3.93	34.07
Y227	J012937.37-733352.4	22.4057	-73.5646	6.99	0.38
Y237	J012920.64-733306.3	22.3360	-73.5518	4.14	3.56
Y240	J012937.99-733352.8	22.4083	-73.5647	6.17	4.24
Y251	J012935.64-733339.5	22.3985	-73.5610	3.50	16.98
Y255	J012937.60-733342.9	22.4067	-73.5619	3.23	4.28
Y264	J012936.99-733336.1	22.4041	-73.5600	3.20	0.29
A270	J012935.51-733330.3	22.3980	-73.5584	6.48	36.11
Y270i	J012935.51-733330.3	22.3980	-73.5584	6.16	251.03
Y271	J012933.43-733323.3	22.3893	-73.5565	2.32	0.68
Y283	J012930.20-733310.5	22.3759	-73.5529	2.23	0.76
Y285	J012923.06-733251.5	22.3461	-73.5476	2.92	0.07
Y287	J012943.36-733343.1	22.4307	-73.5620	3.19	0.24
Y288	J012942.42-733341.6	22.1768	-73.0120	3.58	32.66
Y290	J012937.08-733325.4	22.4045	-73.5570	3.50	0.08
Y312	J012939.17-733318.9	22.4132	-73.5553	3.31	0.39
Y326	J012944.75-733325.2	22.4365	-73.5570	2.79	2.51
Y327	J012935.69-733302.1	22.3987	-73.5506	6.45	2.27
Y340	J012926.56-733230.3	22.3607	-73.5418	2.33	0.17
A340	J012926.56-733230.3	22.3607	-73.5418	2.77	1.89
Y358	J012935.17-733242.5	22.3966	-73.5452	2.19	2.70
Y396	J012924.20-733152.8	22.3508	-73.5313	2.37	0.21
Y493	J012915.95-733017.6	22.3165	-73.5049	1.60	0.03
Y700	J013006.79-733258.9	22.5286	-73.5497	2.56	5.35
K340	J012926.56-733230.3	22.3607	-73.5418	3.53	0.03
U364	J012908.96-733129.5	22.2873	-73.5249	2.41	2.81
U703	J012911.03-733039.6	22.2961	-73.5110	1.21	0.02

Note—IDs are as given in Carlson et al. (2011).

REFERENCES

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 710, 133
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Bertoldi, F., & McKee, C. F. 1992, ApJ, 395, 140
Bolatto, A. D., Leroy, A. K., Rosolowsky, E., Walter, F., & Blitz, L. 2008, ApJ, 686, 948
Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207
Bourke, T. L., Garay, G., Lehtinen, K. K., et al. 1997, ApJ, 476, 781
Bradski, G. 2000, Dr. Dobb's Journal of Software Tools
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Carlson, L. R., Sabbi, E., Sirianni, M., et al. 2007, ApJL, 665, L109
Carlson, L. R., Sewiło, M., Meixner, M., et al. 2011, ApJ, 730, 78
Caselli, P., Benson, P. J., Myers, P. C., & Tafalla, M. 2002, ApJ, 572, 238
Caselli, P., & Myers, P. C. 1995, ApJ, 446, 665
Chen, H. H.-H., Offner, S. S. R., Pineda, J. E., et al. 2020, arXiv e-prints, arXiv:2006.07325, submitted to ApJ
Chen, H. H.-H., Pineda, J. E., Goodman, A. A., et al. 2019, ApJ, 877, 93
Chevance, M., Madden, S. C., Fischer, C., et al. 2020, MNRAS, 494, 5279
Cignoni, M., Sabbi, E., Nota, A., et al. 2009, AJ, 137, 3668
Da Rio, N., & Robberto, M. 2012, AJ, 144, 176
De Marchi, G., Beccari, G., & Panagia, N. 2013, ApJ, 775, 68
Dobbs, C. L. 2015, MNRAS, 447, 3390
Dolphin, A. 2016, DOLPHOT: Stellar photometry
Dolphin, A. E. 2000, PASP, 112, 1383
Elmegreen, B. G. 1989, ApJ, 338, 178
Elmegreen, B. G., & Scalo, J. 2004, ARA&A, 42, 211
Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining (AAAI Press), 226–231
Evans, C. J., Hainich, R., Oskinova, L. M., et al. 2012, ApJ, 753, 173
Fahrion, K., Cormier, D., Bigiel, F., et al. 2017, A&A, 599, A9
Field, G. B., Blackman, E. G., & Keto, E. R. 2011, MNRAS, 416, 710
Finn, M. K., Indebetouw, R., Johnson, K. E., et al. 2022, AJ, 164, 64
Fukui, Y., Habe, A., Inoue, T., Enokiya, R., & Tachihara, K. 2021, PASJ, 73, S1
Fukui, Y., Ohno, T., Tsuge, K., Sano, H., & Tachihara, K. 2020, arXiv e-prints, arXiv:2005.13750, submitted to PASJ
Fulmer, L. M., Gallagher III, J. S., Hamann, W.-R., Oskinova, L. M., & Ramachandran, V. 2020, A&A, 633, A164
Galametz, M., Madden, S., Galliano, F., et al. 2009, A&A, 508, 645
Garden, R. P., Hayashi, M., Gatley, I., Hasegawa, T., & Kaifu, N. 1991, ApJ, 374, 540
Ginsburg, A., Koch, E., Robitaille, T., et al. 2019, radio-astro-tools/spectral-cube: v0.4.4
Glover, S. C. O., & Mac Low, M.-M. 2011, MNRAS, 412, 337
Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, ApJ, 594, 279
Gordon, K. D., Engelbracht, C. W., Rieke, G. H., et al. 2008, ApJ, 682, 336
Gordon, K. D., Bot, C., Muller, E., et al. 2009, ApJ, 690, L76
Gordon, K. D., Meixner, M., Meade, M., et al. 2011, AJ, 142, 102
Gordon, K. D., Roman-Duval, J., Bot, C., et al. 2014, ApJ, 797, 85
Gouliermis, D. A., Quanz, S. P., & Henning, T. 2007, ApJ, 665, 306
Gouliermis, D. A., Schmeja, S., Dolphin, A. E., et al. 2012, ApJ, 748, 64
Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, Sci, 307, 1292
Gritschneder, M., Burkert, A., Naab, T., & Walch, S. 2010, ApJ, 723, 971
Gvaramadze, V. V., Kniazev, A. Y., Gallagher, J. S., et al. 2021, MNRAS, 503, 3856
Haberl, F., Sturm, R., Filipović, M. D., Pietsch, W., & Crawford, E. J. 2012, A&A, 537, L1
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Henize, K. G. 1956, ApJS, 2, 315
Hermelo, I., Lisenfeld, U., Relaño, M., et al. 2013, A&A, 549, A70
Heyer, M., Krawczyk, C., Duval, J., & Jackson, J. M. 2009, ApJ, 699, 1092
Heyer, M. H., & Brunt, C. M. 2004, ApJL, 615, L45
Hilditch, R. W., Howarth, I. D., & Harries, T. J. 2005, MNRAS, 357, 304
Hodge, P. W. 1983, ApJ, 264, 470

Hughes, A., Wong, T., Ott, J., et al. 2010, MNRAS, 406, 2065
Hughes, A., Meidt, S. E., Colombo, D., et al. 2013, ApJ, 779, 46
Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
Hutchings, J. B., Thompson, I. B., Cartledge, S., & Pazder, J. 1991, AJ, 101, 933
Hénault-Brunet, V., Oskinova, L. M., Guerrero, M. A., et al. 2012, MNRAS: Letters, 420, L13
Ikeda, N., Kitamura, Y., & Sunada, K. 2009, ApJ, 691, 1560
Imara, N., & Faesi, C. M. 2019, ApJ, 876, 141
Indebetouw, R., Wong, T., Chen, C.-H. R., et al. 2020, ApJ, 888, 56
Indebetouw, R., Brogan, C., Chen, C.-H. R., et al. 2013, ApJ, 774, 73
Jameson, K. E., Bolatto, A. D., Leroy, A. K., et al. 2016, ApJ, 825, 12
Jameson, K. E., Bolatto, A. D., Wolfire, M., et al. 2018, ApJ, 853, 111
Kalari, V. M., Rubio, M., Saldaño, H. P., & Bolatto, A. D. 2020, MNRAS, 499, 2534
Kauffmann, J., Pillai, T., & Goldsmith, P. F. 2013, ApJ, 779, 185
Kazandjian, M. V., Pelupessy, I., Meijerink, R., Israel, F. P., & Spaans, M. 2016, A&A, 595, A125
Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531
Kepley, A. A., Leroy, A. K., Johnson, K. E., Sandstrom, K., & Chen, C. H. R. 2016, ApJ, 828, 50
Kroupa, P. 2001, MNRAS, 322, 231
Krumholz, M. R. 2013, MNRAS, 436, 2747
Kurucz, R. L. 1993, SYNTHE spectrum synthesis programs and line data
Lada, C. J., Muench, A. A., Rathborne, J., Alves, J. F., & Lombardi, M. 2008, ApJ, 672, 410
Langer, W. D., Velusamy, T., Pineda, J. L., Willacy, K., & Goldsmith, P. F. 2014, A&A, 561, A122
Larson, R. B. 1981, MNRAS, 194, 809
Lebouteiller, V., Cormier, D., Madden, S. C., et al. 2019, A&A, 632, A106
Lee, C., Leroy, A. K., Schnee, S., et al. 2015, MNRAS, 450, 2708
Lee, E. J., Chang, P., & Murray, N. 2015, ApJ, 800, 49
Lee, E. J., Miville-Deschênes, M.-A., & Murray, N. W. 2016, ApJ, 833, 229
Lee, J.-K., Rolleston, W. R. J., Dufton, P. L., & Ryans, R. S. I. 2005, A&A, 429, 1025
Lee, M.-Y., Stanimirović, S., Douglas, K. A., et al. 2012, ApJ, 748, 75
Lehner, N., Howk, J. C., Keenan, F. P., & Smoker, J. V. 2008, ApJ, 678, 219
Leroy, A., Bolatto, A., Stanimirovic, S., et al. 2007, ApJ, 658, 1027
Lin, Y., Wyrowski, F., Liu, H. B., et al. 2022, A&A, 658, A128
MacLaren, I., Richardson, K. M., & Wolfendale, A. W. 1988, ApJ, 333, 821
Madden, S. C., Galliano, F., Jones, A. P., & Sauvage, M. 2006, A&A, 446, 877
Madden, S. C., Cormier, D., Hony, S., et al. 2020, A&A, 643, A141
Mangum, J. G., Ginsburg, A. G., Henkel, C., et al. 2019, ApJ, 871, 170
McKinney, W. 2010, in Proc. 9th Python in Science Conf., ed. Stéfan van der Walt & Jarrod Millman, 56
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Ser., Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127
Meaburn, J. 1980, MNRAS, 192, 365
Meixner, M., Panuzzo, P., Roman-Duval, J., et al. 2013, AJ, 146, 62
Muraoka, K., Homma, A., Onishi, T., et al. 2017, ApJ, 844, 98
Myers, P. C., Hatchfield, H. P., & Battersby, C. 2022, ApJ, 929, 34
Naslim, N., Tokuda, K., Onishi, T., et al. 2018, ApJ, 853, 175
Nayak, O., Meixner, M., Fukui, Y., et al. 2018, ApJ, 854, 154
Nayak, O., Meixner, M., Indebetouw, R., et al. 2016, ApJ, 831, 32
Nigra, L., Gallagher, J. S., Smith, L. J., et al. 2008, PASP, 120, 972
Nikolić, S., Garay, G., Rubio, M., & Johansson, L. E. B. 2007, A&A, 471, 561
Ochsendorf, B. B., Meixner, M., Chastenet, J., Tielens, A. G. G. M., & Roman-Duval, J. 2016, ApJ, 832, 43
Ochsendorf, B. B., Meixner, M., Roman-Duval, J., Rahman, M., & Evans, N. J. 2017, ApJ, 841, 109
Oka, T., Hasegawa, T., Sato, F., et al. 2001, ApJ, 562, 348
O'Neill, T. J., Indebetouw, R., Bolatto, A. D., Madden, S. C., & Wong, T. 2022, ApJ, 933, 179
Oskinova, L. M., Sun, W., Evans, C. J., et al. 2013, ApJ, 765, 73
pandas development team, T. 2020, pandas-dev/pandas: Pandas
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pineda, J. L., Langer, W. D., Velusamy, T., & Goldsmith, P. F. 2013, A&A, 554, A103

Pineda, J. L., Langer, W. D., Goldsmith, P. F., et al. 2017, ApJ, 839, 107

Pirogov, L. E. 2009, ARep, 53, 1127

Rahmah, N., & Sitanggang, I. S. 2016, IOP Conf. Ser.: Earth Environ. Sci., 31, 012012

Ramachandran, V., Hamann, W.-R., Oskinova, L. M., et al. 2019, A&A, 625, A104

Requena-Torres, M. A., Israel, F. P., Okada, Y., et al. 2016, A&A, 589, A28

Ridge, N. A., Francesco, J. D., Kirk, H., et al. 2006, AJ, 131, 2921

Robitaille, T. 2019, APLpy v2.0: The Astronomical Plotting Library in Python

Robitaille, T., & Bressert, E. 2012, APLpy: Astronomical Plotting Library in Python, Astrophysics Source Code Library, ascl:1208.017

Robitaille, T., Rice, T., Beaumont, C., et al. 2019, astrodendro: Astronomical data dendrogram creator

Robitaille, T. P. 2017, A&A, 600, A11

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

Rolleston, W. R. J., Dufton, P. L., McErlean, N. D., & Venn, K. A. 1999, A&A, 348, 728

Rubele, S., Girardi, L., Kerber, L., et al. 2015, MNRAS, 449, 639

Rubele, S., Pastorelli, G., Girardi, L., et al. 2018, MNRAS, 478, 5017

Russell, S. C., & Dopita, M. A. 1992, ApJ, 384, 508

Saigo, K., Onishi, T., Nayak, O., et al. 2017, ApJ, 835, 108

Sandstrom, K. M., Bolatto, A. D., Draine, B. T., Bot, C., & Stanimirović, S. 2010, ApJ, 715, 701

Schmalzl, M., Gouliermis, D. A., Dolphin, A. E., & Henning, T. 2008, ApJ, 681, 290

Schruba, A., Leroy, A. K., Walter, F., et al. 2012, AJ, 143, 138

Schruba, A., Leroy, A. K., Kruijssen, J. M. D., et al. 2017, ApJ, 835, 278

Schultz, G. V., & Wiemer, W. 1975, A&A, 43, 133

Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conference

Seale, J. P., Looney, L. W., Wong, T., et al. 2012, ApJ, 751, 42

Seale, J. P., Meixner, M., Sewiło, M., et al. 2014, AJ, 148, 124

Shetty, R., Beaumont, C. N., Burton, M. G., Kelly, B. C., & Klessen, R. S. 2012, MNRAS, 425, 720

Sidorin, V. 2017, Quickclump: Identify clumps within a 3D FITS datacube

Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, ApJ, 319, 730

Staveley-Smith, L., Sault, R. J., Hatzidimitriou, D., Kesteven, M. J., & McConnell, D. 1997, MNRAS, 289, 225

Szűcs, L., Glover, S. C. O., & Klessen, R. S. 2016, MNRAS, 460, 82

Tatematsu, K., Umemoto, T., Kandori, R., & Sekimoto, Y. 2004, ApJ, 606, 333

Tchernyshyov, K., Meixner, M., Seale, J., et al. 2015, ApJ, 811, 78

Tognelli, E., Prada Moroni, P. G., & Degl'Innocenti, S. 2011, A&A, 533, A109

Valdivia-Mena, M. T., Rubio, M., Bolatto, A. D., Saldaño, H. P., & Verdugo, C. 2020, A&A, 641, A97

van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627

Velusamy, T., Langer, W. D., Pineda, J. L., et al. 2010, A&A, 521, L18

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Waskom, M. L. 2021, Journal of Open Source Software, 6, 3021

Welty, D. E., Xue, R., & Wong, T. 2012, ApJ, 745, 173

Westerlund, B. E. 1964, MNRAS, 127, 429

Williams, J. P., de Geus, E. J., & Blitz, L. 1994, ApJ, 428, 693

Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191

Wolfire, M. G., Hollenbach, D., McKee, C. F., Tielens, A. G. G. M., & Bakes, E. L. O. 1995, ApJ, 443, 152

Wong, T., Hughes, A., Tokuda, K., et al. 2017, ApJ, 850, 139

—. 2019, ApJ, 885, 50

Wong, T., Oudshoorn, L., Sofovich, E., et al. 2022, ApJ, 932, 47

Xu, D., Li, D., Yue, N., & Goldsmith, P. F. 2016, ApJ, 819, 22

Yonekura, Y., Dobashi, K., Mizuno, A., Ogawa, H., & Fukui, Y. 1997, ApJS, 110, 21

4

Identifying PMS Stars in the LMC with Machine Learning

Draft version December 16, 2022
Typeset using LaTeX twocolumn style in AASTeX631

Identifying Pre-main-sequence Stars in the Star-forming Region N159 with Machine Learning

Theo J. O'Neill[1] and Rémy Indebetouw[1, 2]

[1] Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA
[2] National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA

ABSTRACT

We explore machine learning-based techniques to identify pre-main-sequence (PMS) stars in the star-forming region N159 in the Large Magellanic Cloud. We analyze of Hubble Space Telescope optical and infrared photometry of stars near three giant molecular clouds (GMCs) in N159, each of which displays dramatically different star formation activities and histories. We estimate the reddening laws in the regions surrounding these GMCs by fitting the slope of the red clump using the machine learning algorithm RANSAC before deriving an extinction map across N159 from its upper-main-sequence population. We find a mean slope of $R_{555-814}(555)$ =2.5 and typical extinctions A_{555} of 1 – 2.5 mag. We correct our HST photometry for extinction using this map and apply a support vector machine (SVM) classifier to identify probable PMS stars in N159. Our trained machine learning model is able to predict PMS membership with 95% accuracy and yields a sample of 10,000 PMS candidates in N159. PMS candidates are generally associated with the GMCs. Future work will quantitatively analyze the relationship between N159's PMS population, GMCs, and other young stellar objects, and explore the impact of other machine learning methods.

1. INTRODUCTION

Stars form by accreting mass from interstellar gas clouds. Young stars that are actively accreting, or that have only recently stopped accreting but not yet achieved a stable internal structure, are referred to as "pre-main-sequence" (PMS). This is followed by the subsequent and much longer (>10x) evolutionary stage along which stars spend most of their lives — the "main sequence." PMS stars retain memories of their gaseous births in both their relative ages and locations compared to any remaining gas, making them powerful probes of the recent histories and likely futures of star-forming regions. Studying the PMS stage is then critical in understanding the process of star formation.

Current techniques used to identify PMS stars frequently involve drawing strict barriers between observed colors and magnitudes of candidate stars based on evolutionary modeling. This approach breaks down in complex environments. Active star-forming regions typically contain large reservoirs of dust that interfere with estimations of stellar ages by shifting intermediate-mass, non-PMS stars to overlap with true PMS stars in color-magnitude space; this effect is known as extinction. Large uncertainties on distances to individual stars within a region, and difficulties identifying stars along our line of sight that are not associated with the region at all, similarly complicate identification. These factors create significant confusion and make the use of more sophisticated PMS identification methods necessary for accurate analysis of star formation.

In this work, we develop such methods in the star-forming region N159 in the Large Magellanic Cloud (LMC). The LMC is a low metallicity environment ($Z \sim 1/2\ Z_{\odot}$, Russell & Dopita 1992) approximating the conditions under which the earliest stars in the Universe formed. It is one of the Milky Way's nearest companion galaxies (d ∼ 50 kpc, Schaefer 2008) and is relatively shallow along the line of sight, which reduces the effects of uncertain distances on PMS star identification. The LMC then provides an ideal window into how star formation proceeds in environments that are dramatically different from the Milky Way, and to experiment with new methods of PMS star identification.

As shown in Figure 1, N159 is situated along a kiloparsecs-long giant molecular cloud (GMC) complex in the LMC. The active star-forming region 30 Doradus is located north of N159, and a quiescent molecular "ridge" is located to the south. There is a significant gradient in massive star formation activity between these two ends of the cloud complex: 30 Doradus is one of the largest, most active massive star-forming regions ever identified within the local group, while the ridge has almost no massive star formation despite containing 1/3 of the total molecular gas content of the LMC (Mizuno et al. 2001; Indebetouw et al. 2008). Star formation ac-

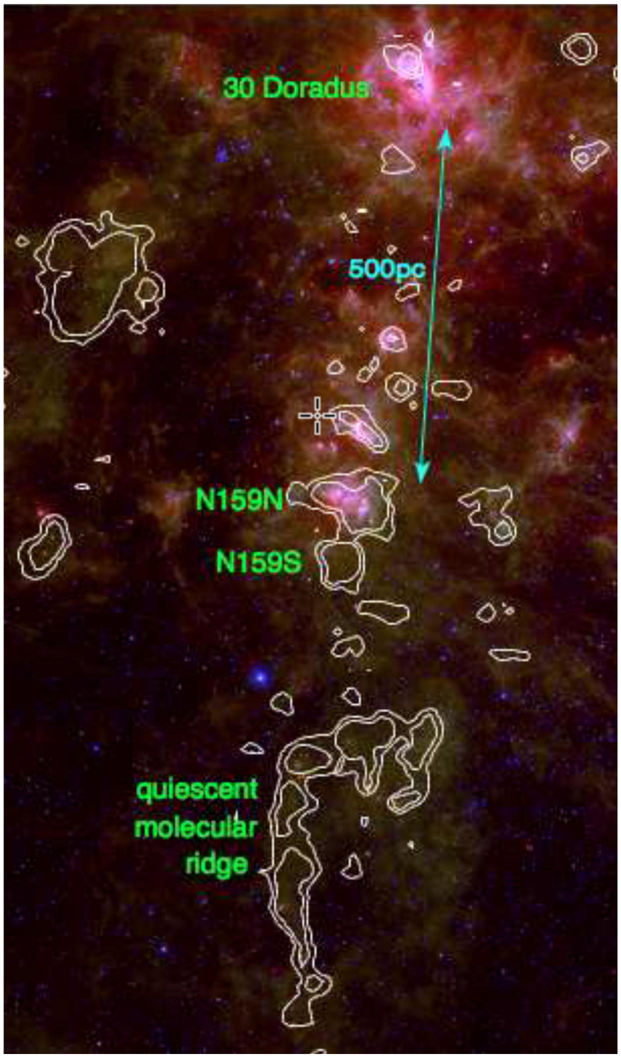

Figure 1. N159 in the LMC, with the star-forming region 30 Doradus to the north and the quiescent molecular ridge to the south. Colors are (R,G,B) = (24, 8, 3.6μm) and white contours are integrated CO emission.

tivity in N159 follows a similar north-south trend. As shown in Figure 2, it contains three GMCs: the southern GMC (N159S) has no massive star formation, and the two northern GMCs (N159E and N159W) both display high levels of massive star-formation but do not appear to share the same evolutionary stage or physical cause for the onset of star formation (Chen et al. 2010; Nayak et al. 2018).

Despite N159's suggestive position along the transition from 30 Doradus to the ridge, it is not yet clear why its three GMCs differ so dramatically in their star formation activities in spite of their physical proximity. By being situated between two extremes of low-metallicity star formation (30 Doradus and the ridge) and by virtue of the proximity of the LMC, N159 is an ideal candidate

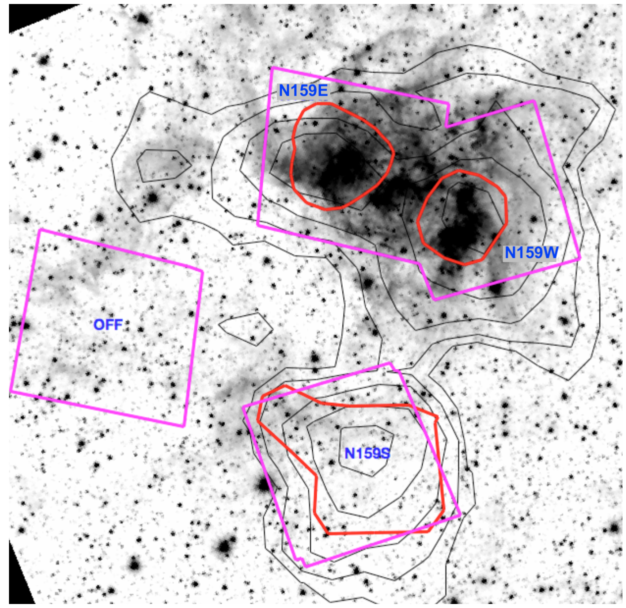

Figure 2. Spitzer 3.6μm image of N159. HST coverage shown in magenta, with black contours of CO intensity overlaid.

to test the effects of feedback and environment on star formation and to refine the next generation of PMS star-identification methods.

In this work, we develop methods to identify PMS stars and uncover the recent star formation history of N159 using supervised machine learning. We describe the data used in this work in §2, which includes Hubble Space Telescope (HST) observations of N159 (§2.1), photometry of bright LMC stars from the Magellanic Cloud Photometric Survey (Zaritsky et al. 2004) (§2.2), and a "training set" of stars in the star cluster R136 in 30 Doradus identified as likely being PMS vs. non-PMS (Ksoll et al. 2018) (§2.3). We measure the slope of the reddening vector in N159 through analysis of red clump stars (§3.1), and use this information to estimate the extinction of HST-identified stars using the Zaritsky et al. (2004) upper main sequence populations. Finally, we train a support vector machine (SVM) using the R136 dataset (§4.1) and apply it to the HST stars in N159 (§4.2). We conclude with a preview of future work in §5.

2. OBSERVATIONS AND ARCHIVAL DATA

2.1. *HST Photometry*

Optical and near-infrared observations of N159 were obtained using the Hubble Space Telescope (HST, PI: R. Indebetouw) in bands F125W (∼J), F160W (∼H), F555W (∼V), and F814W (∼I). Transmission curves for these filters are shown in Figure 3, and Figure 4 shows

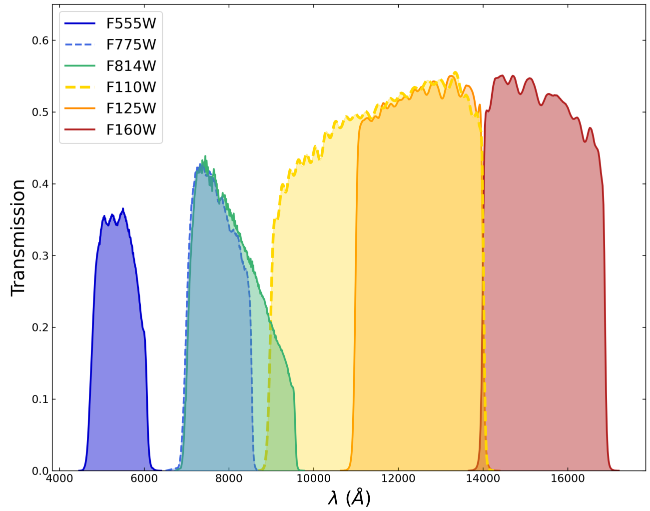

Figure 3. Transmission as a function of wavelength for HST ACS WFC (F555W, F775W, F814W) and WFC3 IR (F110W, F125W, F160W) filters. Filters traced with solid lines were observed in N159, and filters traced with dashed lines are only obtained as part of the the R136 training set.

the F814W and F160W coverage for N159W N159E, and N159S.

Photometry was performed using DOLPHOT (Dolphin 2000). In this work, we only make use of F555W and F814W for classification purposes; we will include F125W and F160W photometry in future analysis. We cross-matched our photometry with a maximum distance between F555W and F814W candidates of 0.2", and additionally required that both bands pass the following requirements:

- Magnitude $<$ 90 mag
- SNR $>$ 10
- Crowding $\leq$ 0.48 (Dalcanton et al. 2012)
- Shape $>$ -0.6
- Color $m_{555} - m_{814} > -0.2$

This yielded a collection of 114000 candidate sources: 33300 in N159W, 36000 in N159E, and 45000 in N159S.

2.2. *Bright Stars in the Upper Main Sequence*

In this work, we follow the methods of Ksoll et al. (2018) to estimate extinction affecting PMS stars by measuring the departure of UMS stars from the zero age main sequence (ZAMS). Unfortunately, our HST photometry is saturated at the brightest magnitudes ($\lesssim$ 18 mag), so we supplement our catalog's missing UMS with sources from Zaritsky et al. (2004, hereafter ZH04)'s photometric survey of the LMC. We select stars in the ZH04 catalog that fall within 15" of an HST candidate, and convert their Johnson V & I photometry to F555W and F814W filters using the relationships derived by Sirianni et al. (2005),

$$m_{555} = m_V - (25.719 - 0.008c_{VI} + 0.043c_{VI}^2) + m_{zpt,555}$$
$$m_{814} = m_I - (25.489 + 0.041c_{VI} - 0.093c_{VI}^2) + m_{zpt,814}, \quad (1)$$

with color $c_{VI} = m_V - m_I$ and synthetic WFC zeropoint magnitudes $m_{zpt,555} = 25.724$ and $m_{zpt,814} = 25.501$. As with the HST photometry, we required $m_{555} - m_{814} > -0.2$. This yields a total of 4320 sources. We plot the combined CMD and spatial distributions of the HST and ZH04 samples in Figure 5.

We used the Pisa PMS evolutionary models (Tognelli et al. 2011) to generate a 20 Myr ZAMS isochrone by way of the IDL program TA-DA (Da Rio & Robberto 2012), which created synthetic HST photometry informed by the Kurucz (1993) atmospheric models. The models were calculated for Z=0.008, Y=0.254, and mixing length parameter $\alpha = 1.2$. We assumed a Galactic foreground $A_V = 0.2$ mag, $R_V = 3.1$, and a distance modulus of $\mu = 18.48$ mag. The isochrone was only calculated for $m_{555} - m_{814} \gtrsim -0.16$ mag, so we extrapolate to the minimum color in our sample using a simple cubic fit; this method will be improved in the future. The resulting ZAMS isochrone is also shown in Figure 5. We define the UMS by identifying stars that fall between the right side of the ZAMS and $c_{555,814} < 0.85$ mag (to exclude the red clump feature). This yields a sample of 2000 UMS stars, which we use in §3.2 to estimate extinction.

2.3. *Training Set for Machine Learning Model*

We teach our machine learning model to identify probable PMS stars using a training set generated by (Ksoll et al. 2018) in the LMC superstar cluster R136. The R136 training set contains 10,440 entries with fit probabilities of PMS membership P(PMS). These example stars are mainly located in the LMS and PMS, but also include 1,200 entries with artificial P(PMS) = 0 to account for other contaminants. The R136 training set includes estimated extinctions generated from analysis of the UMS in a similar process to our methods in §3.2 that we use to deredden the photometry. We define PMS members as having P(PMS) $\geq$ 0.85.

The training set includes photometry using HST filters F555W, F775W, F110W, and F160W. We plot the transmission curves of F775W and F110W, which in our data are replaced by F814W and F125W, in Figure 3. Both F775W and F110W are nearly entirely contained by their matching filters in our data (for F775W, 6800.98 Å $\leq \lambda \leq$ 8627.17 Å with peak λ_{pk} = 7380 Å,

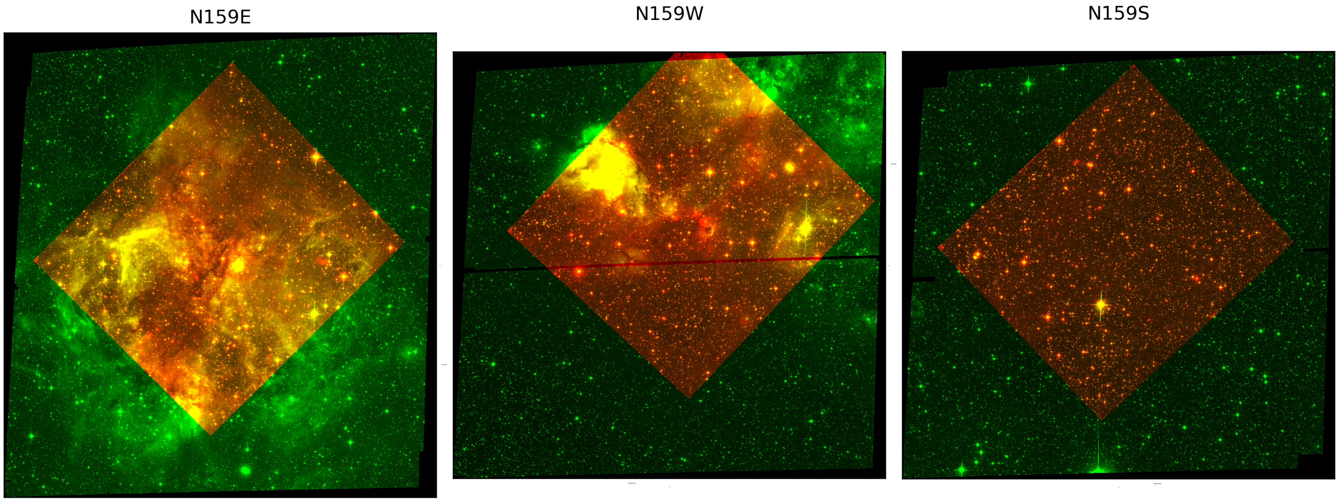

Figure 4. HST images of (from left) N159E, N159W, and N159S, with [R, G] = [F160W, F814W].

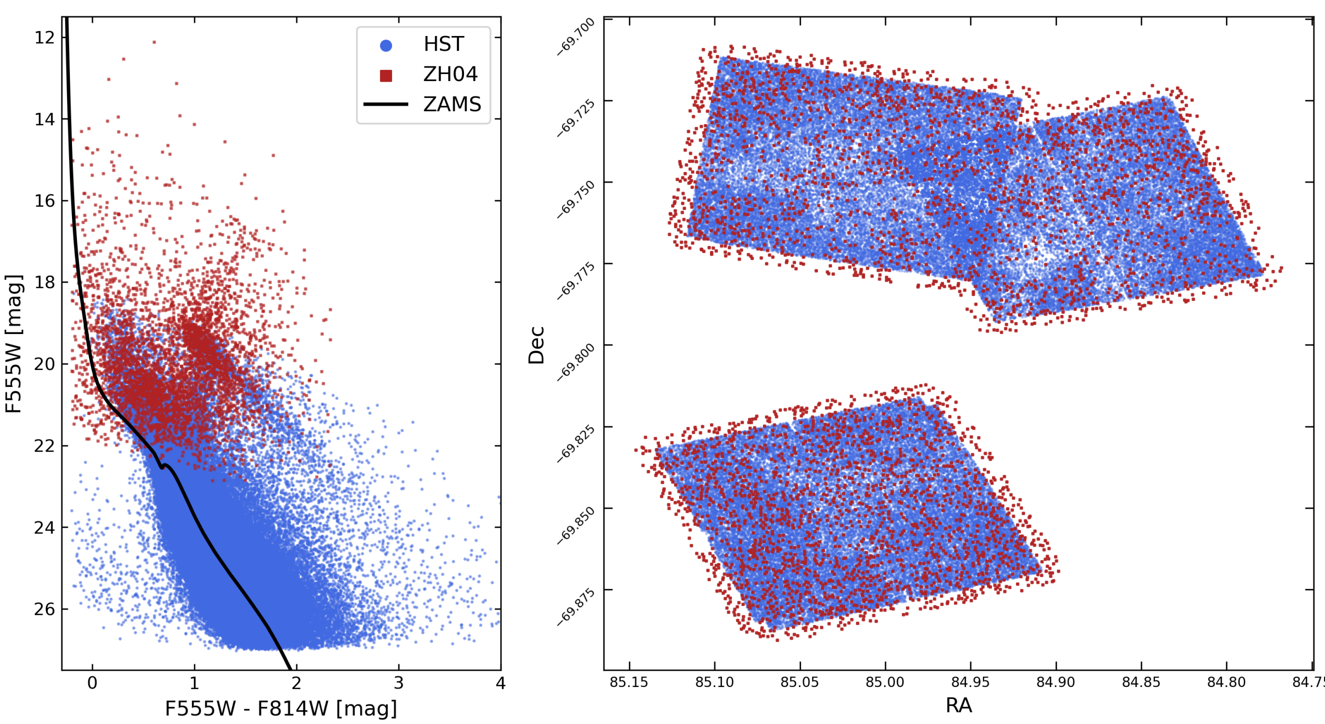

Figure 5. *Left:* F555W as a function of F555W - F814W for complete N159 HST sample (blue circles) and ZH04 sample (red squares). The assumed ZAMS is shown by the black isochrone. *Right:* Spatial distribution of HST vs ZH04 samples, with points colored as in left.

vs for F814W 6867.81 Å $\leq \lambda \leq$ 9626.08 Å with peak λ_{pk} = 7440 Å). In the future we will use synthetic photometry to extrapolate their equivalent F814W and F125W magnitudes. Future work will also explore the effect of different training sets on overall classification results in N159. At present, we use F775W as a direct substitute for F814W to train our classification model for N159.

3. EXTINCTION ESTIMATION

3.1. *RANSAC-ing the Red Clump*

The red clump (RC) is populated by post-MS stars burning helium in their cores. Because the luminosities of these stars are relatively constant with age, the RC is frequently used to assess reddening and distances. Inspection of the N159 regions CMDs reveals clear RC features. If no extinction were present, the RC would be a roughly circular feature branching off of the MS. The RC appears very elongated in N159, which suggests that significant differential reddening from dust and molecular clouds is affecting our photometry. This elongation of the RC leaves a roughly linear feature, and by measuring its slope we can estimate the reddening vector for each region in N159.

Following the methods of Ksoll et al. (2021), we use the RANSAC (RANdom SAmple Consensus Fischler & Bolles 1981) algorithm to fit reddening vector slopes. RANSAC is an iterative fitting algorithm that can robustly derive the slope of and determine probable membership of points to linear features in the presence of outliers.

For each region's CMD, we select points that fall within $[0.85 \leq m_{555} - m_{814} \leq 2.5,\ 18.5 \leq m_{555} \leq 22]$ as candidate RC points. Assuming a distance modulus μ = 18.48 for the LMC (Pietrzyński et al. 2019), we expect the RC's center to fall at V - I = 0.93, V = 19.02 (Nataf et al. 2021); this translates to $m_{555} - m_{814}$ = 0.925, m_{555} = 19.07 via Equation 1. We require the RANSAC fitter to fit lines that pass through this expected RC locus to reduce uncertainties and allow easier comparison of slopes.

We record slopes fit by the RANSAC algorithm over 2000 runs. We also calculate for each point a probability of RC membership, P(RC), as the fraction of times the point is identified as being an inlier vs outlier to the RANSAC-identified linear feature. Figure 6 shows the resulting probabilities and median slopes in CMD space for each region as well as the full sample.

Figure 7 shows kernel density estimations (KDEs) of the distribution of RANSAC slopes for N159W, N159E, and N159S. There are significant variations in slope between regions, with N159S having the lowest slopes followed by N159W, and N159E having the highest. Each pair of regions distributions are significantly different ($p \ll 0.001$) as assessed by Kolmogorov-Smirnov (K-S) tests.

N159E, N159W, N159S, and the full sample have median slopes of $R_{555-814}(555)$ =2.76, 2.57, 2.44, and 2.53, respectively. Using the same methods, Ksoll et al. (2021) derived a steeper $R_{555-814}(555) = 2.8 \pm 0.3$ in the LMC star-forming region N44. This suggests that reddening in N44 and N159E is "grayer" than in N159W and N159S.

3.2. *Upper Main Sequence Extinction Maps*

The upper main sequence (UMS) is populated by young, massive stars that can be reasonably assumed to be spatially correlated with any adjacent PMS populations. This makes the UMS a valuable way to estimate extinction for other stars in our sample and correct our CMD for differential reddening, as informed by our RC slope fitting.

We estimate extinction for each region individually (N159W, N159E, N159S) and as part of the full sample of all stars. We select UMS stars in the ZH04 sample as described in §2.2, and calculate the distance to the ZAMS along the slope of the RANSAC-fitted reddening vector, A_{555}, for that region.

We then perform a K-nearest neighbors regression for each star in the HST sample to estimate their extinctions from the UMS. We take K=10 and assign weights to the UMS stars as (Ksoll et al. 2018),

$$w_i = \frac{1}{d_i^2 + \epsilon^2} \frac{1}{\sum_{n=1}^{N} \frac{1}{d_i^2 + \epsilon^2}}, \tag{2}$$

where d_i is the distance to the ith nearest UMS stars, and ϵ is a smoothing parameter we set equal to 1. We deredden each star's photometry for its inferred A_{555} using the reddening law derived in 30 Doradus by De Marchi et al. (2016).

The extinction maps generated from this process are shown in Figure 9 with a contour of CO emission overlaid to mark the locations of the GMCs. We also present KDEs of the distributions of A_{555} for stars in each region in Figure 8. N159E and N159S have higher average extinctions than N159W, which may reflect the GMCs extending over more of the HST coverage in those regions. Extinction maps generated for individual regions tend to agree very well with the corresponding subsection of the maps derived for the full HST sample, despite different reddening vectors being used. We only consider the full HST sample for the remainder of this work. In the future, we will perform our classification methods on

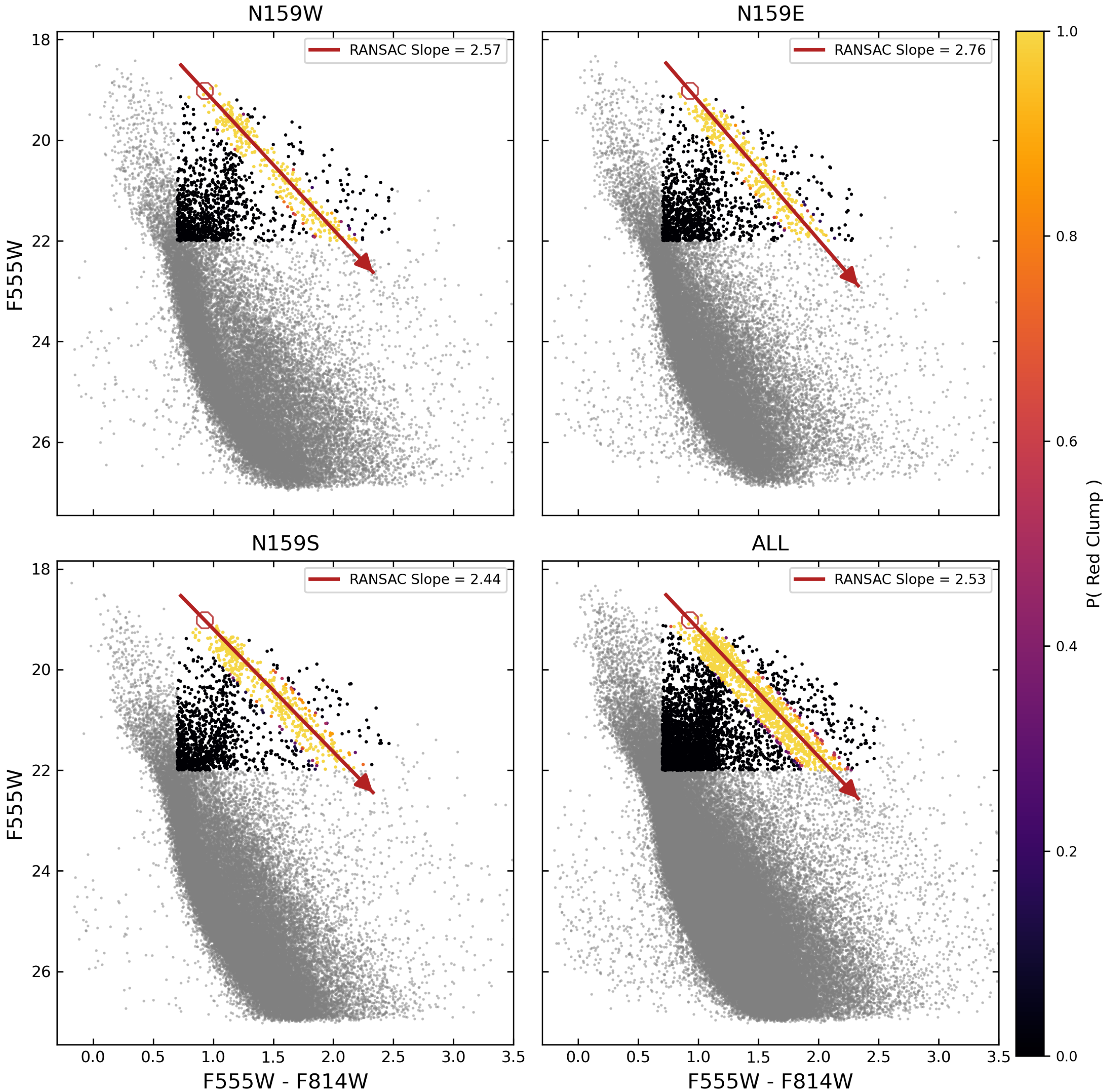

Figure 6. Probability of red clump membership as assessed by the RANSAC algorithm is shown for CMDs of N159W (top left), N159E (top right), N159S (bottom left), and the full HST sample (bottom right). In each subplot, the open red circle marks the expected locus of the red clump and the red arrow shows the median RANSAC-derived red clump slope.

each individually-dereddened region as well to see how large of an effect it might have.

4. IDENTIFYING PMS STARS WITH SUPPORT VECTOR MACHINES

4.1. *Training the Support Vector Machine*

A support vector machine (SVM) identifies the most effective boundaries between classes of a labeled training set (here, PMS vs. non-PMS), and applies these

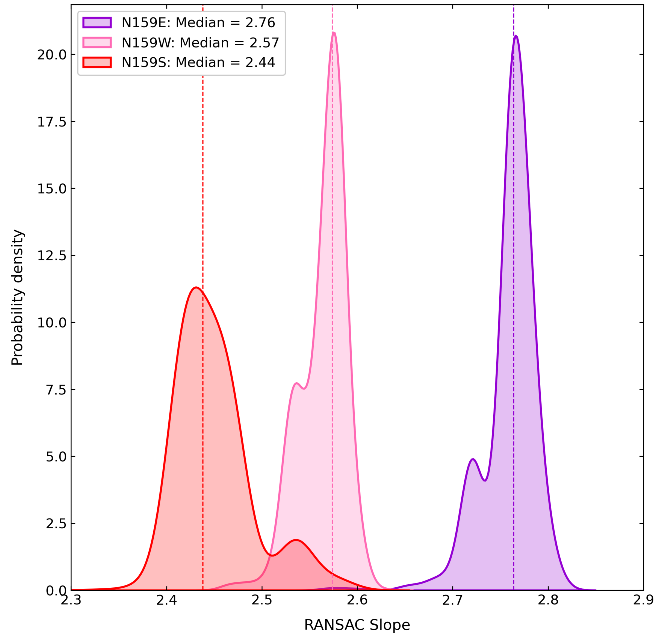

Figure 7. Kernel density estimations of the red clump slopes estimated by the RANSAC algorithm are shown. N159S's KDE is red, N159W's is pink, and N159E's is purple.

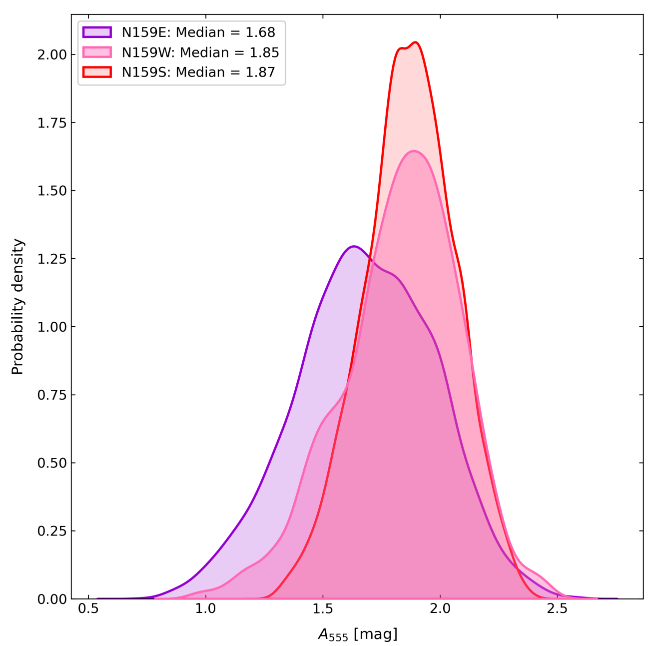

Figure 8. KDEs of extinctions A_{555} inferred via K-nearest neighbors regression for HST stars. N159S's KDE is red, N159W's is pink, and N159E's is purple.

boundaries to the target sources to generate classification probabilities. SVMs are most useful when data are not linearly separable, as is the case in our CMD space. The SVM projects a feature space to a higher dimension

Table 1. Normalized Confusion Matrix

	True PMS	True non-PMS
Predicted PMS	**0.921**	0.041
Predicted non-PMS	0.079	**0.959**

and finds the optimal hyperplane that completely separates classes in the original, lower-dimensional space. The optimal hyperplane is defined as one where the margin (minimum distance of all points to a hyperplane) is largest (i.e., the separation between categories is the largest). This is the maximum margin hyperplane (MMH). Points that are closest to the MMH are referred to as the support vectors - since if they changed position, the hyperplane would need to change in response to maintain the maximum margin. Once obtained, this system is referred to as a support vector clasifier (SVC).

Transforming features to a higher dimensional space requires the choice of a kernel function (e.g., a polynomial) to dictate the nature of the transformation. Using a kernel reduces the computational cost of transforming data to higher dimensions. A SVC using a non-linear kernel function is referred to as a SVM.

SVMs have a variety of hyperparameters which need to be finely tuned to yield an optimal classification method. It is desirable to create a soft margin between classes, which allows some degree of error on what side of the margin different classes fall, since it is in most cases unrealistic to expect that a perfect margin exists. The degree of softness is specified by a cost parameter C that specifies how much error is allowed; larger Cs increase the penalty placed on the SVM when a point is misclassified. Appendix A3 of Ksoll et al. (2018) provides a more detailed overview of the SVM classification and hyperparameter tuning process to which we direct the interested reader. There are several key steps in training an SVM, which we describe here:

(i) *Training/test Split:* We randomly split the R136 catalog into 70% training and 30% testing sets. We assigned points in the training set with an input P(PMS) $>$ 0.85 as belonging to the binary PMS class, and P(PMS) $<$ 0.85 as to the non-PMS class. Both sets are shown in Figure 10. We also scale the considered features (F555W and F775W) to a standard 0–1 range.

(ii) *Hyperparameter Tuning:* We choose a Radial Basis Function (RBF) as our kernel. We selected hyperparameters via a stratified grid search, to reduce variance and offset the imbalance in the number of

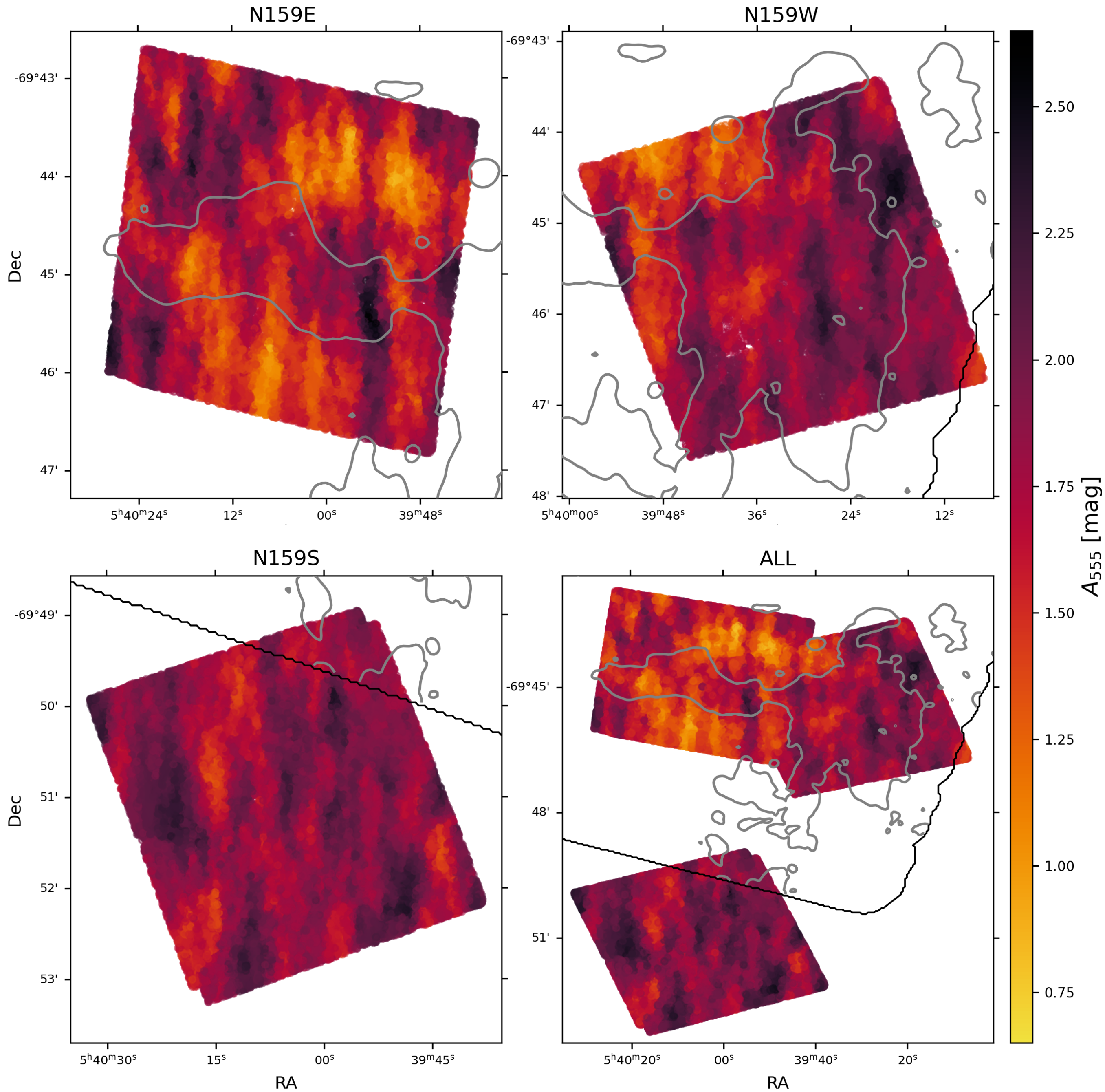

Figure 9. Extinction A_{555} maps for HST stars in N159W (top left), N159E (top right), N159S (bottom left), and the full sample (bottom right). A contour of CO in N159W and N159E at 5 K is overlaid in grey, with the boundaries of the CO map being marked in black.

PMS vs non-PMS stars in the R136 catalog. We found an optimal $C \simeq 20$ and $\gamma \simeq 10$ (where γ controls the curvature of the decision boundary).

(iii) *Fit SVM:* We fit an SVM to the training set using our optimal hyperparameters. The resulting decision boundary is shown in the left panel of Figure 12. We then apply the trained SVM to the reserved test set.

(iv) *Probability Estimation:* We predict probabilities of PMS membership through 5-fold cross validation.

(v) *Performance Validation:* Our classifier achieved 94.9% accuracy on correctly identifying object classes in the test set. Appendix A4 of Ksoll et al.

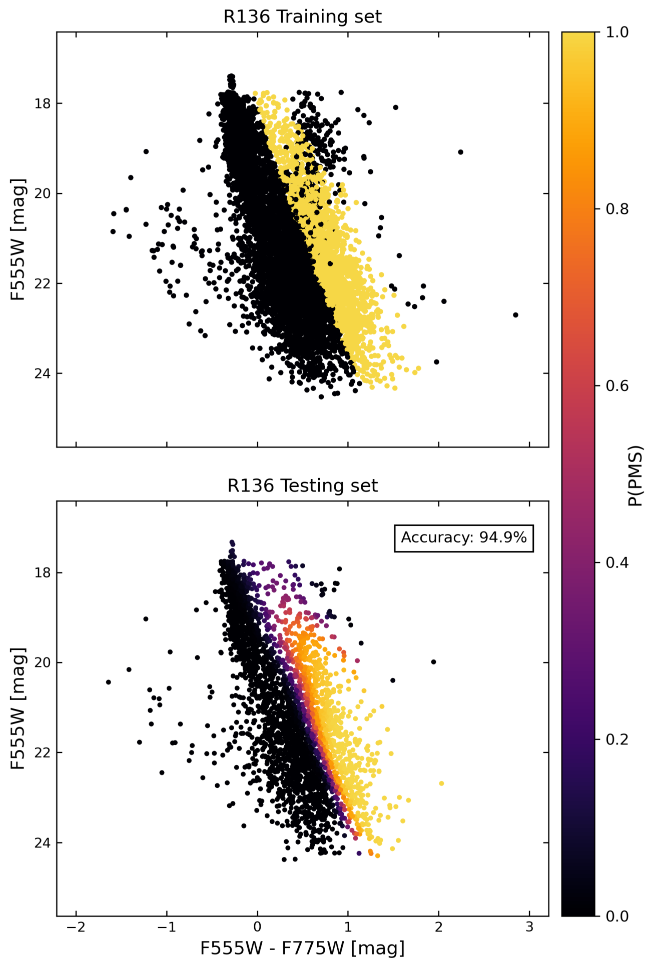

Figure 10. *Top:* Dereddened CMD of randomly selected 70% subset of R136 dataset (the training set for the SVM). Stars with P(PMS) ≥ 0.85 are in yellow, and P(PMS) < 0.85 are in black. *Bottom:* Dereddened CMD of subset of R136 dataset not selected for top panel (the test set), colored by P(PMS) assigned by the trained SVM.

(2018) describes the meaning and interpretation of common metrics of classifier performance. As preliminary analysis, we describe the results of analyzing confusion matrices and ROC curves for our fit SVM.

The normalized confusion matrix (Table 1) gives an overview of the SVM's performance by summarizing the number of samples in each class that were correctly vs incorrectly classified. The upper left and lower right squares show the fraction of true positives (correctly classified PMS stars, P(PMS) ≥ 0.95) and true negatives (non-PMS stars, P(PMS) < 0.95), respectively. The classifier performs slightly better with identifying true negatives (95.5% of non-PMS stars) than true positives (92.1% of PMS stars), although both are excellent

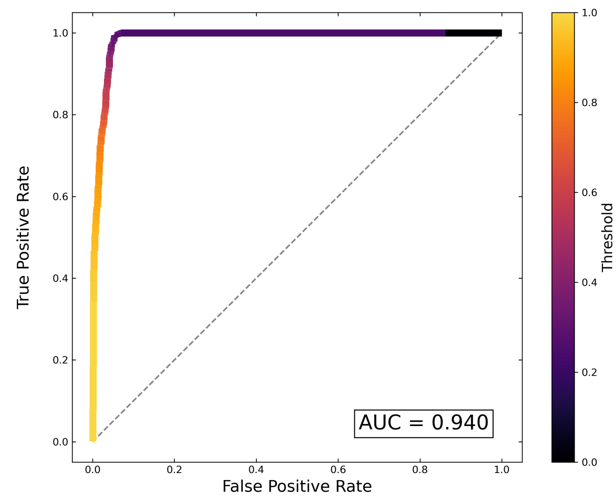

Figure 11. ROC curve for the trained SVM, showing True Positive Rate as a function of False Positive Rate while classification threshold varies (colored line). The gray, dashed 1:1 diagonal line shows the performance of a random classifier.

overall. The lower left and upper right squares show the fraction of false negatives (PMS stars classified as non-PMS stars) and false positives (non-PMS stars classified as PMS stars). The classifier is marginally worse in creating false negatives (7.9% of PMS stars) than false positives (4.1% of non-PMS stars).

The receiver operating characteristic (ROC) curve presented in Figure 11 shows the true positive rate as a function of false positive rate while the discrimination threshold (minimum P(PMS) to be considered as a PMS star) varies. A truly random classifier would have a 1:1 ROC curve, since it classifies with a 50:50 random probability; this is shown in the plot by the dotted gray line. Our SVM significantly outperforms a random classifier, as evidenced by its steep rise in true positive rate at low false positive rates. The area under the curve (AUC) of each classifier can range between 0-1, with 1 being a perfect classifier and 0 being a completely incorrect classifier, and confirms that our SVM outperforms a random classifier with an AUC of 0.94 (vs. 0.5 for random).

4.2. *Classifying PMS Stars in N159*

We then run our trained SVM on our full dereddened and scaled N159 sample. We plot the resulting CMD with points colored by P(PMS) in Figure 12. We also show the SVM's decision boundaries in the scaled m_{555} vs. m_{814} feature space. Our model performs well via visual inspection; the MS and RC have P(PMS) $\simeq 0$,

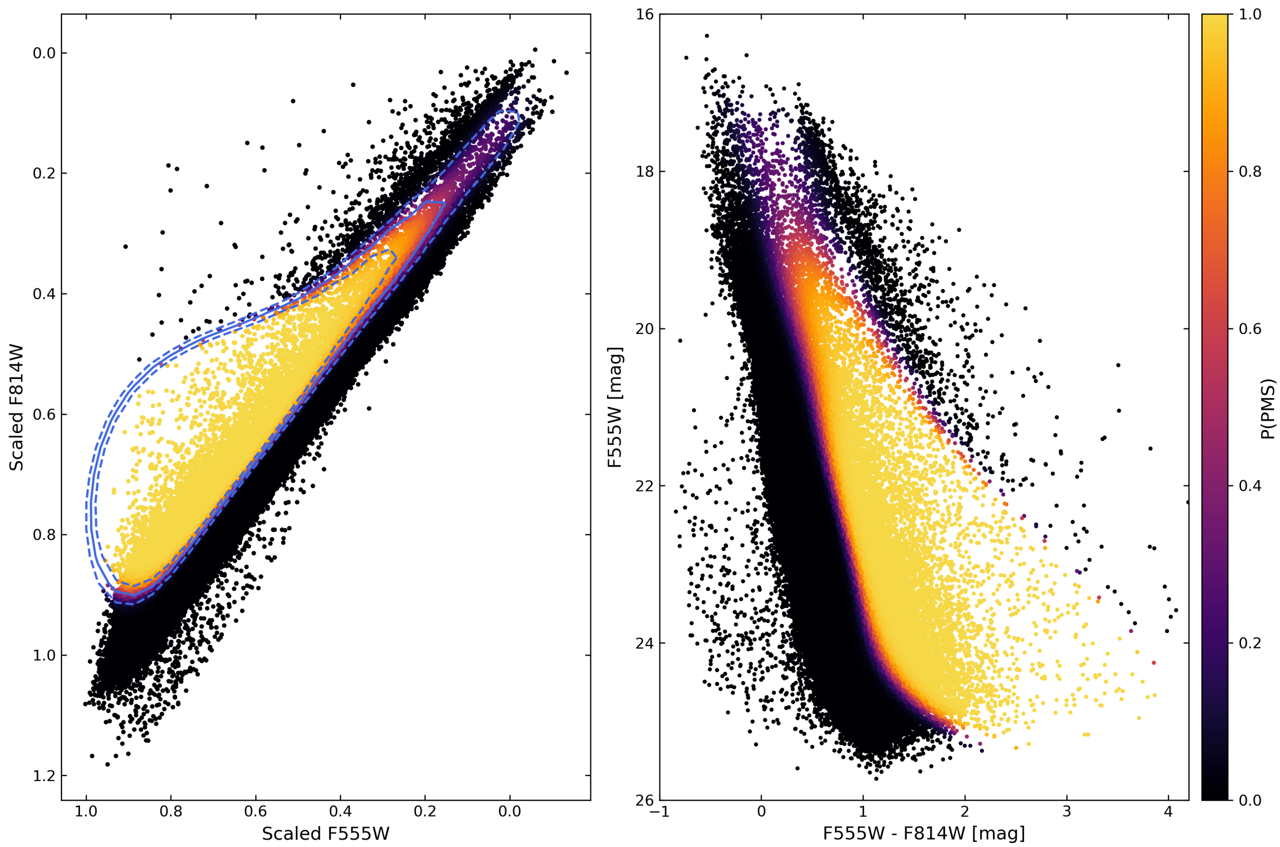

Figure 12. *Left*: Scaled F814W magnitude vs scaled F555W magnitude for full HST sample, with points colored by SVM-derived P(PMS). The SVM's decision boundary is shown by the blue curves. *Right:* Dereddened CMD, with colors as in left.

and an intermediate fuzzy boundary between these features and PMS stars exists. $\sim$9% (9900 stars) of the full HST sample are compelling PMS candidates (P(PMS) ≥ 0.95).

In the left panel of Figure 13, we plot a map of the spatial distribution of P(PMS) for all HST stars, and a map of only those stars with P(PMS) > 0.95 in the right panel. We see that PMS candidates are generally clustered around the high CO regions in N159E and N159W.

5. DISCUSSION AND FUTURE WORK

In this work we have identified candidate pre-main-sequence stars in the LMC star-forming region N159 through a supervised machine learning approach while accounting for differential reddening. We are able to achieve 95% classification accuracy, and our results suggest that the PMS star candidates are closely associated with GMCs in the region.

Future work will experiment with other machine learning methods (e.g., random forests and neural networks), training sets obtained in other LMC regions, extinction estimation methods, and more. With a reliable map of probable PMS stars in N159, we will be able to more directly compare their properties and spatial positions to the three GMCs and to other stellar populations in the region.

Since solar-mass stars can remain PMS for up to 50 Myr, the PMS stars we identify will allow us to probe N159's star formation history for indications of the timescale over which feedback from star formation acts. We will uncover the recent and historical activity of each individual GMC and test the previously observed north-south trends in star formation activity in N159 corresponding to the transition between 30 Doradus and the molecular ridge. This continued work will be the most comprehensive study of star formation in N159 yet performed, and yield insights that can be generalized to the process of star formation in other moderately low-metallicity environments.

Software: Astropy (Astropy Collaboration et al. 2013, 2018); Cmasher (van der Velden 2020); DOLPHOT (Dolphin 2016); Glue (Beaumont et al. 2015; Robitaille et al. 2019); Matplotlib (Hunter 2007); Numpy (Harris et al. 2020); Pandas (McKinney 2010); Seaborn (Waskom 2021); scikit-learn (Pedregosa et al. 2011); Scipy (Virtanen et al. 2020); TA-DA (Da Rio & Robberto 2012)

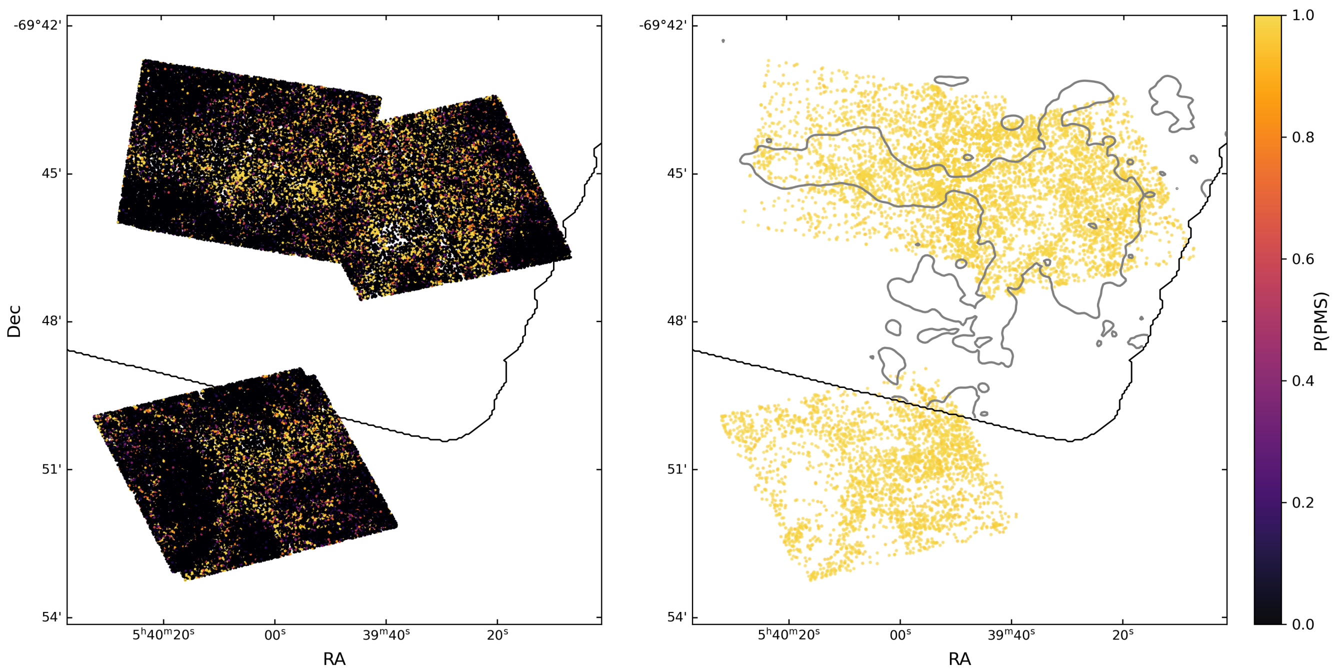

Figure 13. *Left:* Spatial distribution of P(PMS) for all HST stars. *Right:* Spatial distribution of HST stars with P(PMS) > 0.95. CO contour at 5 K is overlaid.

Facilities: ALMA, HST

REFERENCES

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123

Beaumont, C., Goodman, A., & Greenfield, P. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 495, Astronomical Data Analysis Software an Systems XXIV (ADASS XXIV), ed. A. R. Taylor & E. Rosolowsky, 101

Chen, C. H. R., Indebetouw, R., Chu, Y.-H., et al. 2010, ApJ, 721, 1206

Da Rio, N., & Robberto, M. 2012, AJ, 144, 176

Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012, ApJS, 200, 18

De Marchi, G., Panagia, N., Sabbi, E., et al. 2016, MNRAS, 455, 4373

Dolphin, A. 2016, DOLPHOT: Stellar photometry

Dolphin, A. E. 2000, PASP, 112, 1383

Fischler, M. A., & Bolles, R. C. 1981, Commun. ACM, 24, 381–395

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90

Indebetouw, R., Whitney, B. A., Kawamura, A., et al. 2008, AJ, 136, 1442

Ksoll, V. F., Gouliermis, D. A., Klessen, R. S., et al. 2018, MNRAS, arXiv: 1805.07157

Ksoll, V. F., Gouliermis, D., Sabbi, E., et al. 2021, AJ, 161, 256

Kurucz, R. L. 1993, SYNTHE spectrum synthesis programs and line data

McKinney, W. 2010, in Proc. 9th Python in Science Conf., ed. Stéfan van der Walt & Jarrod Millman, 56

Mizuno, N., Yamaguchi, R., Mizuno, A., et al. 2001, PASJ, 53, 971

Nataf, D. M., Cassisi, S., Casagrande, L., Yuan, W., & Riess, A. G. 2021, ApJ, 910, 121

Nayak, O., Meixner, M., Fukui, Y., et al. 2018, ApJ, 854, 154

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pietrzyński, G., Graczyk, D., Gallenne, A., et al. 2019, Nature, 567, 200

Robitaille, T., Beaumont, C., Qian, P., Borkin, M., & Goodman, A. 2019, glueviz v0.15.2: multidimensional data exploration

Russell, S. C., & Dopita, M. A. 1992, ApJ, 384, 508

Schaefer, B. E. 2008, AJ, 135, 112

Sirianni, M., Jee, M. J., Benítez, N., et al. 2005, PASP, 117, 1049

Tognelli, E., Prada Moroni, P. G., & Degl'Innocenti, S. 2011, A&A, 533, A109

van der Velden, E. 2020, The Journal of Open Source Software, 5, 2004

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Waskom, M. L. 2021, Journal of Open Source Software, 6, 3021

Zaritsky, D., Harris, J., Thompson, I. B., & Grebel, E. K. 2004, AJ, 128, 1606

5

Conclusions

The interstellar medium is the host to and catalyst for many of the most important physical processes driving the evolution of our universe. In this thesis, I have explored variations in the physics of star formation and the ISM in two of our nearest low-metallicity cosmic neighbors — the Small and Large Magellanic Clouds.

In Chapter 2, I examined the theoretical effect of "CO-dark" molecular gas on observable molecular cloud properties in low-metallicity and high radiation environments. I derived corrections that can be applied to empirical clump properties to account for this missing material, and demonstrated that correcting for CO-dark gas is critical for accurately comparing the dynamical state and evolution of molecular clouds across diverse environments.

In Chapter 3, I analyzed the recent evolutionary past of the young, low-metallicity star cluster NGC 602 in the SMC. I identified molecular clumps traced by CO emission, and studied their relationship with nearby young stellar objects (YSOs) and pre-main-sequence (PMS) stars. Contrary to expectations, I did not find evidence for a triggered star formation scenario among the youngest ($\lesssim$2 Myr) stellar generations in NGC 602, and instead concluded that a sequential star formation process in which NGC 602 did not directly cause recent star formation in the region is likely.

Finally, I explored machine-learning-based methods of identifying PMS stars in Chapter 4. I used the LMC star-forming region N159 as a pilot program to develop these methods. I estimated reddening laws in the region by fitting the slope of the red clump population, before deriving an extinction map of the region from the upper-main-sequence. I then applied a support vector machine (SVM) classifier to our extinction-corrected photometry to identify a sample of 10,000 probable PMS candidates associated with giant molecular clouds in N159.

It is clear that the properties of the ISM and progression of star formation in the SMC and LMC vary dramatically from their equivalents in our own Milky Way. This thesis investigates just a few of these variations. The advent of the next generation of high resolution interferometers and large surveys creates an enormous range of future opportunities in this area, and promises no shortage of further discoveries in our understanding of the physics of the interstellar medium.