



**HAL**  
open science

## H II regions and high-mass starless clump candidates

S. Zhang, Annie Zavagno, J. Yuan, H. Liu, M. Figueira, D. Russeil, F. Schuller, K. A. Marsh, Y. Wu

► **To cite this version:**

S. Zhang, Annie Zavagno, J. Yuan, H. Liu, M. Figueira, et al.. H II regions and high-mass starless clump candidates: I. Catalogs and properties. *Astronomy and Astrophysics - A&A*, 2020, 637, pp.A40. 10.1051/0004-6361/201936792 . hal-02903854

**HAL Id: hal-02903854**

**<https://hal.science/hal-02903854>**

Submitted on 21 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# H II regions and high-mass starless clump candidates

## I. Catalogs and properties ★,★★

S. Zhang (张骜驹)<sup>1</sup>, A. Zavagno<sup>1,2</sup>, J. Yuan<sup>3</sup>, H. Liu<sup>4,5</sup>, M. Figueira<sup>6</sup>, D. Russeil<sup>1</sup>, F. Schuller<sup>7</sup>,  
K. A. Marsh<sup>8</sup>, and Y. Wu<sup>9</sup>

<sup>1</sup> Aix Marseille Univ., CNRS, CNES, LAM, Marseille, France  
e-mail: [sj.zhang@lam.fr](mailto:sj.zhang@lam.fr)

<sup>2</sup> Institut Universitaire de France (IUF), Paris, France

<sup>3</sup> National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100012, PR China

<sup>4</sup> CASSACA, China-Chile Joint Center for Astronomy, Camino El Observatorio #1515, Las Condes, Santiago, Chile

<sup>5</sup> Departamento de Astronomía, Universidad de Concepción, Av. Esteban Iturra s/n, Distrito Universitario, 160-C, Chile

<sup>6</sup> National Centre for Nuclear Research, ul. Pasteura 7, 02-093, Warszawa, Poland

<sup>7</sup> Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

<sup>8</sup> Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, California 91125, USA

<sup>9</sup> Department of Astronomy, Peking University, 100871 Beijing, PR China

Received 26 September 2019 / Accepted 24 February 2020

### ABSTRACT

**Context.** The role of ionization feedback on high-mass ( $>8 M_{\odot}$ ) star formation is still highly debated. Questions remain concerning the presence of nearby H II regions changes the properties of early high-mass star formation and whether H II regions promote or inhibit the formation of high-mass stars.

**Aims.** To characterize the role of H II regions on the formation of high-mass stars, we study the properties of a sample of candidates high-mass starless clumps (HMSCs), of which about 90% have masses larger than  $100 M_{\odot}$ . These high-mass objects probably represent the earliest stages of high-mass star formation; we search if (and how) their properties are modified by the presence of an H II region.

**Methods.** We took advantage of the recently published catalog of HMSC candidates. By cross matching the HMSCs and H II regions, we classified HMSCs into three categories: (1) the HMSCs associated with H II regions both in the position in the projected plane of the sky and in velocity; (2) HMSCs associated in the plane of the sky, but not in velocity; and (3) HMSCs far away from any H II regions in the projected sky plane. We carried out comparisons between associated and nonassociated HMSCs based on statistical analyses of multiwavelength data from infrared to radio.

**Results.** We show that there are systematic differences of the properties of HMSCs in different environments. Statistical analyses suggest that HMSCs associated with H II regions are warmer, more luminous, more centrally-peaked and turbulent. We also clearly show, for the first time, that the ratio of bolometric luminosity to envelope mass of HMSCs ( $L/M$ ) could not be a reliable evolutionary probe for early massive star formation due to the external heating effects of the H II regions.

**Conclusions.** We show HMSCs associated with H II regions present statistically significant differences from HMSCs far away from H II regions, especially for dust temperature and  $L/M$ . More centrally peaked and turbulent properties of HMSCs associated with H II regions may promote the formation of high-mass stars by limiting fragmentation. High-resolution interferometric surveys toward HMSCs are crucial to reveal how H II regions impact the star formation process inside HMSCs.

**Key words.** stars: formation – H II regions – submillimeter: ISM

## 1. Introduction

The formation of high-mass ( $>8 M_{\odot}$ ) stars is still a mystery as a consequence of their rapid evolution, larger distance, and violent feedback, which are all deeply embedded in molecular clouds. With the benefits of higher resolution and sensitivity from millimeter interferometers such as Atacama Large Millimeter/Submillimeter Array (ALMA), the details happening at the very early evolutionary stages of high-mass star formation (HMSF) have been investigated in recent years. Two popular models for HMSF have been discussed widely: (1) monolithic

collapse (McKee & Tan 2003; Krumholz et al. 2009; Csengeri et al. 2018; Sanna et al. 2019; Motogi et al. 2019) and (2) competitive accretion (Bonnell et al. 2007; Cyganowski et al. 2017; Fontani et al. 2018). It is possible that these models are not strongly and mutually incompatible as described by many current studies; these two models of HMSF are expected to be merged into a unified and consistent star formation model with the progress of observations and theories (Schilke 2015). The new evolutionary scenario proposed by Motte et al. (2018) suggests that the large-scale (0.1–1 pc) gas reservoirs, known as starless massive dense cores (MDCs) or starless clumps, could replace the high-mass analogs of prestellar cores (about 0.02 pc). These mass reservoirs could concentrate their mass into high-mass cores at the same time when stellar embryos are accreting.

\* Full Table 4 and additional data are only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](https://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/637/A40>

\*\* This publication is based on the SEDIGISM and PPMAP data.

The complexities of HMSF result from not only the controversially basic pictures of formation mechanism but also from the environmental impacts on the formation process. The powerful UV radiation of massive stars ionizes hydrogen to form H II regions. Because high-mass stars form in the densest parts of molecular clouds, H II regions are expected to become a typical environment of HMSF. Previous studies found that at least 30% of HMSF in the Galaxy are observed at the edges of H II regions (Deharveng et al. 2010; Kendrew et al. 2016; Palmeirim et al. 2017).

The possible impacts of ionization radiation on star formation have been investigated in many publications. H II regions could trigger star formation with several mechanisms, mainly depending on the density and configuration of the surrounding medium (see Sect. 5 in Deharveng et al. 2010), such as the collect and collapse (C&C) mechanism in IR dust bubbles (Zavagno et al. 2010; Samal et al. 2014; Duronea et al. 2017) or the radiation-driven implosion (RDI) mechanism in pillars or bright rimmed clouds (Bertoldi 1989; Imai et al. 2017; Marshall & Kerton 2019).

The interferometric studies of HMSF generally focus on physical and/or chemical phenomena inside the massive clumps or in a single field of view. This could produce a blind zone for the possible interactions between massive clumps, HMSF, and their various environments. For example, Csengeri et al. (2017a) used ALMA to survey fragmentation at the early evolutionary stage of IR-quiet massive clumps and ignored possible environmental factors, even though environmental factors such as the presence of nearby H II region could impact the fragmentation process (Busquet et al. 2016).

In a series of forthcoming papers, we try to systematically explore how the environment of the H II region changes HMSF, particularly for the HMSF at the very early evolutionary stages. High-mass starless clumps (HMSCs), which are very cold (dust temperature of around 10 to 20 K) and massive (from a few hundreds to several thousands of solar mass at a typical scale of 0.1–1 pc), are thought to be the possible precursors of high-mass stars. These sources are likely at the earliest stage of HMSF (quiescent prestellar stage), and only less than 10% of the clumps with the ability to probably form massive stars are at this stage, indicating their short lifetime (Urquhart et al. 2018). Observationally, HMSCs are dark against their background environment from optical to mid-infrared (mid-IR) owing to the absence of the embedded protostellar objects.

Systematic studies of HMSCs under the effect of H II regions are very crucial to characterize the influence of the H II regions and how the influence impacts very early HMSF and the associated gas dynamics. We investigate the impacts of H II regions on a large sample of 463 HMSC candidates recently selected by Yuan et al. (2017) (written briefly as Y17 in the following) in the inner Galactic plane. As in the first paper, we cross match HMSCs and H II regions based on their mutual correlation in the plane of sky to ascertain their association and then investigate the differences of the physical properties between the HMSCs associated with and those not associated with H II regions.

This paper is organized as follows: the sample selection and archival data acquisition are described in Sect. 2. Section 3 shows how we determine the basic parameters of HMSCs with updated data, followed by cross matching of H II regions and HMSCs in Sect. 4. In Sects. 5–8, we explore and compare the properties of different HMSCs. In Sect. 9, we discuss the results to shed light on the role of ionization feedback on modifying the physical properties of HMSCs. The conclusions are finally presented in Sect. 10.

**Table 1.** Star formation indicators.

Search	IR point sources	Maser	Others
S16	Hi-GAL 70 $\mu\text{m}$ 2–24 $\mu\text{m}$ YSOs	H <sub>2</sub> O masers CH <sub>3</sub> OH masers	UCH II regions
	Hi-GAL 70 $\mu\text{m}$ MIPSGAL sources GLIMPSE YSOs <b>IR sources</b> <b>YSOs (C.)</b> <sup>(b)</sup> <b>PMS (C.)</b> <b>T Tau (C.)</b> <b>HAEBE (C.)</b>	<b>Masers</b> <sup>(a)</sup>	<b>H II regions</b> extended 24 $\mu\text{m}$ <b>cm radio sources</b> <b>HH objects</b> <b>outflows (C.)</b>
Y17			

**Notes.** <sup>(a)</sup>The sources in bold are the star formation indicators in SIMBAD database. <sup>(b)</sup>The abbreviation “C.” means candidate.

## 2. Sample and data

### 2.1. Sample of HMSC candidates

There are several studies dedicated to the search of starless clumps in the Milky Way. Elia et al. (2017) explore the physical properties of *Herschel* infrared Galactic Plane Survey (Hi-GAL, see Molinari et al. 2010) compact sources in the inner Galaxy and termed the starless cores or clumps when they are not associated with 70  $\mu\text{m}$  counterpart. A large number of 76338 starless clump/core candidates are found. About 65, 52, and 3% of these clumps are compatible with massive star formation according to the different thresholds, which are  $M(r) > 870 M_{\odot}(r/\text{pc})^{1.33}$ , or  $\Sigma_{\text{crit}} = 0.2$  and  $1 \text{ g cm}^{-2}$  suggested by Kauffmann & Pillai (2010), Butler & Tan (2012), and Krumholz & McKee (2008), respectively. With Bolocam Galactic Plane Survey data (BGPS; Aguirre et al. 2011), Svoboda et al. (2016) (S16) obtained a sample of 2223 starless clump candidates in  $10^{\circ} < l < 65^{\circ}$  based on the absence of a series of observational signatures of star formation (see Table 1).

The recent search for starless clumps by Y17, which targeted at only massive starless clumps ( $\geq 100 M_{\odot}$  for 90% of these clumps at a typical scale of 0.2–0.8 pc) in the inner Galactic plane ( $|l| < 60^{\circ}$ ,  $|b| < 1^{\circ}$ ), resulted in a sample of 463 HMSC candidates by extracting APEX Telescope Large Area Survey of the Galaxy 870  $\mu\text{m}$  clumps (ATLASGAL; Csengeri et al. 2014) based on a series of selection criteria. Firstly, Y17 selected clumps with a 870  $\mu\text{m}$  peak intensity higher than  $0.5 \text{ Jy beam}^{-1}$  to ensure that the column density is sufficient to form a massive star. Then, these authors removed clumps with star formation indicators such as young stellar objects (YSOs) identified by Galactic Legacy Infrared Midplane Survey Extraordinaire data (GLIMPSE; Churchwell et al. 2009), MIPS Galactic Plane Survey 24  $\mu\text{m}$  data (MIPSGAL; Carey et al. 2009), and *Herschel* 70  $\mu\text{m}$  data (Molinari et al. 2010). Besides, various kinds of star formation indicators in the SIMBAD database<sup>1</sup> were also used. Finally, all candidates were inspected in *Spitzer* images by eye.

A comparison between the samples of Y17 and S16 shows that only 20% of ATLASGAL counterparts of BGPS starless clumps are in the HMSC sample of Y17. This could be due to denser or more reliably starless properties of HMSCs in the sample of Y17. Additional star formation indicators such as outflows (candidates) are considered. Furthermore, the inclusion

<sup>1</sup> <http://simbad.u-strasbg.fr/simbad/sim-display?data=otypes>

of extended  $24\ \mu\text{m}$  could mitigate false identifications due to a strong background. Another factor could be the better resolution of ATLASGAL (about  $20''$ ) compared to BGPS (about  $33''$ ), which may resolve one BGPS clump to more individual structures in ATLASGAL considering that BGPS could detect the similar mass range as ATLASGAL (Ellsworth-Bowers et al. 2015).

We chose the sample by Y17 as our HMSC candidates because it is probably more robust and focused on massive clumps. Even though a larger number of star formation indicators are cross matched, the selected HMSCs are still starless candidates rather than absolutely starless (Pillai et al. 2019). One of the examples is the massive  $70\ \mu\text{m}$  quiet clump, which was considered a more reliable starless candidate in previous, while evidences of embedded low- to intermediate-mass star formation have been found in such clump recently (Traficante et al. 2017; Svoboda et al. 2019; Li et al. 2019; Sanhueza et al. 2019). Basically, HMSCs hereafter represent HMSC candidates unless there are special explanations.

## 2.2. Ionized region sample

The ionized regions created during HMSF produce a shell or layer of photodissociation region (PDR) that can be observed by the PAH<sup>2</sup> emission in the mid-IR, for example, via GLIMPSE  $8\ \mu\text{m}$  or Wide-field Infrared Survey Explorer  $12\ \mu\text{m}$  emission (WISE; Wright et al. 2010). This shell or layer surrounds the hot thermal dust emission of ionized regions, which can be observed in MIPS GAL  $24\ \mu\text{m}$  or WISE  $22\ \mu\text{m}$  emission. The photodissociation shell and the inner hot thermal dust emission in mid-IR images make up the objects that are generally named as IR dust bubbles (Deharveng et al. 2010).

The ionized regions are originally traced by the emission from ionized gas such as the free-free centimeter continuum or radio recombination lines (RRL). The close correlations between the mid-IR hot dust emission and the centimeter radio free-free emission of H II regions have been revealed by several studies (Ingallinera et al. 2014; Makai et al. 2017), which indicate a tight spatial association between H II regions and IR dust bubbles. Churchwell et al. (2006) cataloged 322 IR bubbles by visual examination of GLIMPSE mid-IR images. These authors argued that most of the bubbles are formed by hot young stars in HMSF regions. Only three of their bubbles are identified as known supernova remnants (SNRs). The overlapping fraction between H II regions and bubbles could approach  $\approx 85\%$  or even more in other similar studies (Deharveng et al. 2010; Bania et al. 2010; Anderson et al. 2011), which shed light on the association between H II regions, bubbles, and HMSF.

We used the IR bubbles identified in mid-IR images as our candidates of H II regions. Based on GLIMPSE images, a large catalog of more than 5000 IR bubbles is attained through the visual identification by tens of thousands citizens within Milky Way Project (MWP) (Simpson et al. 2012; Kendrew et al. 2012). This catalog consists of the following two types of bubbles based on angular size: large ( $>30''$ ) and small ( $<30''$ ) bubbles. In this work, we take advantage of the sample of 3744 large bubbles as our H II region candidates. The small bubble candidates are ignored because of their high level of confusion with other astronomical objects unrelated to star formation such as AGB. Another reason is that the resolutions of millimeter and far-infrared (far-IR) images in our studies are close to the scale of

small bubble ( $20''$ – $30''$ ). The unresolved H II regions could mix with nearby HMSCs, leading to the unreal properties of HMSCs.

In addition, Anderson et al. (2014) constructed a catalog of over 8000 Galactic H II regions (candidates), around 2000 of which are known H II regions. They developed an online tool for viewing WISE emission of these H II regions<sup>3</sup>, which is updated when there are new radio continuum and RRL observations such as H II Region Discovery Survey (HRDS) (Anderson et al. 2015, 2018; Wenger et al. 2019). The distant Galactic H II regions are also included in HRDS (Anderson et al. 2015). Anderson et al. (2014) propose that WISE has a sensitivity to detect all Galactic H II regions (see Fig. 2 in their paper). The size and position of H II region (bubble) candidates in this online tool are outlined by characteristic morphology of WISE mid-IR emission. We use this catalog and associated online tool as supplements in searching for our H II region candidates.

## 2.3. Survey data

The main observing parameters of the surveys we used for this work are listed in Table 2. Several types of data are included as follows: (1) infrared data: GLIMPSE mapped the Galactic plane in mid-IR and we used this survey to show the spatial distribution of PAH, hot dust, and YSOs. (2) (Sub)millimeter and far-IR continuum data: ATLASGAL  $870\ \mu\text{m}$  survey ( $|l| < 60^\circ$ ,  $|b| < 1.5^\circ$ ) and Point Process Mapping for Hi-GAL (Hi-GAL PPMAP; Marsh et al. 2015) column density and dust temperature data are used. A more detailed description about Hi-GAL PPMAP is given in Sect. 3.3. (3) (Sub)millimeter molecular line data: Structure, Excitation, and Dynamics of the Inner Galactic InterStellar Medium survey (SEDIGISM; Schuller et al. 2017) mapped the inner Galactic plane ( $-60^\circ \leq l \leq 18^\circ$ ,  $|b| \leq 0.5^\circ$ ) in the  $J = 2-1$  rotational transition of  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  using the SHFI instrument on APEX with a spatial resolution of about  $30''$  and a velocity resolution of  $0.25\ \text{km s}^{-1}$  (Schuller et al. 2017). We used this survey to explore kinematics of the HMSCs. The survey is described in more detail in Sect. 8. (4) Centimeter radio images: Multi-Array Galactic Plane Imaging Survey (MAGPIS, first quadrant the Galactic plane; Helfand et al. 2006), Sydney University Molonglo Sky Survey (SUMSS, southern sky; Bock et al. 1999), and VLA Galactic Plane Survey (VGPS,  $18^\circ < l < 67^\circ$  and  $b < 1.3^\circ$  to  $3.3^\circ$ ; Stil et al. 2006) are used to indicate the H II regions.

## 3. Physical parameters

### 3.1. Velocity and distance

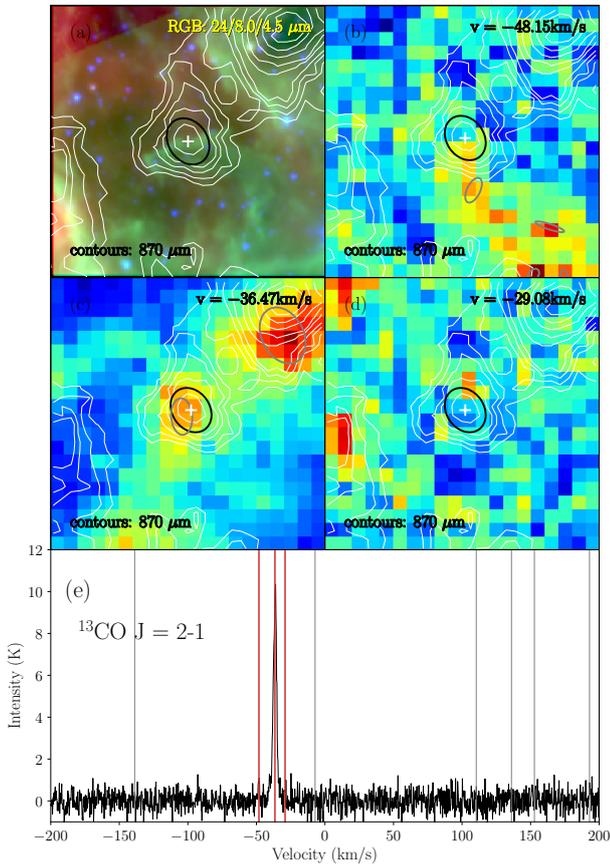
A reliable velocity determination is crucial to estimate a reliable set of other physical properties of the clumps. We take advantage of the results of velocity component identification from the Hi-GAL distance project (Mege et al., in prep.) to revise the systematic velocity ( $v_{\text{lsr}}$ ) estimated by Y17. This project aims to determine the  $v_{\text{lsr}}$  and distance of about 150 000 Hi-GAL-selected compact sources as accurately as possible by using all the molecular tracers available for the sources. In practice, we determine the most reliable  $v_{\text{lsr}}$  of a clump by carefully inspecting the spatial association and morphology between dust and molecular line emission. The velocity component with a morphology most similar to ATLASGAL emission of the HMSCs is selected. An example of the velocity determination of the HMSCs with the results of Hi-GAL distance project is shown in Fig. 1.

<sup>2</sup> Polycyclic aromatic hydrocarbon (PAH) commonly exists in the PDR (Fleming et al. 2010).

<sup>3</sup> <http://astro.phys.wvu.edu/wise/>

**Table 2.** Information about the surveys used in this work.

Survey	Band	Resolutions	Pixel sizes	Facilities	References
GLIMPSE	3.6 $\mu\text{m}$	$\approx 2''$	1.2''	<i>Spitzer</i> IRAC	Churchwell et al. (2009)
	4.5 $\mu\text{m}$				
	5.8 $\mu\text{m}$				
	8.0 $\mu\text{m}$				
ATLASGAL	870 $\mu\text{m}$	19.2''	6''	APEX LABOCA	Schuller et al. (2009)
SEDIGISM	220 GHz	30''	9.5''	APEX SHFI	Schuller et al. (2017)
PPMAP	–	12''	6''	<i>Herschel</i>	Marsh et al. (2015)
MAGPIS	6 cm	1.8''	0.6''	VLA	Helfand et al. (2006)
	21 cm	6.0''	2''		
	90 cm	25''	6''		
SUMSS	35 cm	$43'' \times 43''$ cosec $\delta$	11''	MOST	Bock et al. (1999)
VGPS	20 cm	45''	18''	VLA	Stil et al. (2006)



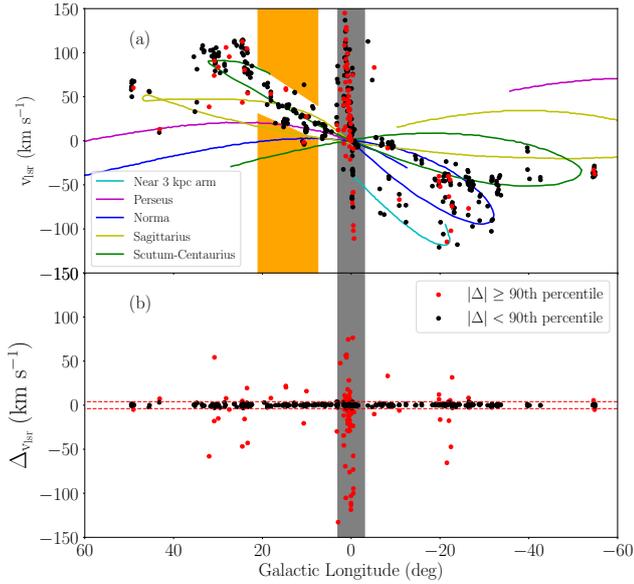
**Fig. 1.** Example of the  $v_{\text{lsr}}$  determination with the results of Hi-GAL distance project for the HMSC G305.1543+0.0477. *Panel a:* ATLASGAL contours overlaid on RGB image constructed from GLIMPSE 24  $\mu\text{m}$  (in red), 8  $\mu\text{m}$  (in green), and 4.5  $\mu\text{m}$  (in blue) images. *Panels b–d:* SEDIGISM  $^{13}\text{CO } J = 2-1$  intensity at the velocity of  $-48.15$ ,  $-36.47$ , and  $-29.08 \text{ km s}^{-1}$ , respectively. Black ellipses indicate the size and position of the HMSCs taken from Y17. Gray ellipses indicate the clumps identified in the corresponding intensity maps of the Hi-GAL distance project. Gray and red solid lines in (e) indicate the  $v_{\text{lsr}}$  components identified with  $^{13}\text{CO } J = 2-1$  spectra of the HMSC by the Hi-GAL distance project. The red lines highlight the  $v_{\text{lsr}}$  for the corresponding intensity maps (b), (c), and (d). According to the mapping results, the most proper  $v_{\text{lsr}}$  for the clump is  $-36.47 \text{ km s}^{-1}$ .

After revising the velocity, we used the same methods as that adopted in Y17 to derive the distance from renewed velocity, which is a parallax-based distance estimator developed by Reid et al. (2016) based on the Bayesian maximum-likelihood approach. We carefully considered kinematic distance ambiguities (KDA) when determining the distance. We applied the KDA solution results of Urquhart et al. (2018) to ascertain the distance. If the HMSCs are not included in Urquhart et al. (2018), we followed their work flows to solve the ambiguities, which use a series of evidence such as the scale height of the Galactic plane, H I self-absorption (H I SA), and extinction (IR dark clouds). More detailed information about KDA solutions is given in Appendix A.

### 3.2. Kinematic uncertainty

The renewed  $v_{\text{lsr}}$  and its comparisons with that in Y17 are shown in Fig. 2. Most of the HMSCs except for the low-longitude sources are located at or close to spiral arms as shown in Fig. 2a. The HMSCs toward the Galactic center have a radial component of the velocity that is only a small fraction of the entire velocity of the clump, making the rotation curve and resulted kinematic distance are just not reliable. Using any kinematic information in extremely low longitude is challenging, as shown by the significant  $l - v_{\text{lsr}}$  structure blending in  $|l| < 3^\circ$  region in Fig. 2a. Furthermore, the commonly existing broad line width (several dozens of  $\text{km s}^{-1}$ ) of molecular lines in this region increases the difficulties to identify a proper  $v_{\text{lsr}}$  (Kruijssen et al. 2015; Ginsburg et al. 2016). The 90th percentiles of the identified  $v_{\text{lsr}}$  difference between Y17 and this paper are  $3.9 \text{ km s}^{-1}$  and  $64.6 \text{ km s}^{-1}$  for  $|l| > 3^\circ$  and  $|l| \leq 3^\circ$ , respectively. This shows the large uncertainty in  $v_{\text{lsr}}$  identification for low-longitude clumps.

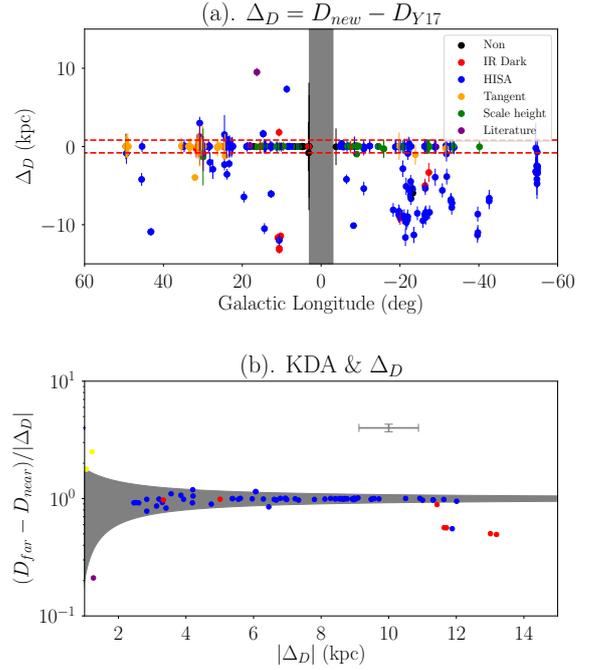
The strong radial streaming motions caused by the Galactic bar, which could also impact the  $l - v_{\text{lsr}}$  structure, are carefully considered by Ellsworth-Bowers et al. (2015). Figure 2a shows that only few HMSCs are located in the exclusion regions proposed by Ellsworth-Bowers et al. (2015) and that these HMSCs are close to the boundaries of exclusion regions. Thus, this kind of uncertainty is not dominating for our HMSCs. Considering the unreliable velocity and kinematic distance in lower longitude, we exclude a total of 125 HMSCs located in  $|l| \leq 3^\circ$  from the following analyses and visualization.



**Fig. 2.** New HMSC  $v_{\text{lsr}}$  (this work) and those in Y17. *Panels a and b:* renewed  $v_{\text{lsr}}$  and difference between renewed  $v_{\text{lsr}}$  and those in Y17,  $\Delta v_{\text{lsr}} = v_{\text{lsr new}} - v_{\text{lsr Y17}}$ , respectively. The 90th percentile of  $|\Delta v_{\text{lsr}}|$  for HMSCs located at higher longitude ( $|l| > 3^\circ$ ) is  $3.9 \text{ km s}^{-1}$ , shown as red dashed lines in (b). The red and black dots represent HMSCs with  $|\Delta v_{\text{lsr}}|$  values larger and smaller than  $3.9 \text{ km s}^{-1}$ , respectively. The gray region indicates the lower-longitude region,  $|l| < 3^\circ$ , where the HMSCs are excluded from following analyses. Orange regions are considered by Ellsworth-Bowers et al. (2015). The four spiral and local arms are shown by different colors and their positions are taken from Urquhart et al. (2018) and its references Taylor & Cordes (1993) and Bronfman et al. (2000). Most of the errors of  $v_{\text{lsr}}$  are less than  $2 \text{ km s}^{-1}$  and the associated error bars are smaller than the point size in (a).

The uncertain kinematics is not the only reason to exclude the lower-longitude HMSCs. Possible different star formation mechanism and history compared to other molecular cloud–star formation complexes are also possible reasons. The central molecular zone (CMZ) is about  $400 \text{ pc}$  in size (angular scale at the distance of the Earth is about  $0.4/8.34 \text{ rad} \approx 3^\circ$ ). The star formation rate (SFR) in the CMZ is an order of magnitude lower than expected (Lu et al. 2019). One of the explanations for the low SFR is that the clouds in CMZ are without active star formation at very early evolutionary stages because of the co-impacts of high turbulence, magnetic field, and a number of other processes (Kruijssen et al. 2014). A higher density is needed for the collapse of the clumps in the CMZ. Therefore, the inhibited HMSF could result in the overproduction of HMSCs (26% HMSCs are in the region  $|l| < 3^\circ$ ). Our study focuses on the impacts of H II regions on HMSCs, thus it is more reliable to only use the higher-longitude HMSCs, which could avoid the complex co-impacts in CMZ.

The 70th percentile of the distance difference  $|\Delta_D|$  between this paper and Y17 is about  $0.8 \text{ kpc}$ , which is shown in Fig. 3a in red dashed lines. Owing to the small  $|\Delta v_{\text{lsr}}|$  for most HMSCs, the improvement in solving KDA problem could be the most possible reason for the large difference between our new distance determination and those in Y17. For HMSCs with a similar  $v_{\text{lsr}}$  but large  $|\Delta_D|$ , their  $|\Delta_D|$  are approximately equal to the differences between far and near distance  $D_{\text{far}} - D_{\text{near}}$  in KDA solutions, as shown in Fig. 3b; this suggests the large  $|\Delta_D|$  are mainly due to the difference of KDA solutions. Nearly 70 HMSCs in total have a different KDA solution compared to those in Y17.



**Fig. 3.** New distance of HMSC (this work) and those in Y17. *Panel a:* distance difference  $\Delta_D = D_{\text{new}} - D_{\text{Y17}}$ . The red dashed lines show  $|\Delta_D| = 0.8 \text{ kpc}$ , representing the 70th percentile of  $|\Delta_D|$ . The red, blue, orange, green, and purple dots mean that their KDA are solved by IR dark, H I SA, tangent, scale height, and literature, respectively (see Table 4 and Appendix A). The black dots are the sources without KDA solution. *Panel b:*  $|\Delta_D|$  and its relation with the difference between resulted far distance and near distance  $D_{\text{far}} - D_{\text{near}}$ . Only HMSCs with  $|\Delta_D| > 0.8 \text{ kpc}$  and  $|\Delta v_{\text{lsr}}| < 3.9 \text{ km s}^{-1}$  are shown. The gray region shows  $(D_{\text{far}} - D_{\text{near}})/|\Delta_D| = 1 \pm 0.8/|\Delta_D|$ , indicating the error of relation. The gray cross shows the typical error.

### 3.3. Column density and dust temperature

One of the classical ways to investigate the dust properties of HMSC is to combine far-IR to millimeter multiwavelength data and then fit the spectral energy distribution (SED) pixel by pixel. All images need to be smoothed to the largest beam size in the data set and then fitted with the modified blackbody function at each pixel. Principally, we made use of the  $\text{H}_2$  column density ( $N_{\text{H}_2}$ ) and dust temperature ( $T_{\text{dust}}$ ) resulting from the Y17 SED fitting to update the physical properties according to the renewed distance. Besides, we also utilized the higher resolution PPMAP Hi-GAL data set in investigations of density and temperature profile.

#### 3.3.1. Results from Y17

In Y17, the Hi-GAL 160, 250, 350, 500  $\mu\text{m}$ , and ATLAS-GAL 870  $\mu\text{m}$  images are smoothed to the largest beam of the data set, which is  $36.4''$ . After carefully removing the foreground/background flux, gray-body functions are fitted pixel-by-pixel with a dust opacity of  $\kappa_\nu = 3.33(\nu/600\text{GHz})^\beta \text{ cm}^2 \text{ g}^{-1}$  and an emissivity index of  $\beta = 2$ . In this paper, physical properties (e.g., clump mass  $M_{\text{clump}}$ , number density  $n_{\text{H}_2}$ , and size  $r_{\text{pc}}$ ) are basically taken from Y17 but revised according to the new distance determination.

#### 3.3.2. Higher resolution results from PPMAP

To explore the density/temperature structures of HMSCs, data with higher resolution is helpful because HMSF regions are

generally located at a distance farther than low mass star formation regions. To avoid degrading resolution when performing fitting process, Marsh et al. (2015) fit the SED based on the nonhierarchical Bayesian procedure point process mapping technique to utilize full instrumental point source functions (PSF). Marsh et al. (2017) applied the PPMAP technique on the Hi-GAL 70, 160, 250, 350, and 500  $\mu\text{m}$  images to get a series of  $T_{\text{dust}}$  – differential  $N_{\text{H}_2}$  planes, which means the column densities are equally spaced in log space between 8 and 50 K and the dust temperature is the third dimension of  $N_{\text{H}_2}$  maps (data cube). The spatial resolution of the resulting PPMAP Hi-GAL  $N_{\text{H}_2}$  and  $T_{\text{dust}}$  images is  $12''$ ; the value is better than that from the conservative SED fitting for Hi-GAL images ( $\approx 36''$ ). The typical errors in the resulting  $N_{\text{H}_2}$  and  $T_{\text{dust}}$  are around 10–50% and around 1 K, respectively.

Comparisons between PPMAP Hi-GAL  $T_{\text{dust}}$  and  $N_{\text{H}_2}$  with those in Y17 are shown in Fig. 4. Generally, the PPMAP Hi-GAL  $T_{\text{dust}}$  value is a higher of 2–3 K than the value obtained in Y17. At the low –  $T_{\text{dust}}$  end ( $< 14$  K for  $T_{\text{dust}}$  of Y17), the difference can reach 5–7 K. The most likely reason for that is the difference in the wavelength of the data used because similar dust opacity and graybody functions are used by PPMAP and Y17. Y17 excluded Hi-GAL 70  $\mu\text{m}$  emission and used additional 870  $\mu\text{m}$  data to fit the SED, which could better constrain the shape of the SED at the long-wavelength end and lead to trace more the coldest parts of the clumps. Besides, Y17 carefully excluded the foreground/background flux when fitting the SED while this is not included in PPMAP Hi-GAL SED fitting.

When the data used in fitting are the same, the PPMAP technique could provide a more accurate estimation of  $N_{\text{H}_2}$  and  $T_{\text{dust}}$ , especially for the maximum of  $N_{\text{H}_2}$ , minimum of  $T_{\text{dust}}$  of the clump (Marsh & Whitworth 2019). Even PPMAP Hi-GAL data should detect more compact structures than Y17 because of the higher resolution of the images; Fig. 4 shows that clumps  $N_{\text{H}_2}$  are normally underestimated while  $T_{\text{dust}}$  are overestimated at the low –  $T_{\text{dust}}$  end compared to Y17. This suggests that the errors created during the SED-fitting process are dominated by the different wavelength data used rather than different techniques or resolutions in our cases.

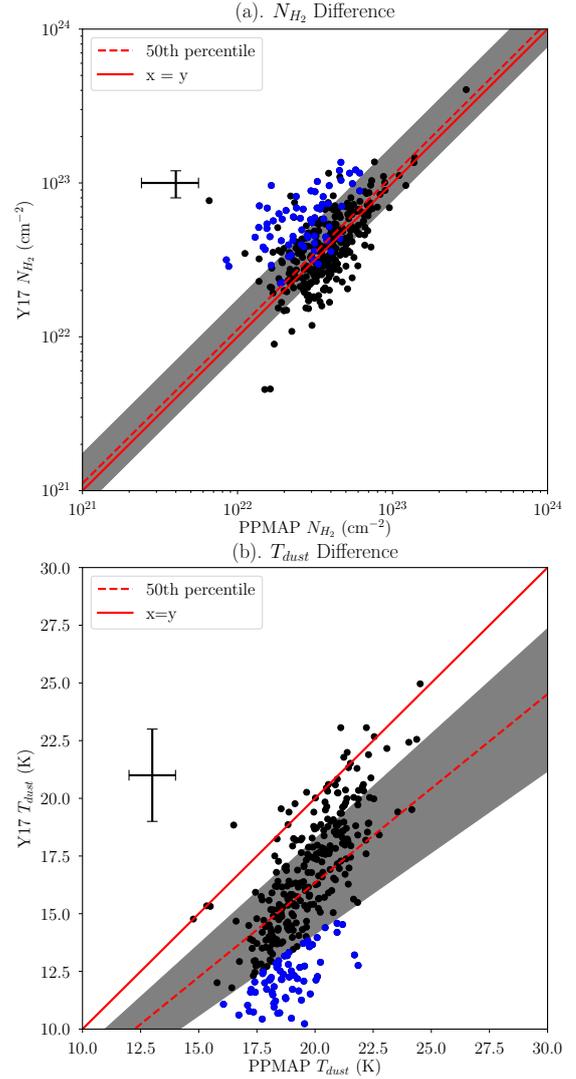
Considering these factors, we only used PPMAP data to estimate the column density/temperature structures (see details in Sect. 7) rather than using these data to derive other global properties of HMSCs such as  $N_{\text{H}_2}$ ,  $T_{\text{dust}}$ , and the ratio of bolometric luminosity to envelope mass  $L/M$ .

## 4. H II region cross matching

### 4.1. Criteria

To explore differences between HMSCs associated with and nonassociated with H II regions, we set two main criteria to ascertain the association relationship.

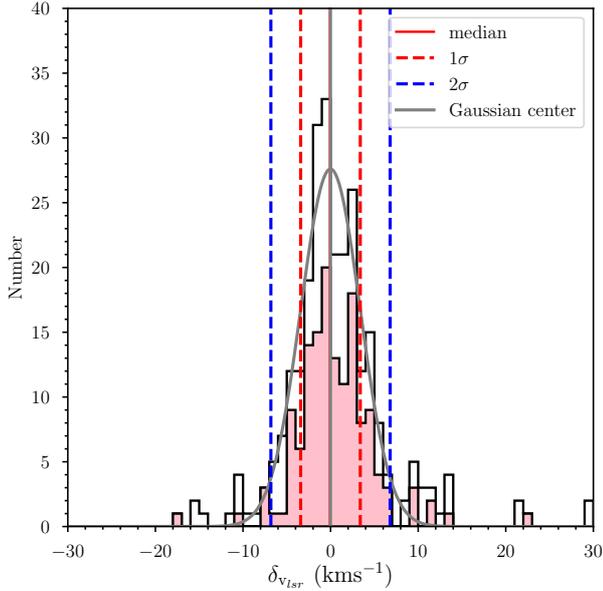
The first criterion is the position in the plane of the sky. Thompson et al. (2012) studied Red MSX Sources (RMSs; see Lumsden et al. 2013) YSOs, and Ultra Compact (UC) H II regions, which are regarded as the signatures of HMSF, around 322 *Spitzer* mid-IR bubbles. With an angular cross-correlation function of RMSs and bubbles, these authors reveal that the location with the highest probability to find RMSs YSOs is around the effective radius  $R_{\text{eff}}$  of the bubble. Furthermore, the RMSs are essentially uncorrelated with bubbles and appear as field RMSs when the separations are larger than  $2R_{\text{eff}}$ . A similar relationship for cold clumps has also been found in the MWP bubbles (Kendrew et al. 2016). If our HMSCs are precursors of



**Fig. 4.** HMSCs  $N_{\text{H}_2}$  and  $T_{\text{dust}}$  derived from PPMAP Hi-GAL data and those in Y17. The red solid lines show 1:1 relations. The red dashed lines show the 50th percentile of the ratios of Y17 derived parameters to PPMAP Hi-GAL derived parameters. The gray regions show the 20th percentile to 80th percentile of the ratios of Y17’s parameters to PPMAP parameters. The blue points are sources with the lowest  $T_{\text{dust}}$  and the largest difference of  $T_{\text{dust}}$  between Y17 and PPMAP.

HMSF that could form massive stars in the future, we can simply assume that the HMSCs associated with bubbles have a distance correlation similar to the RMSs or UC H II regions. A study that looked into 70 000 star-forming objects around bubbles reveals that either the radial source density profiles of starless clumps or those of protostars become quite flat and near the field values when beyond  $2R_{\text{eff}}$  (Palmeirim et al. 2017). This suggests that the nonassociated properties of the clumps outside  $2R_{\text{eff}}$  of H II region are independent of the evolutionary stages of the clumps. Combining these facts, we propose that  $2R_{\text{eff}}$  is a reasonable value for declaring association in sky position between HMSCs and H II regions. We make use of the  $R_{\text{eff}}$  in Kendrew et al. (2012) and Anderson et al. (2014), which is from the measurement of the mid-IR PAH emission of the bubble (see Sect. 2.2).

The second criterion is the  $v_{\text{lsr}}$  difference  $|\delta_{v_{\text{lsr}}}| = |v_{\text{HMSC}} - v_{\text{H II region}}|$  between HMSCs and the H II regions. A proper limitation on  $|\delta_{v_{\text{lsr}}}|$  could mitigate mismatching between H II region



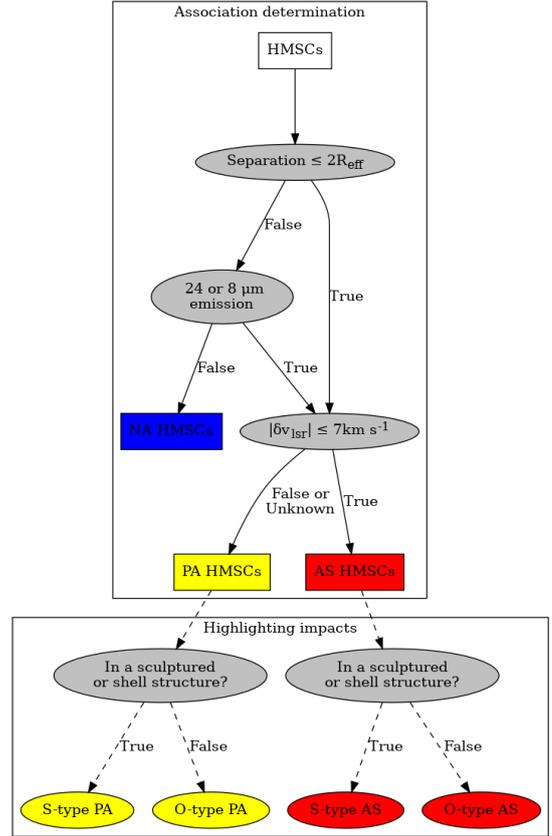
**Fig. 5.** Number distribution of  $v_{\text{lsr}}$  difference between H II region and HMSC  $\delta_{v_{\text{lsr}}} = v_{\text{HMSC}} - v_{\text{H II region}}$  in the  $-30$  to  $30 \text{ km s}^{-1}$  range. The gray lines are the fitted Gaussian and its center. The pink histogram highlights the distribution for S-type HMSCs (see Sect. 4.4). The red and blue dashed lines represent  $1\sigma$  and  $2\sigma$  of Gaussian, respectively. The median value of  $\delta_{v_{\text{lsr}}}$  overlaps with Gaussian center in the figure.

and HMSCs due to the projection in the plane of sky. If a HMSC is indeed associated with an H II region,  $\delta_{v_{\text{lsr}}}$  mainly comes from the molecular cloud turbulence and the expansion of H II region. By Larson’s law (Larson 1981), the velocity dispersion  $\sigma_v$  of a molecular cloud is related with its size. If we assume a cloud size of  $5 \text{ pc}$ , which is similar to the typical size of IR bubbles, the velocity dispersion  $\sigma_v$  is about  $0.9 \times 5^{0.56 \pm 0.02} \text{ km s}^{-1} \approx 2 \text{ km s}^{-1}$  (Heyer & Brunt 2004). This value is a bit smaller than the  $^{13}\text{CO } J = 1-0$  observations of Yan et al. (2016) toward 13 Galactic bubbles, which resulted in a value of about  $2.5\text{--}5 \text{ km s}^{-1}$ . The larger  $\sigma_v$  may be due to the additional energy injection from massive stellar feedback such as stellar winds and UV radiation. Another factor that enlarges  $|\delta_{v_{\text{lsr}}}|$  is the expansion of the H II region. The simulation of Tremblin et al. (2014a) for the expanding process of ionized gas in turbulent molecular clouds indicates an averaged expanding velocity  $\approx 2 \text{ km s}^{-1}$ . Molecular spectral observations toward expanding shells surrounding H II regions indicate an expanding velocity on the order of one to several  $\text{km s}^{-1}$  (Elmegreen 2011).

To ascertain a proper limitation,  $\delta_{v_{\text{lsr}}}$  for all HMSCs in  $2R_{\text{eff}}$  of H II regions with known  $v_{\text{lsr}}$  are shown in Fig. 5. There are only few sources whose  $|\delta_{v_{\text{lsr}}}| > 30 \text{ km s}^{-1}$ , thus they are not shown in Fig. 5. Gaussian fitting to  $\delta_{v_{\text{lsr}}}$  distribution gives a Gaussian standard deviation  $\sigma$  of  $\approx 3.5 \text{ km s}^{-1}$ . Considering all the facts mentioned above, we assume that a  $v_{\text{lsr}}$  difference limit of  $7 \text{ km s}^{-1}$ , about  $2\sigma$  of fitted Gaussian distribution, is a reasonable value to reduce the misclassification due to the projection on the line of sight. We made use of H II region  $v_{\text{lsr}}$  information from the work of Hou & Gao (2014) and Anderson et al. (2014), which are usually derived from  $\text{H}\alpha$ , RRL observations or the average molecular lines of H II regions.

#### 4.2. Workflow

According to the two criteria mentioned, we classified HMSCs into three types that reflect the association relations between



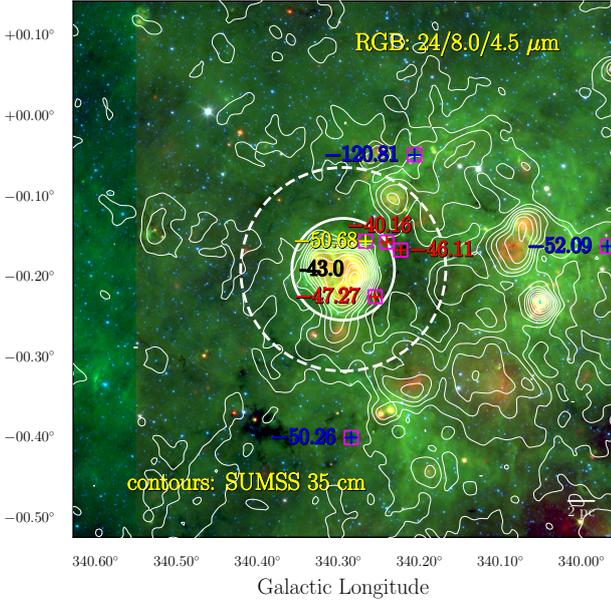
**Fig. 6.** Workflow chart. The first part is the determination of association as described in Sect. 4.2 and the second part is for highlighting impact of H II regions, as described in Sect. 4.4.

HMSCs and H II regions. Generally, when a HMSC is located outside  $2R_{\text{eff}}$  of H II regions or bubbles, we classify it as non-associated HMSC (written as NA HMSC or NA), otherwise it is classified as possibly associated HMSC (written as PA HMSC or PA) or associated HMSC (written as AS HMSC or AS). The criterion for discriminating between PA and AS is whether the absolute value of  $v_{\text{lsr}}$  difference  $|\delta_{v_{\text{lsr}}}|$  between HMSC and H II region is less than  $7 \text{ km s}^{-1}$ . The HMSC is classified as AS if this is the case, otherwise it is classified as PA. When a HMSC is located in  $2R_{\text{eff}}$  but there is no  $v_{\text{lsr}}$  information for the H II region, we also put it in the category of PA HMSCs.

Considering that some of the H II regions or bubbles have an irregular or elliptical shape, a single circular description with  $R_{\text{eff}}$  is not accurate enough. We check GLIMPSE  $24 \mu\text{m}$  hot dust emission and  $8 \mu\text{m}$  PAH emission images to confirm that all NA HMSCs classified by the criterion of  $2R_{\text{eff}}$  are really far away from any  $8 \mu\text{m}$  emission shell structure or diffuse  $24 \mu\text{m}$  emission. If this is the case, we keep it in the NA category, otherwise we reclassify it as PA or AS. A flow chart describing how classification process works is shown in Fig. 6. An example for the identifications of AS, PA, and NA HMSCs surrounding the H II region G340.294–00.193 are shown in Fig. 7.

#### 4.3. Bias and results

Large ( $>30'$ ) and very diffuse (weak at  $24 \mu\text{m}$ ) H II regions are not considered when cross matching H II regions and HMSCs. Some AS and PA could be misclassified as NA owing to the omission of the very diffuse H II regions. We suggest that the omission of very diffuse H II regions does not largely change



**Fig. 7.** H II region G340.294–00.193. The  $R_{\text{eff}}$  and  $2R_{\text{eff}}$  of the H II region are indicated as solid and dashed white circles, respectively. The purple squares highlight the position of HMSCs. The red, yellow, and blue crosses represent AS, PA, and NA HMSCs, respectively. The corresponding  $v_{\text{lsr}}$  in  $\text{km s}^{-1}$  is shown close to the mark. Number  $-43.0$  is the  $v_{\text{lsr}}$  of H II region in  $\text{km s}^{-1}$ . The RGB image is constructed in the same method as Fig. 1. The white contours represent SUMSS 35 cm emission. The  $v_{\text{lsr}}$  of the H II region is taken from Hou & Gao (2014) and Anderson et al. (2014), see Sect. 4.1.

our comparison results between different HMSCs because the impacts of ionized gas of this kind of H II regions are too weak to significantly modify the properties of HMSCs.

Another bias is that some PA with  $|\delta v_{\text{lsr}}| > 7 \text{ km s}^{-1}$  could be due to the very powerful feedback of H II regions, which strongly changes the overall kinematics of molecular clouds, making  $|\delta v_{\text{lsr}}|$  larger than  $7 \text{ km s}^{-1}$ , rather than the properties of nonassociation. It is possible that the chemistry properties of HMSCs could help in our classification. For example,  $\text{C}_2\text{H}$  is a typical molecular tracer for photodissociation, indicating the impacts of H II region. In a forthcoming paper, we find that the  $\text{C}_2\text{H}$  detection rates between AS and NA are significantly different (Zhang et al., in prep.). Detailed case-studies will reveal whether these PA are AS or NA. In this paper, we only use the simplest methods with the velocity and sky position to classify the association of HMSCs and leave the more complex and detailed classifications to future studies.

To summarize, we classified HMSCs into three categories according to their association with H II region: (1) AS: the HMSC is in  $2R_{\text{eff}}$  of the H II region and its  $v_{\text{lsr}}$  difference with the H II region is  $\leq 7 \text{ km s}^{-1}$ ; (2) PA: the HMSC is in  $2R_{\text{eff}}$  of the H II region, but we do not know  $v_{\text{lsr}}$  of the H II region or the  $v_{\text{lsr}}$  difference is  $> 7 \text{ km s}^{-1}$ ; (3) NA: the separation between the HMSC and the H II region is  $> 2R_{\text{eff}}$ .

A summary of the classification results is shown in Table 3. About 60–80% of HMSCs are associated with H II regions, which reveals the close relation between the H II regions and HMSCs. We only use the simplest method with position and velocity to determine the association. The real associated number could be larger if we consider more features, like  $\text{C}_2\text{H}$ . Classification results and physical parameters of HMSCs are given in Table 4.

**Table 3.** Association of HMSCs and H II regions.

Sample	AS HMSCs	PA HMSCs	NA HMSCs
338 HMSCs in $ l  > 3^\circ$	193 (57%)	64 (19%)	81 (24%)

#### 4.4. Morphology features of impact of H II regions on HMSCs

The main goal of this section is to extract the HMSCs with clear impacts from the H II regions using the available data. According to the morphology of cold and hot dust emission, PAH, and ionized gas emission, we check whether HMSC is in a compressed dust shell or in a structure that is being photoionized or photodissociated by UV radiation. Steeper profiles of radio emission, ATLASGAL  $870 \mu\text{m}$  emission, or column density toward the direction of the H II region and/or a denser dust shell surrounding the H II region both indicate the compression by the H II region (Tremblin et al. 2014b, 2015; Li et al. 2018; Marsh & Whitworth 2019). The bright PAH  $8 \mu\text{m}$  emission layer toward the direction of  $24 \mu\text{m}$  hot dust emission or radio emission of the H II regions suggests that photodissociation is working at the surface of clumps and is “sculpturing” the clumps. We classify AS or PA HMSCs with these features as S (sculptured or shell) type AS or S-type PA. For AS and PA without morphology features of significant interaction with the H II regions, we just classify these AS or PA as O- (other) type AS or O-type PA. These O-type AS or PA could be in a filament, a clumpy structure, or even isolated without neighbor of clump (see Appendix B). The related classification process is shown in the work flow presented in Fig. 6. Table 5 shows that about half of AS are found to be probably significantly impacted by H II regions on the mapping data.

## 5. Distributions in the Galaxy

The distribution of HMSCs in the Galaxy is shown in Fig. 8. The trend is that either AS or NA all highlight the spiral arms and that both of these have a similar distribution in the Milky Way. This trend could be partly due to the coherence between HMSCs and spiral arms in the  $v_{\text{lsr}}$  – longitude structures shown in Fig. 2. Another factor could be the nature of Reid et al. (2016) distance calculator, which links the points in the position–position–space to the specified spiral arms. The peak of latitude distribution of AS shifts to  $b = -0.25^\circ$ . By checking our images of the H II regions - HMSCs complex, this shift could be partly due to several star-forming complexes with a number of AS at a relatively higher latitude, such as G010.19–0.35 (eight HMSCs) and G333.48–0.22 (eight HMSCs). Another factor could be due to the offset between solar system’s location and the Galactic midplane. The location of the Sun is thought to be about 20 pc to 25 pc above the Galactic midplane (Humphreys & Larsen 1995).

The  $T_{\text{dust}}$  of AS and PA seems to be dominated by the local environment rather than the Galactic-scale environment. In the same region of spiral arm, AS and PA usually have a higher  $T_{\text{dust}}$  than that of NA, which probably indicates that  $T_{\text{dust}}$  of AS and PA are deeply impacted by the associated H II regions. The relations between  $T_{\text{dust}}$  and the Galactocentric distance  $R_{\text{GC}}$  for AS and NA HMSCs are shown in Fig. 9. The NA occupy the lower –  $T_{\text{dust}}$  end in most distance ranges. Linear regressions to  $T_{\text{dust}}$  for AS and NA indicate the Galactocentric gradient of  $T_{\text{dust}}$  is weakly different between AS and NA HMSCs. The NA show a weak trend that  $T_{\text{dust}}$  decreases with  $R_{\text{GC}}$ , which conforms with the previous results about the Galactic cold dust temperature gradient owing to the decreasing star formation activities

**Table 4.** Classification and physical parameters of HMSCs.

Name <sup>(1)</sup>	$T_{\text{dust}}$ <sup>(2)</sup> (K)	$n_{\text{H}_2}$ ( $10^3 \text{ cm}^{-3}$ )	$N_{\text{H}_2}$ <sup>(3)</sup> ( $10^{21} \text{ cm}^{-2}$ )	$M_{\text{clump}}$ ( $M_{\odot}$ )	$L_{\text{clump}}$ ( $L_{\odot}$ )	$L/M$ ( $L_{\odot}/M_{\odot}$ )	$r_{\text{pc}}$ (pc)	$v_{\text{lsr}}$ ( $\text{km s}^{-1}$ )	Distance <sup>(4)</sup> (kpc)	KDA <sup>(5)</sup>	Association	Morphology <sup>(6)</sup>
G003.0371−0.0582	13.8	1.17(0.35)	53.6(8.0)	$1.21(0.64) \times 10^4$	$8.28(5.07) \times 10^3$	0.68(0.23)	3.3	85	13(3.3)		NA	C
G003.0928+0.1680	14.2	1.96(0.58)	80.2(12.0)	$2.42(1.28) \times 10^4$	$1.75(1.06) \times 10^4$	0.72(0.24)	3.5	150(0.12)	16(4.1)		NA	C
G003.1447+0.4733	14.3	1.90(0.29)	15.4(2.3)	$9.17(1.39) \times 10^2$	$7.16(2.70) \times 10^2$	0.78(0.29)	1.2	70(1.9)	11(0.12)		NA	F
G003.2110+0.6465	14.5	5.06(0.80)	22.9(3.4)	$1.89(0.34) \times 10^2$	$1.59(0.62) \times 10^2$	0.84(0.32)	0.51	18(0.25)	2.9(0.15)		NA	F
G003.2278+0.4924	14.4	6.14(0.97)	31.4(4.7)	$2.34(0.43) \times 10^2$	$1.88(0.72) \times 10^2$	0.8(0.29)	0.51	20(0.88)	2.9(0.15)	IR DARK	NA	F
G003.2413+0.6334	13	19.2(8.1)	59.7(9.0)	$4.99(4.03) \times 10^2$	$2.15(1.83) \times 10^2$	0.43(0.14)	0.45	41	4(1.6)		NA	F
G003.2702+0.4446	13.8	6.02(0.95)	36.9(5.5)	$3.87(0.70) \times 10^2$	$2.34(0.85) \times 10^2$	0.6(0.21)	0.61	27	2.9(0.15)		NA	C
G004.4076+0.0993	18.1	123(19)	20.8(3.1)	$8.99(1.63) \times 10^1$	$3.60(1.71) \times 10^2$	4(1.9)	0.14	10(0.35)	2.9(0.15)		AS	S
G004.7473−0.7623	12.7	6.53(0.99)	71.0(10.7)	$4.89(0.78) \times 10^3$	$1.81(0.56) \times 10^3$	0.37(0.11)	1.4	210	9.4(0.24)		NA	I
G005.3876+0.1874	12.9	52.7(8.3)	38.3(5.7)	$1.76(0.32) \times 10^2$	$7.44(2.52) \times 10^1$	0.42(0.14)	0.23	11(0.74)	2.9(0.15)		AS	F
G005.8357−0.9958	15.4	6.59(1.05)	23.9(3.6)	$1.25(0.23) \times 10^2$	$1.50(0.62) \times 10^2$	1.2(0.48)	0.4	13	2.9(0.15)	Z H	PA	I
G005.8523−0.2397	12.1	12.2(1.9)	45.0(6.7)	$2.30(0.42) \times 10^2$	$7.28(2.29) \times 10^1$	0.32(0.09)	0.4	17(0.13)	3(0.15)	HISA	NA	I
G005.8799−0.3591	16.5	30.0(4.8)	51.2(7.7)	$2.43(0.44) \times 10^2$	$4.38(1.80) \times 10^2$	1.8(0.72)	0.3	5.8(0.1)	2.9(0.15)	HISA	AS	S
G005.8893−0.4565	19.4	97.0(15.4)	46.6(7.0)	$2.16(0.39) \times 10^2$	$1.16(0.53) \times 10^3$	5.4(2.4)	0.2	9.7(0.23)	2.9(0.15)	IR DARK	AS	S
G005.8917−0.3567	17	20.6(3.3)	35.4(5.3)	$2.46(0.45) \times 10^2$	$5.44(2.36) \times 10^2$	2.2(0.93)	0.35	5.8(0.13)	2.9(0.15)	IR DARK	AS	S
G005.9394−0.3754	17.2	14.6(2.3)	21.0(3.2)	$1.32(0.24) \times 10^2$	$3.27(1.49) \times 10^2$	2.5(1.1)	0.31	6.1(0.1)	2.9(0.15)	HISA	AS	S
G006.2130−0.5937	15.8	23.6(3.7)	34.9(5.2)	$1.67(0.30) \times 10^2$	$2.44(1.00) \times 10^2$	1.5(0.58)	0.29	18(0.11)	3(0.15)	Z H	AS	F
G006.4609−0.3902	17.9	6.96(1.10)	4.58(0.69)	$2.02(0.37) \times 10^1$	$7.12(3.54) \times 10^1$	3.5(1.7)	0.22	19	3(0.15)		NA	I
G006.4916−0.3322	17.7	9.88(1.57)	4.54(0.68)	$2.17(0.40) \times 10^1$	$6.81(3.35) \times 10^1$	3.1(1.5)	0.2	2.2	2.9(0.15)		NA	I
G006.4982−0.3222	17.6	1.94(0.31)	8.96(1.34)	$5.66(1.03) \times 10^1$	$1.65(0.79) \times 10^2$	2.9(1.4)	0.46	5.1(0.25)	2.9(0.15)		NA	I

**Notes.** <sup>(1)</sup>Full table is available at the CDS. <sup>(2)</sup>The typical error of  $T_{\text{dust}}$  estimated by Y17 is about 2–3 K. <sup>(3)</sup>The typical error of  $N_{\text{H}_2}$  estimated is about 15% from Y17. <sup>(4)</sup>The distance error is taken from the results of distance calculator and its typical value is less than 1 kpc. The distance errors are propagated to the calculation of other physical parameters. <sup>(5)</sup>We solve the KDA with several ways: IR DARK. IR extinction. H I SA. H I self-absorption. Z H. The scale height in the Galactic plane. TANGENT. The tangent line of sight. LITERAT. The literature in [Urquhart et al. \(2018\)](#). See Appendix A for more information. <sup>(6)</sup>The detailed explanation about the morphology is in Appendix B.

**Table 5.** Morphology statistics.

Types	S-type	O-type
AS HMSCs (193)	104 (54%)	89 (46%)
PA HMSCs (64)	23 (36%)	41 (64%)

from the Galactic inner to outer regions ([Misiriotis et al. 2006](#); [Paradis et al. 2012](#)), whereas AS HMSCs show a weak positive trend. The Spearman’s correlations<sup>4</sup> result in a coefficient of 0.15 and  $-0.16$  of  $R_{\text{GC}} - T_{\text{dust}}$  relation for AS and NA, respectively, showing the results of different  $R_{\text{GC}} - T_{\text{dust}}$  relations are weak. The probably different  $T_{\text{dust}}$  gradient may indicate the importance of impacts of H II regions on the local environment. In the following sections, we confirm the  $T_{\text{dust}}$  difference of HMSCs in different environments.

## 6. Basic physical properties

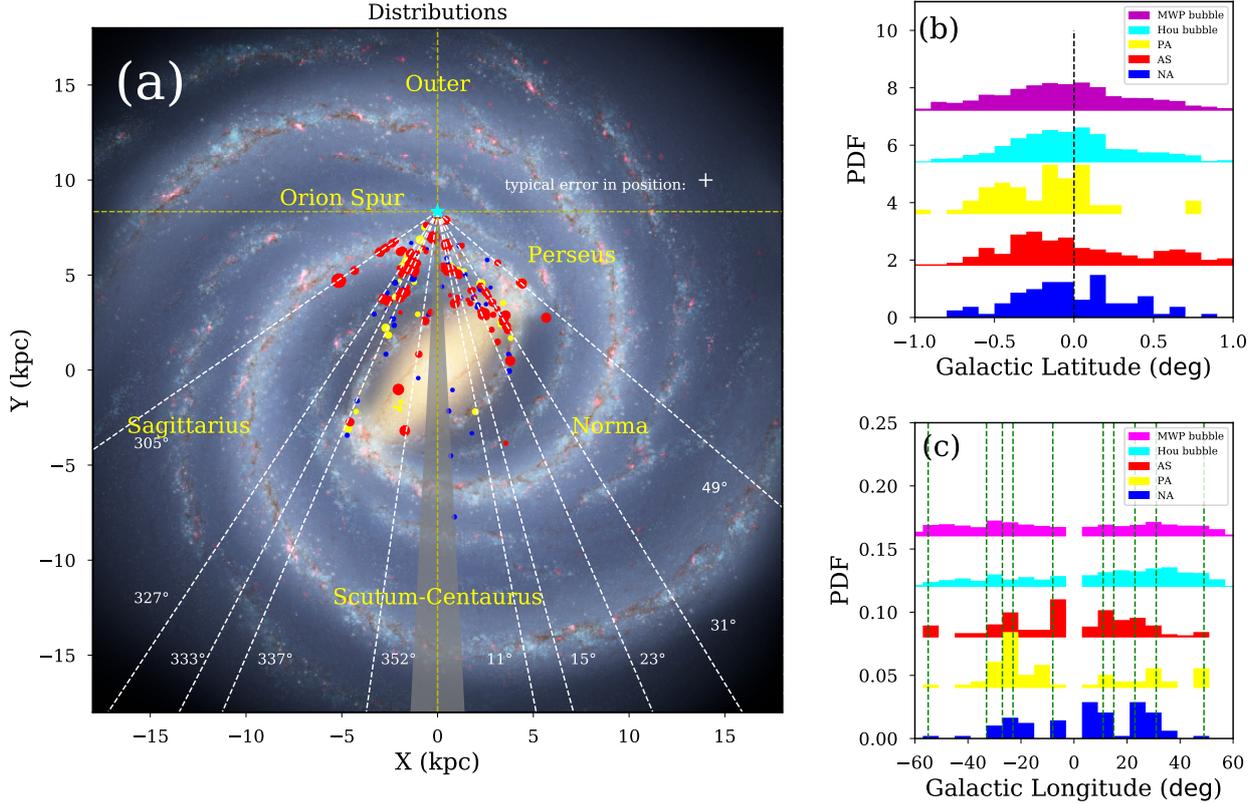
To explore physical differences of HMSCs in different environments with a more reliable sample, we remove HMSCs with a mass smaller than  $100 M_{\odot}$  to confirm that the selected clumps are at massive end of the sample. The PA are not included in the analyses because of their uncertain association with H II regions. AS HMSCs are split into S-type and O-type AS HMSCs to highlight the impacts of H II regions as mentioned in Sect. 4.4.

<sup>4</sup> Spearman’s correlation is used for assessing how well the relations between two variables could be described by a monotonic function. The value range of the correlation coefficient is from  $-1$  to  $1$ . A coefficient of  $1$  or  $-1$  means that the variable is a perfect monotonic function of the other variable.

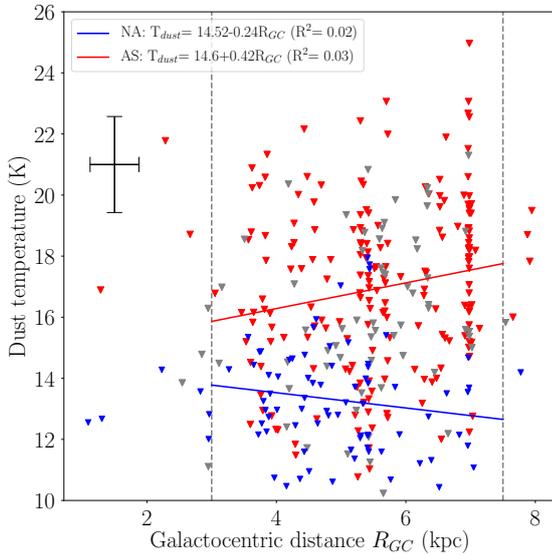
### 6.1. Distance bias

Our samples of HMSCs cover a broad range of distances, from 1–13 kpc, which could be a bias when claiming the differences in physical properties of different HMSCs. Eight physical properties and their relations with distance are shown in Fig. 10. The quantities  $T_{\text{dust}}$ ,  $L/M$  ratio, and eccentricity have a flat relation with distance. Number density  $n_{\text{H}_2}$  and probably also  $N_{\text{H}_2}$  show a decreasing trend, whereas  $M_{\text{clump}}$ ,  $L_{\text{clump}}$ , and size show an increasing trend. The possible reason is that our HMSCs have a similar angular size ( $20$ – $40''$ ) and column density ( $10^{22}$ – $10^{23} \text{ cm}^{-2}$ ), while their distances cover a broad range.

[Baldeschi et al. \(2017\)](#) studied the distance bias when deriving physical properties by fitting SED to Hi-GAL data. They find that the physical properties of the clump, such as temperature, the contribution of inter-core emission in the clumps, and the fractions of starless clumps to protostellar or all clumps, could change with distance. The fractions of HMSCs to massive ATLASGAL clumps in different distance ranges are shown in Fig. 11. The decreasing trend of starless fraction with distance coincides with [Baldeschi et al. \(2017\)](#), who explained it as the combined effects of the unresolved clumps (similar angular size) and their embedded multiple cores (protostellar but also starless). The HMSCs located at large distance are at the large size end (about 1 pc) of all HMSCs and the emission of embedded compact core(s) is more easily confused with the emission of diffuse gas, leading to the error in the determination of physical parameters. We roughly separate HMSCs into three distance ranges, which are  $<4$  kpc,  $4$ – $8$  kpc, and  $>8$  kpc according to their distributions. The HMSCs in these distance ranges should be individually considered when comparing the properties of the clump, which show a consistent increasing or decreasing change among distance.



**Fig. 8.** Galactic distributions for different types of HMSCs. *Panel a:* top view of HMSCs in the Galaxy. The red, yellow, and blue dots represent AS, PA, and NA HMSCs, respectively. The size of the dot increases with  $T_{\text{dust}}$  of HMSCs. The gray region represents  $|l| < 3^\circ$ . The typical error in position is shown as a white cross. *Panels b and c:* latitude and longitude distributions for bubble (see Kendrew et al. 2012; Hou & Gao 2014) and different types of HMSCs. The green dashed lines in *panel c* indicate the tangent directions of the spiral arms, which are shown as white dashed lines in *panel a*.



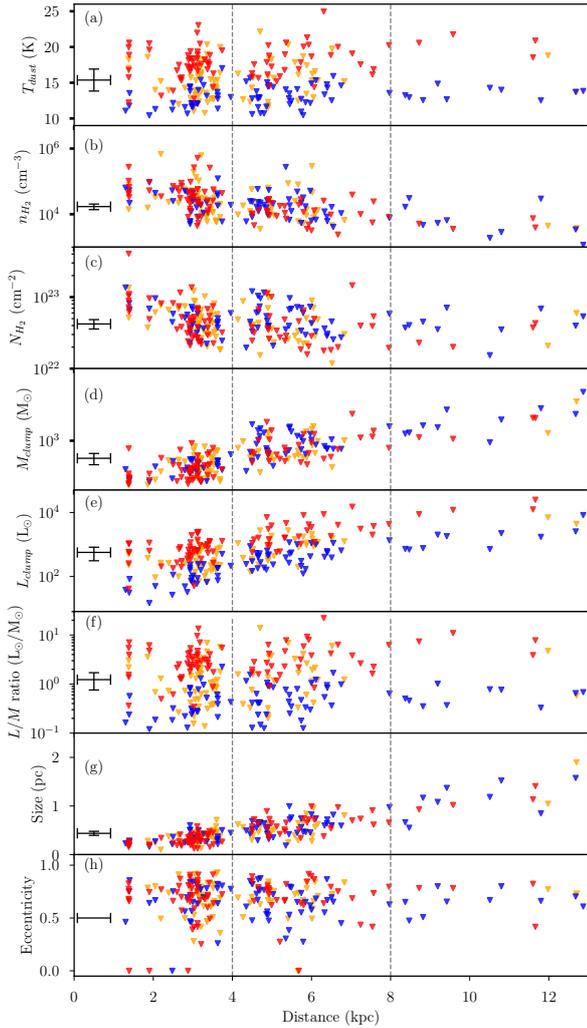
**Fig. 9.** Galactocentric distance  $R_{\text{GC}}$  and  $T_{\text{dust}}$  of HMSCs. Red, gray, and blue dots represent AS, PA, and NA HMSCs, respectively. The typical error is shown as a black cross. The HMSCs used for performing regression and Spearman's correlations are limited to 3–7.5 kpc. The red and blue solid lines represent the results of linear regressions of AS and NA.

Firstly, we perform Kolmogorov-Smirnov (K-S) tests to the physical properties with a flat relation with distance, which are  $T_{\text{dust}}$ ,  $L/M$  ratio, and possibly also  $N_{\text{H}_2}$ . If the properties have a

significant increasing or decreasing trend with distance, it could become a bias for K-S tests. The K-S test conclusions that two considered groups (AS and NA) are different could be just due to the different distance distributions. K-S tests are performed to three combinations, which are (1) S-type AS and NA, (2) S-type AS and O-type AS, and (3) O-type AS, and NA. When K-S test  $p$  value is smaller than 0.05, we reject the null hypothesis that the two samples are from the same distributions. The results of K-S tests shown in Table 6 suggest that  $T_{\text{dust}}$  and  $L/M$  ratio are significantly different for S-type AS, O-type AS, and NA, while  $N_{\text{H}_2}$  is similar for the different types of HMSCs.

For those properties with a consistent increasing or decreasing trend with distance, we need to carry out the K-S test within the same distance ranges. However, the small size of the sample in a certain distance range decreases the K-S test reliability, for example, only about 20 HMSCs are in the distance range of  $>8$  kpc. Therefore, we define a dimensionless factor,  $\delta_{1,2} = \sqrt{2} \times (M_1 - M_2) / \sqrt{\sigma_1^2 + \sigma_2^2}$ , where  $M_1$  and  $M_2$  are the median values while  $\sigma_1$  and  $\sigma_2$  are the standard deviations of the corresponding number distributions of properties. The numbers 1 and 2 are the labels for two types of HMSCs. This factor shows the ratio of median value difference to the combined standard deviations of two distributions, which allows for a simplified comparison between two types of HMSCs. For example, if two Gaussian distributions have the same profile but with a shift of standard deviation  $\sigma$  in Gaussian center, the resulted  $\delta$  is equal to 1, indicating the difference at  $1\sigma$  level.

The number distributions of eight properties are presented in Figs. 12–14 and Figs. C.1–C.5 (see Appendix C). Generally, only

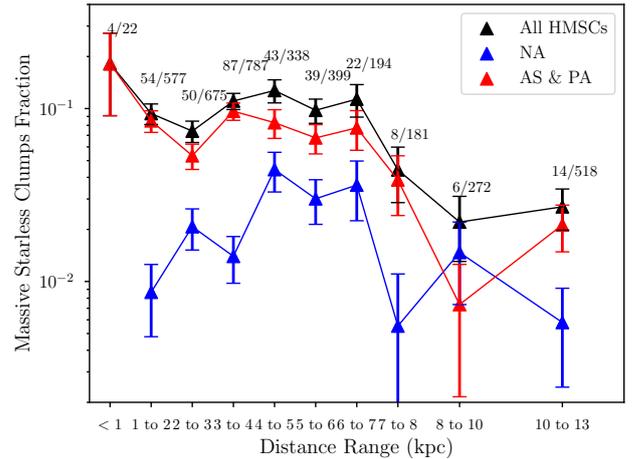


**Fig. 10.** Basic properties and distance. *a–h*: dust temperature ( $T_{\text{dust}}$ ),  $\text{H}_2$  number density ( $n_{\text{H}_2}$ ),  $\text{H}_2$  column density ( $N_{\text{H}_2}$ ), mass ( $M_{\text{clump}}$ ), luminosity ( $L_{\text{clump}}$ ), luminosity-to-mass ratio ( $L/M$ ), size, and eccentricity ( $e_{\text{clump}}$ ) of HMSCs and their distances. The red, orange, and blue triangles indicate S-type AS, O-type AS, and NA HMSCs. The gray dashed lines indicate the distances of 4 and 8 kpc.

$T_{\text{dust}}$ ,  $L_{\text{clump}}$ , and  $L/M$  differences keep on the order of  $1\text{--}2\sigma$  in each distance range, which could be seen in Fig. 15 by the  $\delta$  values of S-type AS and NA ( $\delta_{\text{S-AS, NA}}$ ) and that of O-type AS and NA ( $\delta_{\text{O-AS, NA}}$ ) in three ranges of distance. If a physical property is different under the effect of the H II regions, it should show a consistent increasing or decreasing trend from S-type AS to O-type AS and then NA owing to the gradually weaker impacts of the H II regions. If we set the criteria of significant difference as  $\delta$  larger than 1 or K-S tests suggest that the distributions are different, only  $T_{\text{dust}}$ ,  $L_{\text{clump}}$ , and  $L/M$  show significant differences between different types of HMSCs. The AS types have a higher  $T_{\text{dust}}$ ,  $L_{\text{clump}}$ , and  $L/M$  ratio than NA. When we set the criteria of similarity as  $\delta < 0.5$  and there is an inconsistent trend of parameter changing from S-type AS to O-type AS and then to NA,  $n_{\text{H}_2}$ ,  $N_{\text{H}_2}$ ,  $M_{\text{clump}}$ , size, and eccentricity are similar between different HMSCs.

## 6.2. Difference in properties

The properties with significant difference between different HMSCs, which are  $T_{\text{dust}}$ ,  $L/M$  ratio, and  $L_{\text{clump}}$ , are described



**Fig. 11.** Fraction of HMSCs. The fraction is calculated by the number ratio of HMSCs to massive ATLASGAL clumps in each distance range, as shown in the numbers in figure. The sample of ATLASGAL clumps is taken from the catalog in Csengeri et al. (2014) and the clump distance is taken from Urquhart et al. (2018). ATLASGAL clumps with a  $870\ \mu\text{m}$  peak intensity higher than  $0.5\ \text{Jy beam}^{-1}$  are extracted to be in line with the column density criterion for massive star formation in Sect. 2.1. The black, red, and blue lines indicate the changes of fractions with distance.

in detail in this section. Other properties that are similar for AS and NA are briefly described and more detailed descriptions for these are given in Appendix C.

**Dust temperature  $T_{\text{dust}}$ .** The changes of  $T_{\text{dust}}$  with distance are very flat according to Fig. 10 and therefore we could compare  $T_{\text{dust}}$  without consideration of distance factor. The K-S tests show that  $T_{\text{dust}}$  distributions are different between S-type AS, O-type AS, and NA. From S-type AS to NA, a decreasing trend of  $T_{\text{dust}}$ , from about 18–12 K, is shown in Fig. 12. The  $T_{\text{dust}}$  difference is larger than the typical  $T_{\text{dust}}$  error in the SED-fitting process, which is about 2–3 K. Furthermore,  $\delta$  values indicate that the differences are statistically significant, from  $1\sigma$  level between O-type AS and NA to  $2\sigma$  level between S-type AS and NA. The results show the probably stronger external heating effect of nearby H II regions on S-type AS. We discuss this in more detail in Sect. 9.1.

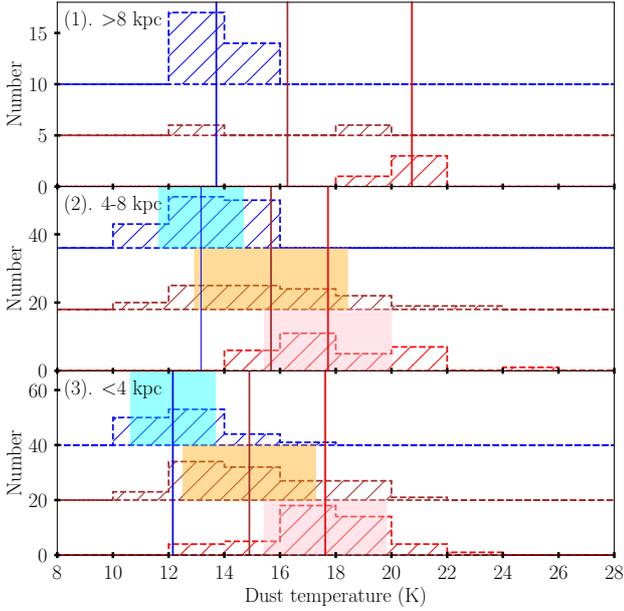
**Clump luminosity-to-mass ratio  $L/M$ .** The  $L/M$  ratio has a flat relation with distance in Fig. 10, thus we could directly compare it. The K-S tests suggest that their distributions between AS and NA HMSCs are different. The  $L/M$  ratios decrease from about  $3\ L_{\odot}/M_{\odot}$  to  $1\ L_{\odot}/M_{\odot}$  and then  $0.4\ L_{\odot}/M_{\odot}$  for S-type AS, O-type AS, and NA, respectively. The typical error of  $L/M$  ratio is less than 40%, thus the difference is larger than error in the calculation. Meanwhile,  $\delta$  values indicate that the difference is on the order of  $1\sigma$  between O-type AS and NA, while it increases to  $2\sigma$  between S-type AS and NA, showing the likely stronger impacts of H II regions.

The  $L/M$  ratio is not an independent parameter if its calculation is related with far-IR/(sub)millimeter SED fitting because of the similar gray-body functions used in the fitting process (Pitts et al. 2019). The  $L/M$  ratios are in proportion to the ratio of integrated flux of gray body function to  $N_{\text{H}_2}$  (see Eqs. (4) and (7) in Y17), while  $N_{\text{H}_2}$  is derived by fitting the gray-body function. A higher  $T_{\text{dust}}$  results in a higher  $L/M$  ratio. In the regime of 12 K to 28 K, an increment of 3 K to 4 K in  $T_{\text{dust}}$  could increase  $L/M$  about 2–3  $L_{\odot}/M_{\odot}$  for a clump with a mass of  $400\ M_{\odot}$  that is the

**Table 6.** K-S test for S-type AS, O-type AS, and NA HMSCs.

Properties	S-type AS vs. NA				S-type AS vs. O-type AS				O-type AS vs. NA			
	K-S result	K-S $D^{(1)}$	K-S $p^{(1)}$	$\delta$	K-S result	K-S $D$	K-S $p$	$\delta$	K-S result	K-S $D$	K-S $p$	$\delta$
$T_{\text{dust}}$	Different	0.848	$4.44 \times 10^{-16}$	2.5	Different	0.48	$1.45 \times 10^{-8}$	1	Different	0.487	$1.78 \times 10^{-8}$	1.14
$N_{\text{H}_2}$	Same	0.192	0.0989	-0.285	Same	0.0956	0.829	-0.197	Same	0.125	0.556	-0.0862
$L/M$	Different	0.847	$4.44 \times 10^{-16}$	2.51	Different	0.476	$1.98 \times 10^{-8}$	0.97	Different	0.461	$1.39 \times 10^{-7}$	1.11

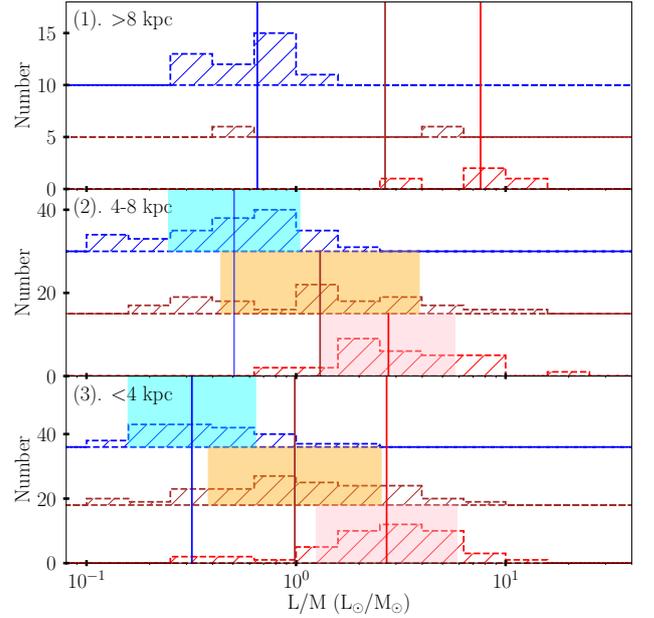
**Notes.** Three combinations are used for the test. <sup>(1)</sup>  $D$  and  $p$  are the K-S statistic and the two-tailed  $p$  value of K-S test, respectively.



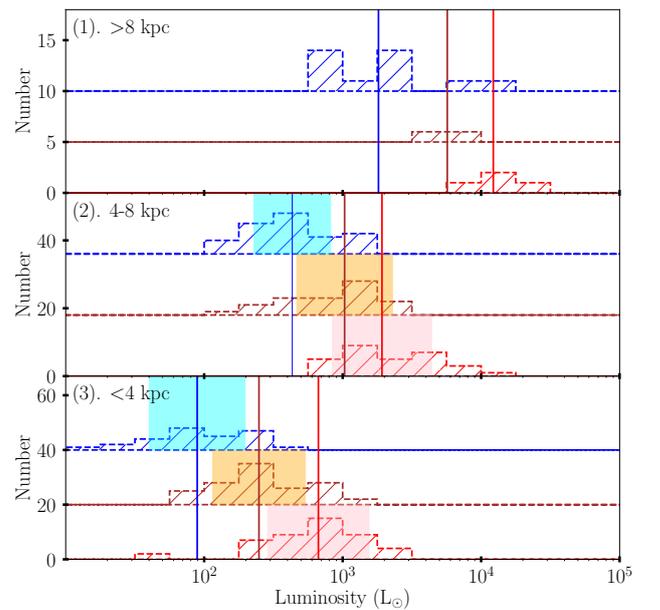
**Fig. 12.** Number distributions of  $T_{\text{dust}}$  of HMSCs. Three panels from bottom to top: number distributions in the distance ranges of <4 kpc, 4–8 kpc, and >8 kpc, respectively. The red, dark brown, and blue histograms represent the number distributions for S-type AS, O-type AS, and NA, respectively. The solid lines indicate their median values. The pink, brown, and cyan filled color regions have a full width of double standard deviation  $2\sigma$  and center at median value of corresponding types of HMSCs.

median mass of the sample. Thus, the  $L/M$  ratio difference could be well explained by the  $T_{\text{dust}}$  difference.

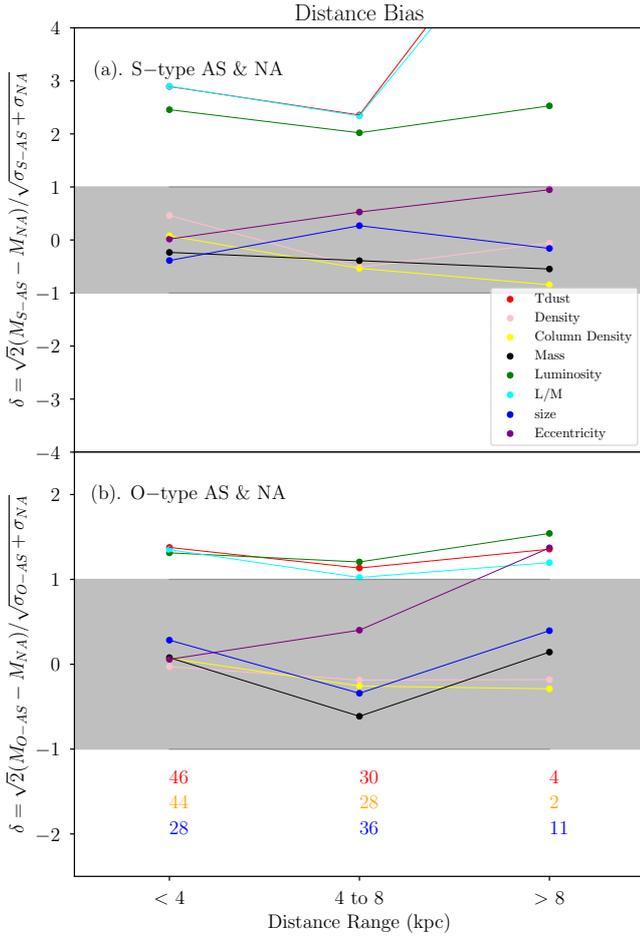
With the single-dish observed thermometer molecule  $\text{CH}_3\text{C}_2\text{H}$ , Molinari et al. (2016) suggest that  $L/M$  ratio could be an indicator of star formation evolutionary stage for massive clumps. These authors define three  $L/M$  intervals, which are  $L/M \lesssim 1 L_\odot/M_\odot$ ,  $1 L_\odot/M_\odot \lesssim L/M \lesssim 10 L_\odot/M_\odot$  and  $L/M \gtrsim 10 L_\odot/M_\odot$ , respectively. These three intervals correspond to three evolutionary stages from (1) very early stages when only low-mass YSOs are forming; to (2) the stage in which relatively low-mass YSOs build up the luminosity of the clump; and then to (3) the stage in which massive stars are forming and warming up the inner clump gas. Most of S-type AS have a  $L/M$  ratio larger than the first interval. Some of these could even arrive  $10 L_\odot/M_\odot$ . The  $L/M$  ratios of O-type AS are around first interval while most of NA are below the first interval. Recently, Molinari et al. (2019) connected the mass fraction locked in the dense cores of massive clumps to the  $L/M$  ratios of clumps. The expression  $L/M \lesssim 1 L_\odot/M_\odot$  classifies the protoclusters with a dense core mass fraction  $\lesssim 0.1$ , whereas  $L/M \gtrsim 10 L_\odot/M_\odot$  excludes dense core fraction  $\lesssim 0.1$ . The significant  $L/M$  change between AS and



**Fig. 13.** Number distributions of  $L/M$  ratio of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for  $L/M$  ratio.



**Fig. 14.** Number distributions of  $L_{\text{clump}}$  of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for  $L_{\text{clump}}$ .



**Fig. 15.** Distance bias. *Panels a and b:*  $\delta$  values for the combination of S-type AS and NA, and the combination of O-type AS and NA HMSCs, respectively. All sources are grouped into three distance bins, which are <4 kpc, 4–8 kpc, and >8 kpc. Dots with different colors mean different properties of HMSCs. The red, orange, and blue numbers in (b) are the numbers of S-type AS, O-type AS, and NA HMSCs in the corresponding distance ranges. The gray regions are  $|\delta| \leq 1$ , where we suggest that the differences are insignificant.

NA indicates that it is uncertain to classify the early evolutionary stages of massive clumps with the only observed  $L/M$  ratio. We discuss this point in Sect. 9.1.

**Clump luminosity  $L_{clump}$ .** The  $L_{clump}$  significantly increases with distance and therefore we could not directly compare it. In the certain distance bins,  $\delta$  values show that S-type AS have a larger  $L_{clump}$  than NA at  $2\sigma$  level while the difference between O-type AS and NA is at  $1\sigma$  level, suggesting the probably different strength of external heating of H II regions on HMSCs between S-type AS and O-type AS.

**Other properties.** The  $N_{H_2}$ ,  $n_{H_2}$ ,  $M_{clump}$ , size, and eccentricity are similar for AS and NA. Their  $\delta$  values generally indicate that the difference between AS and NA is less than  $0.5\sigma$ . The gradual changes of properties from S-type AS to O-type AS and then NA are not clear for these properties. Owing to the selection criteria of S-type AS, which are the features of being photodissociated/photoionized or being in a compressed structure, we probably could expect that S-type AS could be denser than NA. The results show that the difference is not robust. It could be due to the beam dilution effects that narrow the density differences (see Appendix C).

## 7. Column density and temperature structure

### 7.1. Calculation

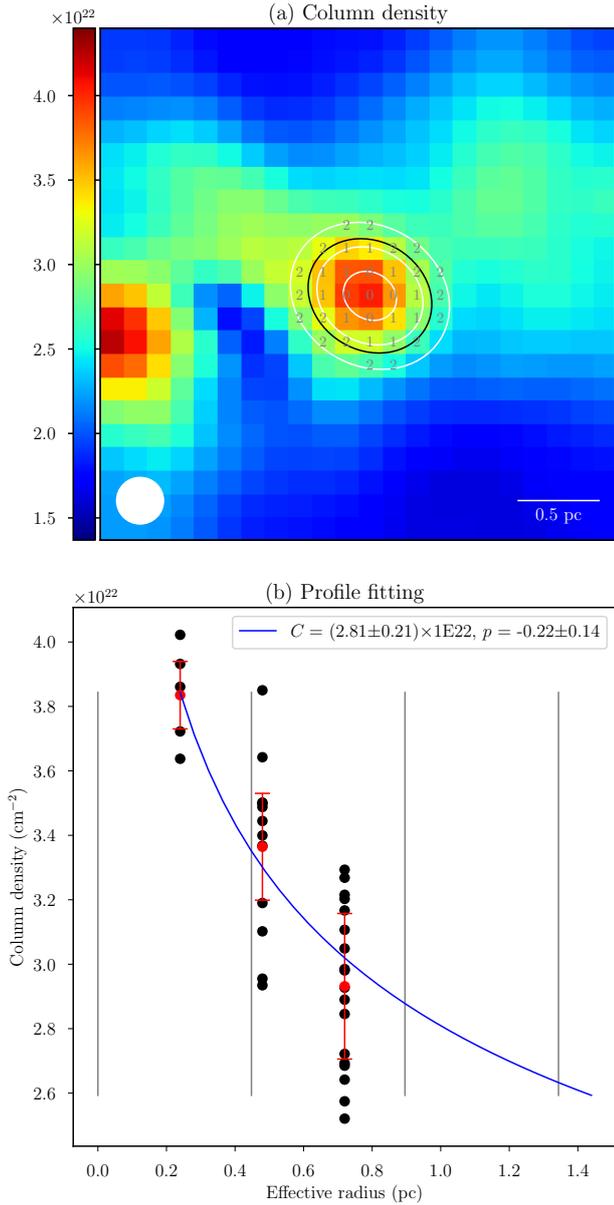
In this section, we compare the differences between the structures of different HMSCs. An isothermal, spherically symmetric prestellar clump with gravitational contraction is suggested have a clump radial density profile consisting of two regions: an inner core  $n_{H_2} = \text{constant}$  and envelope  $n_{H_2} \propto r^{-2}$  (Vázquez-Semadeni et al. 2019). The corresponding radial  $N_{H_2}$  profiles are  $N_{H_2} \propto r$  and  $N_{H_2} \propto r^{-1}$ , respectively. Owing to the limited resolution, a flat central region and truncated radius are usually set when fitting the  $N_{H_2}$  profile to the observational data (Juvela et al. 2018; Tang et al. 2018). The Galactic cold cores  $N_{H_2}$  structures investigated with *Herschel* data indicate that radial  $N_{H_2}$  distribution follows a power law of  $N_{H_2} \propto r^{-1}$  in spite of the broad variety of clump morphology, which suggests the universality of the  $r^{-1}$  profile for cold cores with gravitational contraction (Juvela et al. 2018; Li 2018).

We compare  $N_{H_2}$  or  $T_{dust}$  radial profile of different HMSCs with PPMAP data. Beam size is a crucial factor that could affect the clump structures we observed. A larger beam smooths the steep profile to make it flatter. Furthermore, beam size determines the lower size limit of the flat central region. The resolution of PPMAP Hi-GAL data, which is  $12''$ , is more likely to detect the envelope rather than the inner core of clumps at a distance of several kpc. We simply fit a single power law radial  $N_{H_2}$  profile  $N_{H_2} = C \times r^{p_{N_{H_2}}}$  to the PPMAP  $N_{H_2}$  data rather than with a more complex function.

Firstly, only AS and NA HMSCs with a mass larger than  $100 M_{\odot}$  are considered to ensure the selection of massive sources. The HMSCs with an angular size  $<24''$  and distance  $>8.34$  kpc are removed to mitigate the resolution issue. Secondly, we separate the PPMAP pixels into a series of concentric elliptical shells with a shell width of  $6''$  in minor axis, which is the pixel size and the half beam size of PPMAP data. The aspect ratios and position angles of each elliptical shell of HMSCs are the same as the ellipses outlining the shapes of HMSCs in Y17. The radius of the largest shell is slightly larger than the clump size. For the pixels crossing the edges of the shells, we simply assign the associated shells according to the centers of these pixels. Thirdly, the median  $N_{H_2}$  in one shell is used as the  $N_{H_2}$  in this shell whereas the 25th and 75th percentiles of the pixel  $N_{H_2}$  of the shell are used as the error ranges of  $N_{H_2}$  for this shell. Finally, the median  $N_{H_2}$  in different shells are fitted with a power-law profile  $N_{H_2} = C \times r^{p_{N_{H_2}}}$ . An example of  $N_{H_2}$  profile fitting is shown in Fig. 16. Similar operations are carried out on PPMAP  $T_{dust}$  images to look at the temperature profile and get a power-law index  $p_{T_{dust}}$ .

### 7.2. Properties and results

The  $N_{H_2}$  and  $T_{dust}$  power-law indexes ( $p_{N_{H_2}}$  and  $p_{T_{dust}}$ ) are shown in Figs. 17a and b, respectively. The median value of  $p_{N_{H_2}}$  is  $-0.17$ , which is much larger than the value of  $-0.85$  in Juvela et al. (2018). There are several reasons for this difference. The clumps/cores they studied have a more compact morphology with a typical size of  $0.075$  pc compared to  $\approx 0.5$  pc for our HMSCs. Their beam size is equal to  $0.03$  pc with a typical distance of  $250$  pc, but our beam size is  $0.4$  pc with a typical distance of  $4$  kpc. Our relatively poor physical resolution could create a higher  $p_{N_{H_2}}$  value. Furthermore, their fitting field is larger than the clump scale, which includes the weaker background emission that results in a smaller  $p_{N_{H_2}}$  value. Our



**Fig. 16.** Column density profile fitting for HMSC G010.9823–0.3677. *Panel a:* PPMAP  $N_{\text{H}_2}$ . The black and white ellipses represent the clump and its shells used for extracting pixels. The gray numbers denote the corresponding shells of the pixels. The beam of PPMAP is shown in bottom left. *Panel b:* profile fitting. The black and red dots indicate the PPMAP  $N_{\text{H}_2}$  and their median values in the corresponding shells. The error of  $N_{\text{H}_2}$  value of each shell is estimated by the 25th and 75th percentiles of  $N_{\text{H}_2}$  of each shell, which is shown by the red error bar. The blue line is the fitting result for median values. The gray lines show the triple beams of PPMAP.

$p_{N_{\text{H}_2}}$  values are very similar to the results of [Giannetti et al. \(2013\)](#). These authors simply fitted a power-law profile to 1.2 mm continuum emission of high-mass clumps with a beam size of  $20''$ . They find a typical  $p_{N_{\text{H}_2}}$  value of  $-0.2$  to  $-0.4$  for high-mass quiescent clumps, whereas values smaller than  $-1$  for high-mass star-forming clumps. Their results indicate that our fitting process is reliable at a similar resolution of  $12''$ . We note that our main goal is not to derive a specific and exact value of  $p_{N_{\text{H}_2}}$ , but to search for a possible difference between AS and NA HMSCs.

The median values of  $p_{N_{\text{H}_2}}$  are  $-0.27$ ,  $-0.19$ , and  $-0.12$  for S-type AS (30), O-type AS (26), and NA (34), respectively.

Their standard deviations of distributions are 0.13, 0.1, and 0.05, respectively. The  $\delta$  values suggest that S-type AS have a smaller value of  $p_{N_{\text{H}_2}}$  compared with that of O-type AS and NA at  $0.7\sigma$  and  $1.5\sigma$  level, respectively, showing the significant differences between S-type AS and NA. All HMSCs with a  $p_{N_{\text{H}_2}}$  smaller than  $-0.3$  are AS HMSCs except for one NA. A  $\chi^2$  test to AS and NA with  $p_{N_{\text{H}_2}}$  larger and smaller than  $-0.2$  results in 99% to reject the hypothesis of independence between the association of H II region and  $p_{N_{\text{H}_2}}$  value. Therefore, a smaller  $p_{N_{\text{H}_2}}$  value, which suggests a steeper  $N_{\text{H}_2}$  profile, is closely related with the association of H II region.

The results of  $T_{\text{dust}}$  profile shown in Fig. 17b are very similar to the results of  $N_{\text{H}_2}$  profile. The S-type AS have  $T_{\text{dust}}$  profiles steeper than those of NA HMSCs; the median values and standard deviations are 0.021 and 0.01 for S-type AS, 0.016 and 0.01 for O-type AS, and 0.01 and 0.05 for NA HMSCs. The  $\delta$  values indicate the differences between AS and NA are at about  $1\sigma$  level.

In brief, AS HMSCs have steeper  $N_{\text{H}_2}$  and  $T_{\text{dust}}$  profiles than those of NA HMSCs and the difference between AS and NA types enlarges with a more drastic feedback of the H II regions.

### 7.3. Bias

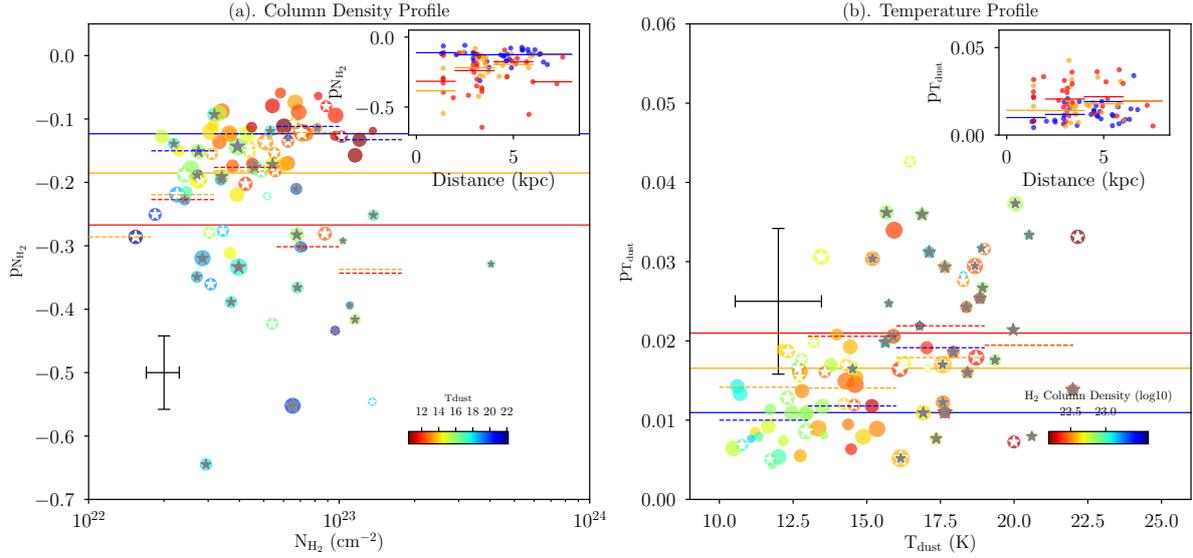
The distance bias could impact our comparison of  $p_{N_{\text{H}_2}}$  values between HMSCs with different distances. The small panel in the top right of Fig. 17a shows the trend that  $p_{N_{\text{H}_2}}$  increase with distance. A larger distance indicates a larger physical size of HMSCs when the angular size is similar (about  $30''$ ), thus the profile of the most distant clump could be less easily resolved and more easily mixed with the surrounding density distribution. To check how distance bias affects our results, we separate HMSCs into a series of distance bins with a width of 2 kpc. The results indicate a consistent trend in each distance bin that AS occupy the lowest  $p_{N_{\text{H}_2}}$  value end compared with NA. The panel in the top right of Fig. 17b suggests the trend that AS occupy the uppermost  $p_{T_{\text{dust}}}$  value end is also consistent in each distance bin.

The number of shells could affect the  $p_{N_{\text{H}_2}}$  value because more shells in a clump means that the structure of the clump is better resolved and could result in a steeper observed profile. Of the 90 HMSCs for which we estimated the profile, only six clumps have four shells while the other HMSCs have three shells. Furthermore, these six clumps with four shells occupy the high  $p_{N_{\text{H}_2}}$  value end ( $-0.2$  to  $0$ ), which means that their profiles are flatter. Thus, we suggest that the effects of shell number do not significantly change our overall results.

Another factor changing the observed profile is the nature of Hi-GAL PPMAP data. It has been shown in Sect. 3.3.2 that Hi-GAL PPMAP overestimates  $T_{\text{dust}}$  and underestimate  $N_{\text{H}_2}$  for the clumps at the coldest end ( $<14$  K). For starless clump, its center is normally expected to have a lower  $T_{\text{dust}}$  compared to the envelope. The Hi-GAL PPMAP data could overestimate  $T_{\text{dust}}$  and underestimate  $N_{\text{H}_2}$  for the center of starless clump, making the profile flatter.

### 7.4. Links between L/M and steeper profile

A steeper density profile indicates that the clump is more centrally peaked. Observational studies have shown that the density structure of the clump becomes more and more centrally peaked with the evolution from the starless stage to protostellar stage ([Butler & Tan 2012](#); [Tang et al. 2018](#)). During the formation of protostars, the heating effect of inner protostars and the collapse/accretion make the clump warmer and more compact,



**Fig. 17.** Power-law index of  $N_{\text{H}_2}$  and  $T_{\text{dust}}$  profile. The whole  $T_{\text{dust}}$  and  $N_{\text{H}_2}$  of the clump are taken from updated Y17 properties in this paper. The  $p$  values are derived from Hi-GAL PPMAP data. *Panels a and b:* column density and dust temperature power-law index  $p$ , respectively. The crosses represent the typical errors. The dots with gray stars, white stars, and without stars represent the S-type AS, O-type AS, and NA HMSCs, respectively. The red, orange, and blue solid lines represent the median values for S-type AS, O-type AS, and NA, respectively. The red, orange, and blue dashed lines represent the median values for S-type AS, O-type AS, and NA in the corresponding column density or temperature bins. Sub-figures in the top right show the power-law index and its relation with distance. The red, orange, and blue dots represent S-type AS, O-type AS, and NA, respectively. The red, orange, and blue short lines show the median values in different distance bins for the corresponding types of HMSCs.

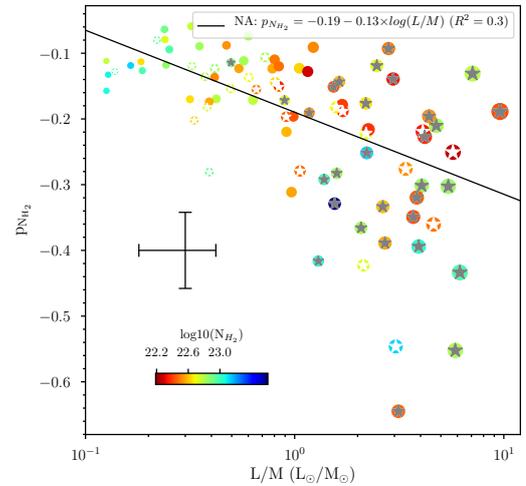
thus a positive correlation between  $L/M$  and centrally peaked properties of the clump is expected.

The HMSCs impacted more deeply by nearby H II region are compressed and heated more deeply by the ionized gas, thus the clumps are warmer and more centrally peaked. A weak positive relation between  $T_{\text{dust}}$  and steepness of profile can be found in Fig. 17a by the color of dots and the  $p_{N_{\text{H}_2}}$  value. There is also a weak trend that  $p_{N_{\text{H}_2}}$  values become smaller with increasing  $N_{\text{H}_2}$  as shown in Fig. 17a. The median  $N_{\text{H}_2}$  and  $T_{\text{dust}}$  are larger for the most centrally peaked AS ( $p_{N_{\text{H}_2}} < -0.3$ ,  $6.7 \times 10^{22} \text{ cm}^{-2}$  and 18.3 K) than those of other AS ( $4.7 \times 10^{22} \text{ cm}^{-2}$  and 15.7 K) and NA ( $4.5 \times 10^{22} \text{ cm}^{-2}$  and 12.9 K).

A similar positive correlation between  $p_{N_{\text{H}_2}}$  and  $L/M$  of HMSCs is shown in Fig. 18. The reason might be the tight relation between  $T_{\text{dust}}$  and impacts of H II region, as mentioned, rather than star formation activities inside HMSCs. The strengthened heating effect working on AS leads to a higher  $T_{\text{dust}}$  and  $L/M$  ratio. This result indicates that even for the  $p_{N_{\text{H}_2}}$ , the environment of the clump should be checked carefully in advance when these values are used as evolutionary probes for early massive clumps.

## 8. Turbulence and virial status

In this section, we compare the turbulence level in AS and NA and seek to understand whether turbulence plays a dominant role in shaping massive star formation. Recent modeling results based on the global hierarchical collapse in molecular clouds argue that chaotic, multiscale infall due to gravity could dominate at all scales (Vázquez-Semadeni et al. 2019). By analyzing a sample of 70  $\mu\text{m}$  quiet clumps, Traficante et al. (2020) propose a scenario that is a continuous interplay between turbulence and gravity. The turbulence creates multiscale structure and then gravity becomes dominant when the critical density threshold is reached. Based on the hypothesis that mass distributions of

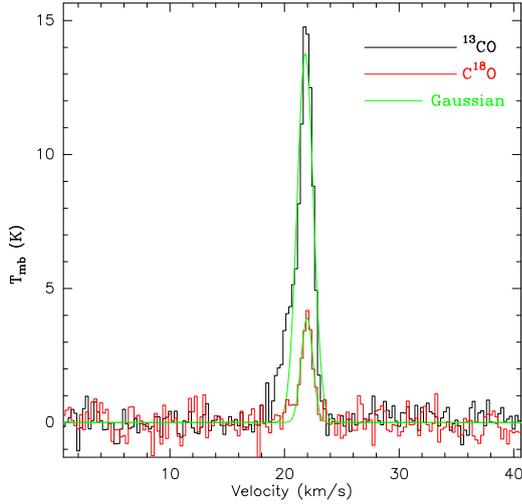


**Fig. 18.** Plot of  $L/M$  and  $p_{N_{\text{H}_2}}$ . The meanings of different dots are similar to Fig. 17a. The colors of dots indicate  $N_{\text{H}_2}$  of HMSCs. The size of the dot is proportional to the  $T_{\text{dust}}$ . The cross represents the typical error. The black line indicates the linear regression to the data.

stars result from turbulent fragmentation, Padoan et al. (2019) propose a scenario in which large, converging, inertial flows in the environments of supersonic turbulence assemble massive stars. These studies shed light on the particular importance of understanding the sources that drive turbulence, their relative importance at the clump scale and the impacts of turbulence on regulating star formation (Lo et al. 2015).

### 8.1. Mach number estimation

To show the turbulence level in HMSCs, we use SEDIGISM data to estimate the Mach number. SEDIGISM survey mapped the inner Galactic plane in the  $J = 2-1$  rotational transition of

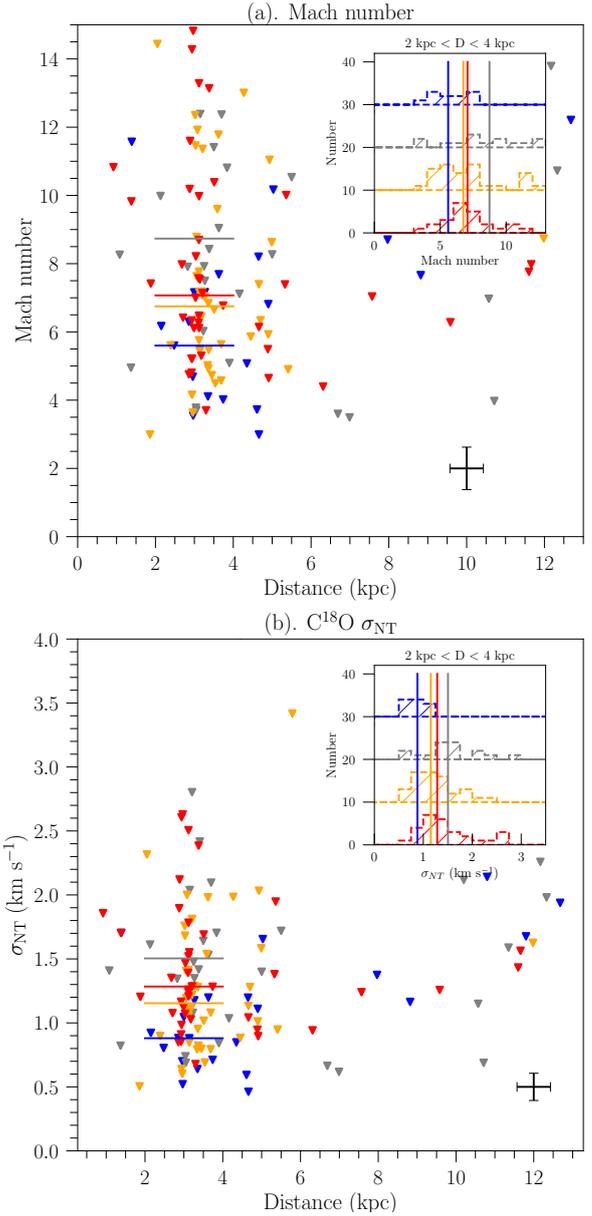


**Fig. 19.** Example of Gaussian fitting to  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$   $J = 2-1$  of the HMSC G015.2169–0.4267.

$^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$  with a spatial and a velocity resolution of about  $30''$  and  $0.25 \text{ km s}^{-1}$ , respectively. We extract the average spectra in one beam centered on the HMSC position from a spectral cube and then fit it with Gaussian using the CLASS package<sup>5</sup>. An example of Gaussian fitting to  $^{13}\text{CO}$  and  $\text{C}^{18}\text{O}$   $J = 2-1$  spectra of the HMSC G015.2169–0.4267 is presented in Fig. 19. We only use  $\text{C}^{18}\text{O}$   $J = 2-1$  rather than  $^{13}\text{CO}$   $J = 2-1$  to estimate the Mach number because  $\text{C}^{18}\text{O}$  is likely to be optically thin and traces denser regions.  $\text{C}^{18}\text{O}$  has a lower abundance than  $^{13}\text{CO}$  with a  $^{13}\text{CO}/\text{C}^{18}\text{O}$  abundance ratio of  $\approx 5-8$  (Milam et al. 2005; Mattern et al. 2018; Paron et al. 2018).

The velocity dispersion  $\sigma_v$  is combined by a thermal component  $\sigma_{\text{Therm}} = (kT_{\text{ex}}/m_{\text{C}^{18}\text{O}})^{1/2}$  and nonthermal component  $\sigma_{\text{NT}}$  with equation  $\sigma_v^2 = \sigma_{\text{Therm}}^2 + \sigma_{\text{NT}}^2$ .  $\sigma_{\text{NT}}/\sigma_{\text{Therm}}$  for  $\text{C}^{18}\text{O}$  lines of HMSCs in our cases are in a range of 10 to 40, suggesting the dominance of turbulence in HMSCs. We set  $\sigma_{\text{NT}}$  as the turbulence flow velocity, thus the Mach number  $M = \sqrt{3}\sigma_{\text{NT}}/c_s$ , where  $\sqrt{3}$  is the correction from 1D to 3D and  $c_s = \sqrt{\gamma R_{\text{gas}}T/m_{\text{mol}}}$ . The quantities  $\gamma$ ,  $R_{\text{gas}}$ , and  $m_{\text{mol}}$  are the adiabatic index, gas constant, and molar mass of the gas, respectively. Orkisz et al. (2017) use a similar method to estimate the Mach number with  $^{13}\text{CO}$   $J = 1-0$  in the Orion B region. They assume that the opacity broadening, natural, and thermal widths of lines are negligible compared to turbulent broadening. We use  $\sigma_{\text{NT}}$  as the turbulent flow velocity rather than the velocity dispersion of spectra  $\sigma_v$  even if thermal widths make up only 2% to 10% width of the  $\text{C}^{18}\text{O}$   $J = 2-1$  lines in our HMSC sample. Opacity broadening is less important because  $\text{C}^{18}\text{O}$   $J = 2-1$  is more optically thin compared to  $^{13}\text{CO}$   $J = 1-0$ .

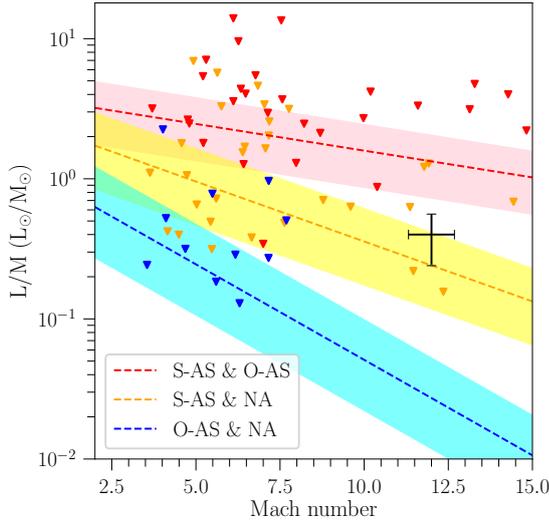
The Mach number, nonthermal component and their correlations with distance are shown in Fig. 20. As shown in Sect. 6.1, HMSCs at a larger distance have a larger physical scale  $r_{\text{pc}}$ . An increase in physical scale  $r_{\text{pc}}$  could increase  $\sigma_v$  according to Larson’s law  $\sigma_v \propto r_{\text{pc}}^{0.56 \pm 0.02} \text{ km s}^{-1}$  (Larson 1981). An increasing trend of  $\sigma_{\text{NT}}$  from 2 kpc to 13 kpc shown in Fig. 20b proves that the bias of distance exists when comparing  $\sigma_v$ . Mach numbers or nonthermal components of the full sample studied do not show a clear difference between AS and NA. The  $\delta_{S-AS,NA}$  value indicates a slight difference ( $0.1\sigma$ ) between S-type AS and NA in the full sample. To compare the Mach number more



**Fig. 20.** Turbulence. *Panels a and b:* Mach number and non-thermal component of HMSCs, respectively. The red, orange, gray and blue triangles represent S-type AS, O-type AS, PA, and NA, respectively. The corresponding median values for different HMSCs in the 2 to 4 kpc range are shown in short lines with corresponding colors. The histograms in top right show the number distributions in the range of 2 to 4 kpc and their median values.

reliably, we extract HMSCs at a distance of 2–4 kpc. More than half of  $\text{C}^{18}\text{O}$  studied HMSCs (85 in all 135 HMSCs, including 11 NA, 16 PA, 30 O-type AS, and 28 S-type AS) are included. The HMSCs in 2–4 kpc are better resolved and the numbers of different types of HMSCs are relatively statistically significant compared with other distance ranges with a width of 2 kpc. The median values of the Mach number in 2–4 kpc are 5.6, 6.7, and 7.1 for NA, O-type AS, and S-type AS, respectively. The quantity  $\delta \approx 1$  between S-type AS and NA, indicating that S-type AS have a Mach number statistically larger than that of NA. The turbulence level of HMSCs could be between supersonic regime (1.3–5) to hypersonic regime (5–10). The AS HMSCs are more likely to be hypersonic while NA types are more likely to be supersonic.

<sup>5</sup> <https://www.iram.fr/IRAMFR/GILDAS/doc/html/class-html/class.html>



**Fig. 21.** Turbulence and  $L/M$  ratios of HMSCs in 2 kpc to 4 kpc. Typical errors of  $L/M$  and Mach number are about 40% of  $L/M$  and 1–2, respectively. The uncertainties of  $T_{\text{dust}}$  and line width are propagated to the calculations of Mach number. The triangles with different colors are similar to Fig. 20. The blue, orange, and red dashed lines indicate the intervals of (1) O-type AS and NA, (2) S-type AS and NA, and (3) S-type AS and O-type AS. The cyan, yellow, and pink regions show the uncertainties of the boundaries. The red and blue lines roughly outline three regions: (i) low  $L/M$  and Mach number, (ii) intermediate  $L/M$  and Mach number, and (iii) high  $L/M$  and Mach number.

A positive correlation between  $\sigma_{\text{NT}}$  and  $L/M$  for prestellar and protostellar dense clumps has been found with  $\text{NH}_3$  lines of about 1000 clumps by Merello et al. (2019). These authors propose that it could be due to the additional energy injection from the embedded protostars. As shown in Fig. 21, NA, O-type AS and S-type AS HMSCs roughly occupy the (i) low Mach number –  $L/M$  regime, (ii) intermediate Mach number –  $L/M$  regime, and (iii) high Mach number –  $L/M$  regime, respectively. The boundaries are estimated by the algorithm of support-vector clustering (SVC<sup>6</sup>). Our identified boundaries are just tentative visualizations for different regimes rather than a rigorous calculation of the margin due to the small size of the sample (less than 70 HMSCs). The roughly positive relation between  $L/M$  and Mach number for our HMSCs between the range 2–4 kpc has a mechanism different from that in star-forming clumps because of the absence of the embedded protostars. The external heating and turbulent energy injection by H II regions are the most likely reasons for the positive relations and these two impacts of H II regions on HMSCs are probably interconnected with each other.

## 8.2. Virial status

We find that AS are more turbulent than NA. In this section, we analyze whether HMSCs are gravitationally bounded and search for the possible differences between AS and NA by estimating virial mass ratio,

$$R_{\text{virial}} = \frac{5}{\alpha\beta} \frac{r\sigma^2}{GM_{\text{clump}}}, \quad (1)$$

where  $\alpha$  and  $\beta$  are factors related with the eccentricity and density profile, respectively (Li et al. 2013);  $\alpha = (1 + a/3)/(1 + 2a/5)$ , where  $a$  is for the power-law index of number density

<sup>6</sup> <https://scikit-learn.org/stable/modules/svm.html>

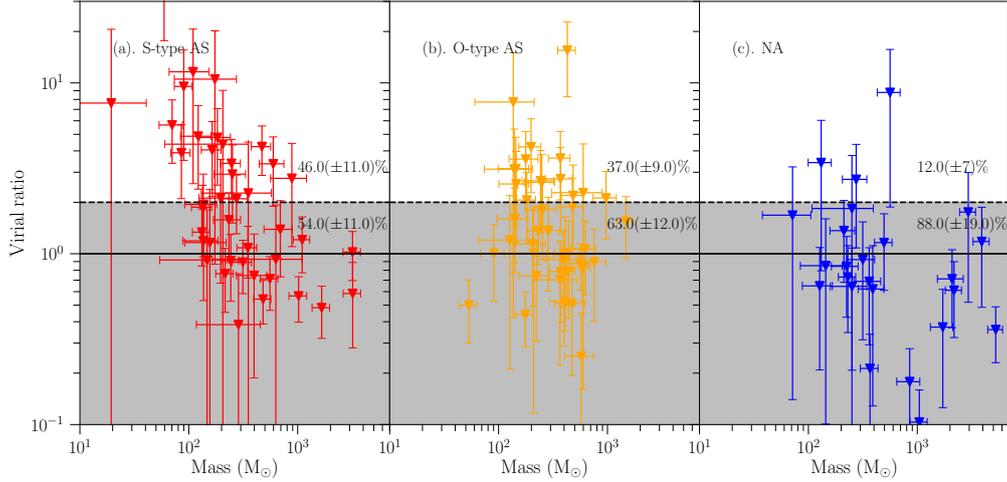
profile  $n_{\text{H}_2} \propto r^a$ ;  $\beta = \arcsine/e$  is the geometry factor determined by eccentricity  $e$ ; and  $a$  is approximately equal to  $p_{\text{N}_{\text{H}_2}} - 1$ . As shown in Sect. 7, AS HMSCs have a lower value of  $p_{\text{N}_{\text{H}_2}}$  than NA HMSCs, especially for the S-type AS HMSCs. The nature and limited resolution of PPMAP data could make us underestimate the real profile difference. To show the differences in the power-law index when calculating the virial mass ratio, we set a unique  $p_{\text{N}_{\text{H}_2}}$  value of  $-0.6$  for NA, which is the value for an isothermal clump in equilibrium (Li et al. 2013), whereas we set another value of  $-0.8$  for AS HMSCs. The difference of 0.2 is estimated by the median values in Fig. 17.

The virial ratio and clump mass for all  $\text{C}^{18}\text{O}$  studied HMSCs are shown in Fig. 22. The trend that virial ratios decrease with clump mass is well observed by other similar studies (Csengeri et al. 2017b; Wielen et al. 2018). The fraction of HMSCs with a virial ratio of less than 2 decreases from about 90% for NA to 50% for S-type AS, which may imply a different virial status between AS and NA. A virial ratio of less than 2 suggests a gravitationally bound structure for the clump (Bertoldi & McKee 1992). We note that the surface pressure and magnetic field are not included in our virial calculation. For HMSCs with a virial ratio  $\gg 1$ , these clumps could be unbound and transitory but still possible to be in bound status if we consider surface pressure or magnetic field. AS HMSCs are believed to be deeply impacted by nearby H II regions. The external pressure from H II regions could assist to confine the clumps. The  $\delta$  values indicate that the difference between S-type AS and NA for all sample is at  $0.9\sigma$  level. An Anderson-Darling test shows that the null hypothesis that AS and NA have the same distributions can be rejected ( $p$  values equal to 0.0018 for S-type AS and NA, and 0.015 for O-type AS and NA). To avoid the bias of distance, we show the virial ratio distribution for HMSCs in the 2–4 kpc range in Fig. 23, the median value for S-type AS, O-type AS, and NA are 2.0, 1.3 and 0.7, respectively. The  $\delta$  value between S-type AS and NA is about 1.5, indicating a more significant difference of virial ratios compared to full sample.

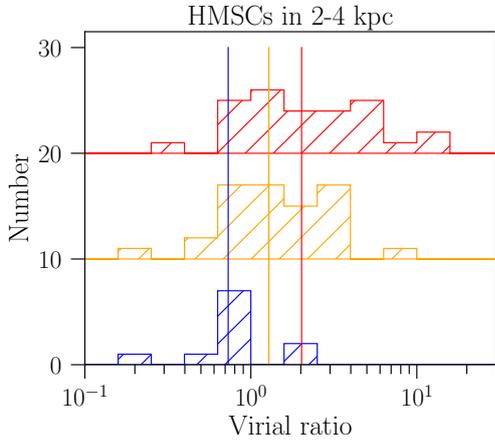
The median virial ratios of AS and NA are in the transition regime between virialized to subvirial, even if the virial ratios for AS are statistically higher than NA. Recent simulations of Rosen et al. (2019) on the collapse of subvirial ( $R_{\text{virial}} = 0.1$ ) and virialized ( $R_{\text{virial}} = 1$ ) turbulent massive cores (0.1 pc) suggest that the most massive stars likely form from subvirial massive cores by the rapid monolithic collapse without fragmentation, while other less massive stars likely form through slow collapse of highly turbulent cores with a larger number of fragments. We note that AS HMSCs are not isolated and are probably being compressed by H II regions. The pressure from H II regions is ignored in the calculations. Furthermore, the virial ratios indicate a similar virial status when move to the massive end ( $>300 M_{\odot}$ ) rather than totally different virial status like in the simulations of Rosen et al. (2019). Therefore, we propose that even if a different virial status probably exists between AS and NA, it is hard to find out whether it could dominate the star-forming process such as fragmentation (see Sect. 9.3) in HMSCs.

## 9. Discussion

The HMSCs are cold and dense clumps and probably at the earliest stage of HMSF. Even in such short-living objects, we observed a systematic difference between different HMSCs, depending on their environment (associated or not with an H II region). This sample of HMSCs – H II regions combinations could be a useful database for future case studies about the impact of H II regions on HMSF, especially for very early stages.



**Fig. 22.** Virial ratios. The red, yellow, and blue triangles are similar to Fig. 20. The black dashed lines and gray regions show the region of virial ratio  $< 2$ . The solid lines represent virial ratio = 1. The numbers report the fractions of clumps with a virial ratio  $\geq 2$  and  $< 2$ . The uncertainties of clump mass and line width are propagated to the calculations of virial ratio.



**Fig. 23.** Number distribution of virial ratios in 2–4 kpc. The red, orange, and blue histograms represent S-type AS, O-type AS, and NA, respectively. The solid lines show the median values of corresponding distributions.

### 9.1. $T_{\text{dust}}$ and $L/M$ ratio

We observed a higher  $T_{\text{dust}}$  for AS than that of NA (about 4 to 6 K). The  $T_{\text{dust}}$  difference could be larger when considering the details of environment such as whether the HMSCs are in shells created by the expansion of H II regions or whether there are any implications of compression of H II region on HMSCs. The decreasing trend of  $T_{\text{dust}}$  from S-type AS to NA could be well explained with a weaker external heating of H II regions. Ultraviolet radiation could photodissociate the molecular gas at the surface of HMSCs, which has been shown by the higher detection rate of the typical photodissociation tracer  $\text{C}_2\text{H}$  for AS HMSCs (Zhang et al., in prep.). Furthermore, AS with  $\text{C}_2\text{H}$  detection have a higher  $T_{\text{dust}}$  than AS without  $\text{C}_2\text{H}$  detection (Zhang et al., in prep.). These facts reveal the tight relation between photodissociation and external heating.

For HMSCs, the inner heating effect of potentially embedded low mass protostars could be negligible compared with the feedback from nearby H II regions. This leads to a more complex explanation for  $L/M$  ratio of HMSCs. The AS HMSCs have a  $L/M$  higher than NA about 3–4  $L_{\odot}/M_{\odot}$ . Furthermore, AS could have a  $L/M$  ratio of  $>10 L_{\odot}/M_{\odot}$  in extreme cases.

We point out that  $L/M$  ratio alone is not a reliable evolutionary probe for early massive clumps without knowledge of the environment of the clump. This is particularly true for the very early clumps (IR quiescent).

Molinari et al. (2016) used  $\text{CH}_3\text{C}_2\text{H}$  to estimate the  $L/M$  intervals for different evolutionary stages of massive clumps (see Sect. 6.2). We carried out an independent check to see whether the  $L/M$  ratios of HMSCs could be significantly changed under the effect of H II regions. We took advantage of  $L/M$  ratio and  $\text{CH}_3\text{C}_2\text{H}$  kinematic temperature data of dense massive clump ( $>1000 M_{\odot}$ ) in Giannetti et al. (2017). Our check shows that the heating effect of H II region makes the  $L/M$  ratios of IR quiescent massive clumps increase from  $\lesssim 0.5 L_{\odot}/M_{\odot}$  to  $\gtrsim 2 L_{\odot}/M_{\odot}$ , even to  $\gtrsim 10 L_{\odot}/M_{\odot}$  in extreme cases (see details in Appendix D). The improved  $L/M$  ratios of IR quiescent massive clumps overlap the  $L/M$  regimes of clumps with weak IR sources, such as 70  $\mu\text{m}$  compact sources.  $\text{CH}_3\text{C}_2\text{H}$  is a good thermometer for dense gas ( $\gtrsim 10^5 \text{ cm}^{-3}$ ), thus it traces mostly the inner region of the clump rather than the envelope. Under the effect of the H II region, the increase of the inner  $\text{CH}_3\text{C}_2\text{H}$  temperature is limited compared to the  $L/M$  for IR quiescent clumps. In extreme cases, for example, the clump located at the edge of an H II region, the inner gas temperature of clump could be significantly increased to  $>30$  K owing to the impact of the H II region.

We explore whether it is possible that the  $L/M$  properties of star-forming clumps could be significantly changed by the presence of a nearby H II region. Pitts et al. (2019) studied the  $L/M$  ratios of Carina Nebula and RCW 64 by fitting SED pixel by pixel to *Herschel* and APEX far-IR to submillimeter data. They find that the  $L/M$  ratio of the pixels in the clump show a complex feature. A number of clumps with ongoing star formation activities show some pixels with a higher  $L/M$  or  $T_{\text{dust}}$  at lower  $N_{\text{H}_2}$ , which is inconsistent with the expected positive correlation between  $L/M$  and  $N_{\text{H}_2}$  in star formation theory. We propose that this could be due to the feedback of the Carina and RCW 64 regions that heats the surface and surrounding of the star-forming clump, making some pixels at the surface to have a higher  $L/M$  ratio or  $T_{\text{dust}}$  with lower  $N_{\text{H}_2}$ , for example, clump BYF 77 and BYF 109 in Fig. 21 of Pitts et al. (2019). The  $L/M$  properties of starless clumps are more easily significantly changed due to the external heating and the absence of internal heating sources.

**Table 7.** Summary of the properties of AS compared to NA.

Properties	AS compared to NA
$T_{\text{dust}}$ and $L/M$	Higher ( $\delta \approx 1$ to 2)
Density profile	Steeper ( $\delta \approx 1$ )
Turbulence (Mach number)	Higher ( $\delta \approx 1$ )
Virial ratios	Higher ( $\delta \approx 1$ )

In a word, the suggestions in [Molinari et al. \(2016\)](#) that  $L/M$  ratio could be used as an evolutionary probe of massive clumps at very early stages should be carefully treated, especially when there are impacts from H II regions. A detailed check of the environment of clumps is crucial before using  $L/M$  ratios because the  $L/M$  could be partly created by external heating of nearby H II regions rather than forming stars inside the clump. The single dish observation cannot resolve the compact cores embedded in HMSCs. Observations by ALMA toward a series of HMSCs in different environment are needed to address whether the embedded cores remain higher  $T_{\text{dust}}$  and  $L/M$  when they are associated with nearby H II regions.

### 9.2. Nature of HMSCs in different environments

We have shown that a large part of HMSCs are associated with H II regions (about 60–80%), which means that the H II regions are common environments for HMSCs. A simple look at a series of maps indicates that about half of the AS HMSCs probably show the morphology features of interaction with H II regions such as the compressed layer. A systematic difference has been found between the properties of HMSCs associated and non-associated with H II regions. A short summary of the main differences we found for the properties of AS compared to NA is given in Table 7. The quantity  $T_{\text{dust}}$  is one of the key differences that has been investigated. A higher  $T_{\text{dust}}$  causes a higher  $L/M$  ratio. When  $T_{\text{dust}} > 20$  K, evaporation of CO isotopologs dramatically increases and changes the chemistry in HMSCs, for example, more CO promotes the destruction of  $\text{N}_2\text{H}^+$  (Zhang et al., in prep.). Using chemistry as the evolutionary probe of very early massive clumps, especially using single dish observations with low resolution, creates bias. Recent observations toward Carina nebular reveal similar findings that the environmental changes due to the HMSF have modified the chemical composition of nearby dense clumps ([Contreras et al. 2019](#)). Besides, AS are expected to have different density structures and kinematics compared to NA. The compression effect of ionized gas makes HMSCs located close to H II regions become more centrally peaked. At the same time, the additional energy injection from H II regions makes HMSCs more turbulent.

Even if some properties of HMSCs indicate a Galactocentric gradient with a large scatter, it is still possible that some properties of HMSCs are predominated by the local environment rather than the Galactic scale environments, for example, the  $T_{\text{dust}}$  changes among different types of HMSCs are more significant than the changes among different Galactocentric distances. Some implications from other studies also indicate that the very early HMSF is more likely impacted by the local environment rather than the Galactic-scale environment. [Djordjevic et al. \(2019\)](#) used the ratio of Lyman luminosity to clump mass  $L_{\text{Ly}}/M_{\text{clump}}$  and ratio of bolometric luminosity to clump mass  $L_{\text{bol}}/M_{\text{clump}}$  as the proxies for massive star formation efficiency (massive SFE) and the overall SFE for UCH II regions. They find a much significantly decreasing trend among the Galactocentric distance for  $L_{\text{bol}}/M_{\text{clump}}$  than that of  $L_{\text{Ly}}/M_{\text{clump}}$ , suggesting that

HMSF is relatively independent compared to low or overall star formation.

The S-type AS HMSCs are excellent samples for studying the triggered HMSF at very early stages because they are in a compressed structure or being strongly photoionized/photo-dissociated. We want to understand how the H II region could impact the following HMSF in HMSCs according to the differences we have found.

### 9.3. Massive star formation

Fragmentation is a key process that could determine the initial mass of cores (or fragments) embedded in HMSCs ([Lin et al. 2019](#); [Svoboda et al. 2019](#); [Sanhueza et al. 2019](#); [Louvet et al. 2019](#); [Li et al. 2019](#)). According to the fragmentation theory of molecular clump in the simplest case ([Jeans 1902](#)), which is a self-gravitating medium without turbulent motions that is nonmagnetic, isothermal, infinite, and homogeneous, the clump fragments into smaller cores with mass on the order of the Jeans mass,

$$M_{\text{J}} \propto c_{\text{s}}^3 n_{\text{H}_2}^{-1/2} \propto T^{3/2} n_{\text{H}_2}^{-1/2}, \quad (2)$$

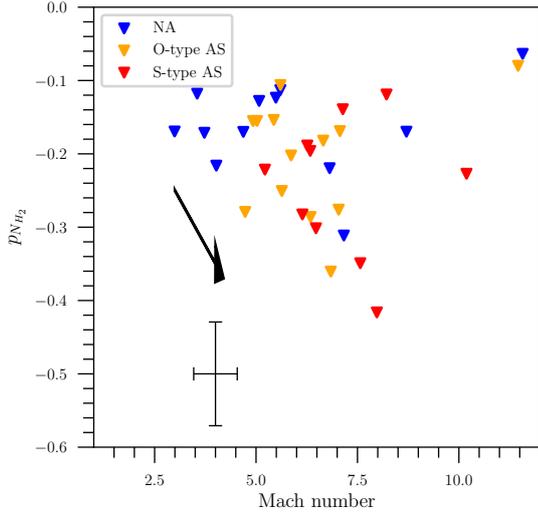
thus a higher temperature or/and lower density generally lead to a massive fragment. The results of Mach number show that HMSCs are predominated by turbulence from supersonic to hypersonic. With the assumptions that turbulence could be treated as additional pressure and that the thermal component  $\sigma_{\text{Therm}}$  only occupies a limited proportion compared to the non-thermal component  $\sigma_{\text{NT}}$  ([Mac Low & Klessen 2004](#)), Eq. (2) can be transformed to

$$M_{\text{J}} \propto \sigma_{\text{NT}}^2. \quad (3)$$

The Jeans mass is higher for a more turbulent clump. Thus, the fragmented cores in AS HMSCs are probably more massive than those of NA HMSCs owing to more turbulent properties. The limited fragmentation at 0.01 pc scale compared to the thermal fragmentation (Eq. (2)) has been found in the most massive clump surrounding the H II region RCW 120 ([Figueira et al. 2018](#)). The fragmented cores are significantly more massive than the expected values with Eq. (2). It is well explained with the additional turbulence from H II region by [Figueira et al. \(2018\)](#).

Our previous results have shown that AS HMSCs have a steeper density profile owing to the compression of H II region. The question is whether a steeper density profile promotes the formation of a massive central object ([McGuire et al. 2016](#)). The simulations of [Girichidis et al. \(2011\)](#) reveal that the cores/clumps with an initial density profile that is strongly centrally peaked are more likely to form one or a few massive stars in the center surrounded by a number of low mass stars, while cores/clumps with a flatter density profile are more likely to form a large number of low mass stars. Furthermore, these simulations suggest that the initial density profile of star-forming clumps/cores could be the most important factor for fragmentation and evolution of HMSF regions. [Palau et al. \(2014\)](#) observed a number of massive dense cores to study the relation between fragmentation and density structure of massive cores and they find that steeper density profiles tend to show lower fragmentation. These results probably indicate that AS more easily form massive stars by limiting fragmentation.

The properties of more turbulent and steeper profiles of AS HMSCs are both implications of the impacts of H II regions and should interconnect with each other well. Such a positive correlation is shown in Fig. 24. Therefore, the combined effects of more



**Fig. 24.** Power-law index of column density profile  $p_{N_{H_2}}$  and Mach number. Only HMSCs with good density profile fitting are included; see details in Sect. 7.1. The blue, orange, and red triangles are similar to Fig. 20. The arrow indicates the trend.

turbulent and more centrally peaked properties of AS probably increase the mass of the fragmented cores and promote HMSF.

As a pilot test to study the possible impacts of H II regions on fragmentation in HMSCs, we select two HMSCs with similar distance (3 kpc), single-dish derived clump mass ( $\approx 300 M_{\odot}$ ), and  $N_{H_2}$  ( $6 \times 10^{22} \text{ cm}^{-2}$ ), but one is S-type AS at the shell structure and another is NA far away from H II regions (Zhang et al., in prep.). The S-type AS HMSC has a higher  $L/M$  ratio of  $2.12 \pm 0.8 L_{\odot}/M_{\odot}$  and Mach number  $8.7 \pm 0.5$ , while the NA HMSC has a lower  $L/M$  ratio  $0.27 \pm 0.08 L_{\odot}/M_{\odot}$  and Mach number  $5.6 \pm 2.3$ . The ALMA observations (ALMA Cycle 4, Project code: 2016.1.01346.S, PI: Thushara Pillai) with a resolution of  $1.5''$  (0.02 pc) show that even though the numbers of fragmented cores are similar between AS and NA (between two to four), the fragmented cores are much denser in S-type AS compared to NA (Zhang et al., in prep.). This pair of AS and NA probably suggests that the impacts of the H II region could favor the formation of massive fragments in clumps at the earliest evolutionary stages.

Recent interferometric studies of a sample of eight massive clumps (0.5 pc scale) close to the H II regions have revealed a possible link between the impacts of H II regions and asymmetric spatial features that high density regions (fragmented cores) drift to off-center position of clump, indicating that the thick envelope of the clump has been deeply compressed (see Fig. 8 in the paper of Zhang et al. 2019). This probe of impacts of H II regions on massive clumps has a drawback that the projection effect probably minimizes the drifted features. For example, if the impacts of H II regions are along the line of sight, the impacted clump still shows a symmetric feature. The density profile could be more reliable than the ways suggested by Zhang et al. (2019). We try to search the relation between density profiles of these eight massive clumps as well as the mass or number of fragmented cores. No clear trend between the density profile power-law index  $p_{N_{H_2}}$  and the mass or number of fragmented cores at 0.02 pc scale has been found, which could be due to the limited number of the sample and/or the broad difference in the properties of these massive clumps, for example, starless or star forming.

Magnetic field is another key factor in the fragmentation and the collapse of clumps (Dall’Olio et al. 2019). Studies suggest that a stronger magnetic field could inhibit the fragmentation

(Fontani et al. 2016; Beuther et al. 2018). The relative importance of magnetic field compared to turbulence could be estimated by Alfvén Mach number  $M_A = \sigma_{NT}/v_A$ , where  $v_A = B/\sqrt{4\pi\rho}$  is Alfvén speed and  $B$  is the total magnetic field. Observations toward massive starless cores/clumps indicate a  $B$  value on the order of 0.1 to 1 mG (Tan et al. 2013; Tamaoki et al. 2019). Even if we can expect that the clumps close to H II region have a higher  $B$  value, whether it should impact the fragmentation requires dedicated observations.

Following our findings, we propose that the properties of AS HMSCs inherited from the impact of H II regions, probably favor the formation of high-mass stars. High-resolution interferometric observations are really needed to ascertain this.

## 10. Conclusions

We investigate the statistical differences between HMSCs close to H II regions (AS or AS HMSCs) and those far away from H II regions (NA or NA HMSCs), based on the following two selection criteria used to ascertain the association: the separations between HMSCs and H II regions in the plane of the sky are  $< 2R_{\text{eff}}$  of the H II region and the velocity differences are  $< 7 \text{ km s}^{-1}$ . The results show that more than half of the HMSCs are AS HMSCs. Among these AS HMSCs, 30–50% of clumps are probably in a compressed structure or photodissociated structure, indicating the powerful feedback of H II regions on HMSCs. This sample of HMSCs constitutes a useful database for future studies dedicated to the impact of H II regions on the earliest HMSF.

We found the following systematic differences between AS and NA:

1. The AS HMSCs are warmer and more luminous than NA types because of the external heating of H II regions. The increment in  $L/M$  of AS HMSCs, which could be up to  $10 L_{\odot}/M_{\odot}$ , can cause a large uncertainty when  $L/M$  ratio is used as the evolutionary probe of massive clumps at very early stages without knowledge of the environment.
2. A steeper column density profile, which means a more centrally peaked density structure, is found in AS compared to NA. The compression by H II region could be the possible reason for this difference.
3.  $C^{18}O$   $J = 2-1$  lines reveal that AS have a higher Mach number than NA, indicating that AS are more turbulent. We propose that this is due to the additional energy injection from H II regions. The AS HMSCs probably have a greater virial ratio compared to NA HMSCs if we do not consider the surface pressure and magnetic field.

The implications of impacts of H II regions on HMSCs, such as positive correlations between the  $L/M$  ratio, the Mach number, and the power-law index of column density profile, are intercorrelated. We propose that the combined effects of more turbulent and steeper density profiles probably make AS HMSCs favor the formation of high-mass stars by limiting fragmentation. Future interferometric surveys toward a statistically significant number of HMSCs in different environments are needed to ascertain this.

*Acknowledgements.* A.Z. thanks the support of the Institut Universitaire de France. We want to thank the anonymous referee for the constructive comments that helped to improve the quality of the paper. The SEDIGISM data is acquired with the Atacama Pathfinder EXperiment (APEX) under programmes 092.F-9315(A) and 193.C0584(A). APEX is a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory. *Herschel* Hi-GAL data processing, map production and source catalogue generation is the result of a multi-year effort thanks to Contracts I/038/080/0 and I/029/12/0 from ASI (Agenzia Spaziale Italiana).

## References

- Aguirre, J. E., Ginsburg, A. G., Dunham, M. K., et al. 2011, *ApJS*, 192, 4
- Anderson, L. D., Bania, T. M., Balsler, D. S., & Rood, R. T. 2011, *ApJS*, 194, 32
- Anderson, L. D., Bania, T. M., Balsler, D. S., et al. 2014, *ApJS*, 212, 1
- Anderson, L. D., Armentrout, W. P., Johnstone, B. M., et al. 2015, *ApJS*, 221, 26
- Anderson, L. D., Armentrout, W. P., Luisi, M., et al. 2018, *ApJS*, 234, 33
- André, P., Di Francesco, J., Ward-Thompson, D., et al. 2014, *Protostars and Planets VI* (Tucson: University of Arizona Press), 27
- Baldeschi, A., Elia, D., Molinari, S., et al. 2017, *MNRAS*, 466, 3682
- Bania, T. M., Anderson, L. D., Balsler, D. S., & Rood, R. T. 2010, *ApJ*, 718, L106
- Barnes, P. J., Muller, E., Indermuehle, B., et al. 2015, *ApJ*, 812, 6
- Bertoldi, F. 1989, *ApJ*, 346, 735
- Bertoldi, F., & McKee, C. F. 1992, *ApJ*, 395, 140
- Beuther, H., Soler, J. D., Vlemmings, W., et al. 2018, *A&A*, 614, A64
- Bock, D. C.-J., Large, M. I., & Sadler, E. M. 1999, *AJ*, 117, 1578
- Bonnell, I. A., Larson, R. B., & Zinnecker, H. 2007, *Protostars and Planets V* (Tucson: university of Arizona Press), 149
- Bronfman, L., Casassus, S., May, J., & Nyman, L. Å. 2000, *A&A*, 358, 521
- Busquet, G., Estalella, R., Palau, A., et al. 2016, *ApJ*, 819, 139
- Butler, M. J., & Tan, J. C. 2012, *ApJ*, 754, 5
- Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, *PASP*, 121, 76
- Churchwell, E., Povich, M. S., Allen, D., et al. 2006, *ApJ*, 649, 759
- Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, *PASP*, 121, 213
- Clark, P. C., Klessen, R. S., & Bonnell, I. A. 2007, *MNRAS*, 379, 57
- Contreras, Y., Rebollo, D., Breen, S. L., Green, A. J., & Burton, M. G. 2019, *MNRAS*, 483, 1437
- Csengeri, T., Urquhart, J. S., Schuller, F., et al. 2014, *A&A*, 565, A75
- Csengeri, T., Bontemps, S., Wyrowski, F., et al. 2017a, *A&A*, 600, L10
- Csengeri, T., Bontemps, S., Wyrowski, F., et al. 2017b, *A&A*, 601, A60
- Csengeri, T., Bontemps, S., Wyrowski, F., et al. 2018, *A&A*, 617, A89
- Cyganowski, C. J., Brogan, C. L., Hunter, T. R., et al. 2017, *MNRAS*, 468, 3694
- Dall'Olio, D., Vlemmings, W. H. T., Persson, M. V., et al. 2019, *A&A*, 626, A36
- Deharveng, L., Schuller, F., Anderson, L. D., et al. 2010, *A&A*, 523, A6
- Djordjevic, J. O., Thompson, M. A., Urquhart, J. S., & Forbrich, J. 2019, *MNRAS*, 487, 1057
- Duronea, N. U., Cappa, C. E., Bronfman, L., et al. 2017, *A&A*, 606, A8
- Elia, D., Molinari, S., Schisano, E., et al. 2017, *MNRAS*, 471, 100
- Ellsworth-Bowers, T. P., Rosolowsky, E., Glenn, J., et al. 2015, *ApJ*, 799, 29
- Elmegreen, B. G. 2011, *EAS Pub. Ser.*, 51, 45
- Figueira, M., Bronfman, L., Zavagno, A., et al. 2018, *A&A*, 616, L10
- Fleming, B., France, K., Lupu, R. E., & McCandliss, S. R. 2010, *ApJ*, 725, 159
- Fontani, F., Commerçon, B., Giannetti, A., et al. 2016, *A&A*, 593, L14
- Fontani, F., Commerçon, B., Giannetti, A., et al. 2018, *A&A*, 615, A94
- Giannetti, A., Brand, J., Sánchez-Monge, Á., et al. 2013, *A&A*, 556, A16
- Giannetti, A., Leurini, S., Wyrowski, F., et al. 2017, *A&A*, 603, A33
- Ginsburg, A., Henkel, C., Ao, Y., et al. 2016, *A&A*, 586, A50
- Girichidis, P., Federrath, C., Banerjee, R., & Klessen, R. S. 2011, *MNRAS*, 413, 2741
- Helfand, D. J., Becker, R. H., White, R. L., Fallon, A., & Tuttle, S. 2006, *AJ*, 131, 2525
- Heyer, M. H., & Brunt, C. M. 2004, *ApJ*, 615, L45
- Hou, L. G., & Gao, X. Y. 2014, *MNRAS*, 438, 426
- Humphreys, R. M., & Larsen, J. A. 1995, *AJ*, 110, 2183
- Imai, R., Sugitani, K., Miao, J., et al. 2017, *ApJ*, 845, 99
- Ingallinera, A., Trigilio, C., Umama, G., et al. 2014, *MNRAS*, 437, 3626
- Jackson, J. M., Simon, R., Shah, R., et al. 2004, *ASP Conf. Ser.*, 317, 49
- Jeans, J. H. 1902, *Phil. Trans. R. Soc. London, Ser. A*, 199, 1
- Juvela, M., Malinen, J., Montillaud, J., et al. 2018, *A&A*, 614, A83
- Kauffmann, J., & Pillai, T. 2010, *ApJ*, 723, L7
- Kendrew, S., Simpson, R., Bressert, E., et al. 2012, *ApJ*, 755, 71
- Kendrew, S., Beuther, H., Simpson, R., et al. 2016, *ApJ*, 825, 142
- Klessen, R. S. 2001, *ApJ*, 556, 837
- Kruijssen, J. M. D., Longmore, S. N., Elmegreen, B. G., et al. 2014, *MNRAS*, 440, 3370
- Kruijssen, J. M. D., Dale, J. E., & Longmore, S. N. 2015, *MNRAS*, 447, 1059
- Krumholz, M. R., & McKee, C. F. 2008, *Nature*, 451, 1082
- Krumholz, M. R., Klein, R. I., McKee, C. F., Offner, S. S. R., & Cunningham, A. J. 2009, *Science*, 323, 754
- Larson, R. B. 1981, *MNRAS*, 194, 809
- Li, G.-X. 2018, *MNRAS*, 477, 4951
- Li, D., Kauffmann, J., Zhang, Q., & Chen, W. 2013, *ApJ*, 768, L5
- Li, H., Li, J.-Z., Yuan, J.-H., Huang, Y.-F., & Ren, Z.-Y. 2018, *Res. Astron. Astrophys.*, 18, 122
- Li, S., Zhang, Q., Pillai, T., et al. 2019, *ApJ*, 886, 130
- Lin, Y., Csengeri, T., Wyrowski, F., et al. 2019, *A&A*, 631, A72
- Liu, M., Tan, J. C., Cheng, Y., & Kong, S. 2018, *ApJ*, 862, 105
- Lo, N., Wiles, B., Redman, M. P., et al. 2015, *MNRAS*, 453, 3245
- Louvet, F., Neupane, S., Garay, G., et al. 2019, *A&A*, 622, A99
- Lu, X., Zhang, Q., Kauffmann, J., et al. 2019, *ApJ*, 872, 171
- Lumsden, S. L., Hoare, M. G., Urquhart, J. S., et al. 2013, *ApJS*, 208, 11
- Mac Low, M.-M., & Klessen, R. S. 2004, *Rev. Mod. Phys.*, 76, 125
- Makai, Z., Anderson, L. D., Mascoop, J. L., & Johnstone, B. 2017, *ApJ*, 846, 64
- Marsh, K. A., & Whitworth, A. P. 2019, *MNRAS*, 483, 352
- Marsh, K. A., Whitworth, A. P., & Lomax, O. 2015, *MNRAS*, 454, 4282
- Marsh, K. A., Whitworth, A. P., Lomax, O., et al. 2017, *MNRAS*, 471, 2730
- Marshall, B., & Kerton, C. R. 2019, *MNRAS*, 489, 4809
- Mattern, M., Kauffmann, J., Csengeri, T., et al. 2018, *A&A*, 619, A166
- McGuire, C., Fuller, G. A., Peretto, N., et al. 2016, *A&A*, 594, A118
- McKee, C. F., & Tan, J. C. 2003, *ApJ*, 585, 850
- Merello, M., Molinari, S., Rygl, K. L. J., et al. 2019, *MNRAS*, 483, 5355
- Milam, S. N., Savage, C., Brewster, M. A., Ziurys, L. M., & Wyckoff, S. 2005, *ApJ*, 634, 1126
- Misiriotis, A., Xilouris, E. M., Papamastorakis, J., Boumris, P., & Goudis, C. D. 2006, *A&A*, 459, 113
- Molinari, S., Swinyard, B., Bally, J., et al. 2010, *PASP*, 122, 314
- Molinari, S., Merello, M., Elia, D., et al. 2016, *ApJ*, 826, L8
- Molinari, S., Baldeschi, A., Robitaille, T. P., et al. 2019, *MNRAS*, 486, 4508
- Motogi, K., Hirota, T., Machida, M. N., et al. 2019, *ApJ*, 877, L25
- Motte, F., Bontemps, S., & Louvet, F. 2018, *ARA&A*, 56, 41
- Myers, P. C. 2009, *ApJ*, 700, 1609
- Olmi, L., Elia, D., Schisano, E., & Molinari, S. 2018, *MNRAS*, 480, 1831
- Orkisz, J. H., Pety, J., Gerin, M., et al. 2017, *A&A*, 599, A99
- Padoan, P., Pan, L., Juvela, M., Haugbølle, T., & Nordlund, Å. 2019, *ApJ*, submitted [arXiv:1911.04465]
- Palau, A., Estalella, R., Girart, J. M., et al. 2014, *ApJ*, 785, 42
- Palmeirim, P., Zavagno, A., Elia, D., et al. 2017, *A&A*, 605, A35
- Paradis, D., Paladini, R., Noriega-Crespo, A., et al. 2012, *A&A*, 537, A113
- Paron, S., Areal, M. B., & Ortega, M. E. 2018, *A&A*, 617, A14
- Pillai, T., Kauffmann, J., Zhang, Q., et al. 2019, *A&A*, 622, A54
- Pitts, R. L., Barnes, P. J., & Varosi, F. 2019, *MNRAS*, 484, 305
- Rathborne, J. M., Jackson, J. M., & Simon, R. 2006, *ApJ*, 641, 389
- Reid, M. J., Dame, T. M., Menten, K. M., & Brunthaler, A. 2016, *ApJ*, 823, 77
- Rigby, A. J., Moore, T. J. T., Plume, R., et al. 2016, *MNRAS*, 456, 2885
- Rosen, A. L., Li, P. S., Zhang, Q., & Burkhardt, B. 2019, *ApJ*, 887, 108
- Samal, M. R., Zavagno, A., Deharveng, L., et al. 2014, *A&A*, 566, A122
- Sanhueza, P., Contreras, Y., Wu, B., et al. 2019, *ApJ*, 886, 102
- Sanna, A., Kölligan, A., Moscadelli, L., et al. 2019, *A&A*, 623, A77
- Schilke, P. 2015, *EAS Pub. Ser.*, 75, 227
- Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, *A&A*, 504, 415
- Schuller, F., Csengeri, T., Urquhart, J. S., et al. 2017, *A&A*, 601, A124
- Simpson, R. J., Povich, M. S., Kendrew, S., et al. 2012, *MNRAS*, 424, 2442
- Stil, J. M., Taylor, A. R., Dickey, J. M., et al. 2006, *AJ*, 132, 1158
- Svoboda, B. E., Shirley, Y. L., Battersby, C., et al. 2016, *ApJ*, 822, 59
- Svoboda, B. E., Shirley, Y. L., Traficante, A., et al. 2019, *ApJ*, 886, 36
- Tamaoki, S., Sugitani, K., Nguyen-Luong, Q., et al. 2019, *ApJ*, 875, L16
- Tan, J. C., Kong, S., Butler, M. J., Caselli, P., & Fontani, F. 2013, *ApJ*, 779, 96
- Tang, M., Liu, T., Qin, S.-L., et al. 2018, *ApJ*, 856, 141
- Taylor, J. H., & Cordes, J. M. 1993, *ApJ*, 411, 674
- Thompson, M. A., Urquhart, J. S., Moore, T. J. T., & Morgan, L. K. 2012, *MNRAS*, 421, 408
- Traficante, A., Fuller, G. A., Billot, N., et al. 2017, *MNRAS*, 470, 3882
- Traficante, A., Fuller, G. A., Duarte-Cabral, A., et al. 2020, *MNRAS*, 491, 4310
- Tremblin, P., Anderson, L. D., Didelon, P., et al. 2014a, *A&A*, 568, A4
- Tremblin, P., Schneider, N., Minier, V., et al. 2014b, *A&A*, 564, A106
- Tremblin, P., Audit, E., Minier, V., Schmidt, W., & Schneider, N. 2015, *Highlights of Astronomy* (Cambridge: Cambridge University Press), 16, 590
- Urquhart, J. S., König, C., Giannetti, A., et al. 2018, *MNRAS*, 473, 1059
- Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, *MNRAS*, 490, 3061
- Veltchev, T. V., Donkov, S., & Klessen, R. S. 2013, *MNRAS*, 432, 3495
- Wenger, T. V., Dickey, J. M., Jordan, C. H., et al. 2019, *ApJS*, 240, 24
- Wienen, M., Wyrowski, F., Menten, K. M., et al. 2015, *A&A*, 579, A91
- Wienen, M., Wyrowski, F., Menten, K. M., et al. 2018, *A&A*, 609, A125
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868
- Yan, Q.-z., Xu, Y., Zhang, B., et al. 2016, *AJ*, 152, 117
- Yuan, J., Wu, Y., Ellingsen, S. P., et al. 2017, *ApJS*, 231, 11
- Yuan, J., Li, J.-Z., Wu, Y., et al. 2018, *ApJ*, 852, 12
- Zavagno, A., Russeil, D., Motte, F., et al. 2010, *A&A*, 518, L81
- Zhang, S.-J., Wu, Y., Li, J. Z., et al. 2016, *MNRAS*, 458, 4222
- Zhang, C.-P., Csengeri, T., Wyrowski, F., et al. 2019, *A&A*, 627, A85

## Appendix A: Details about $v_{\text{lsr}}$ determination and KDA solutions

In this section, we describe the details of  $v_{\text{lsr}}$  determination and KDA solutions mentioned in Sect. 3.1. In the Hi-GAL distance project, a series of inner Galactic surveys are used. These include SEDIGISM, JCMT  $^{13}\text{CO}/\text{C}^{18}\text{O}$   $J = 3-2$  HARP Inner Milky-Way Plane Survey (CHIMPS,  $|b| < 0.5^\circ$  and  $28^\circ < l < 46^\circ$ , with an angular resolution of  $15''$ ; Rigby et al. 2016),  $^{13}\text{CO}$   $J = 1-0$  FCRAO Galactic Ring Survey (FCRAO-GRS,  $18^\circ < l < 55^\circ$  with a resolution of  $46''$ ; Jackson et al. 2004), The Three-mm Ultimate Mopra Milky Way Survey (ThrUMMS  $^{12}\text{CO}/^{13}\text{CO}/\text{C}^{18}\text{O}$   $J = 1-0$  with a resolution of  $72''$ ; Barnes et al. 2015). Hi-GAL distance project identified the  $v_{\text{lsr}}$  components of Hi-GAL compact sources with the observed lines. Generally, we take the  $v_{\text{lsr}}$  components from the spectral lines tracing denser gas. For example, we take the  $\text{C}^{18}\text{O}$   $J = 2-1$  rather than  $^{13}\text{CO}$   $J = 2-1$  in SEDIGISM. For HMSCs that are not included in the Hi-GAL distance project (82 in 463), we keep the  $v_{\text{lsr}}$  identified by Y17 or update it with the results of Urquhart et al. (2018).

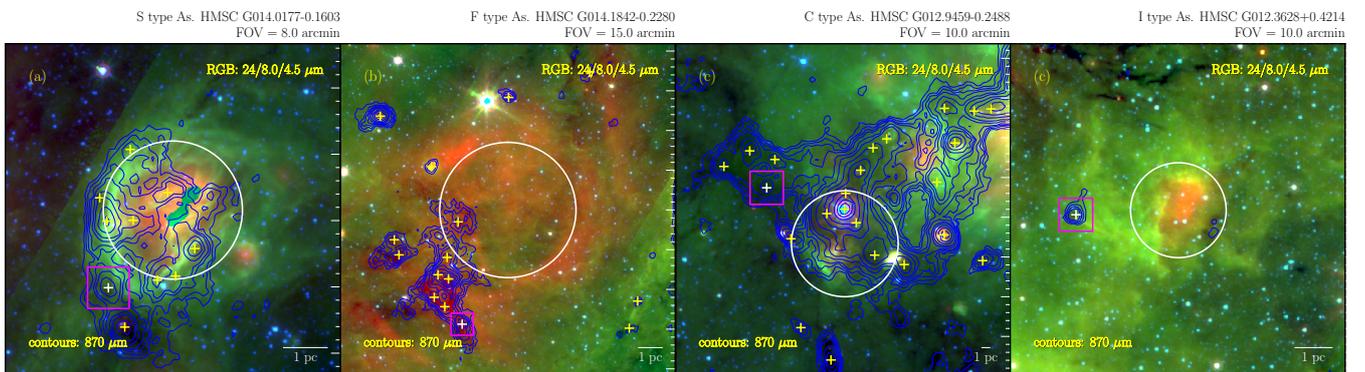
For KDA solutions, we follow the work flows in Urquhart et al. (2018), which test several criteria one by one to ascertain distance. For our cases, we firstly check whether the associated z-distance to the Galactic midplane for far distance is larger than 120 pc. If this is the case, the near distance is chosen to be the distance of HMSCs. Secondly, we check if there is any H I SA feature. A clear H I SA feature indicates that the clump is at near distance. Finally, if there are no H I data or the H I SA feature is inconclusive, we check our GLIMPSE 24/8/4.5  $\mu\text{m}$  RGB images and find out whether these images have a clear extinction feature. The clump has a higher probability to be in near distance if it shows extinction. The extinction criterion is carefully used when HMSC is associated with an H II region. The photodissociation effect of H II region could cover the extinction features of the clumps. Thus we also check the size of associated H II regions; if the size of the H II region is too huge (several dozens of pc) when put the HMSC at far distance, we put HMSC in near distance.

## Appendix B: Morphology of surroundings of HMSCs

We introduce the subtypes for the HMSCs without significant interaction with H II regions in our series of maps. They belong to O-type AS, O-type PA, or NA HMSCs, which are partly described in Sect. 4.4. Basically, there are three subtypes for these HMSCs, which are F (filament) type, C (clumpy) type, and I (isolated) type HMSCs. Observational studies have shown that the massive protostars or protoclusters located at the center (“spoke”) of converging filaments (“hub”) could accrete mass along the filaments (Myers 2009; André et al. 2014; Yuan et al. 2018). We classify the HMSC as F-type if the HMSC is in an elongated molecular cloud structure (several pc) and the morphology of the structure is not in accordance with the ionized front of H II region. When there are several ATLASGAL clumps close to the HMSC and their combined morphology does not indicate an elongated or filamentary structure, we classify it as C-type HMSC. Finally, if the HMSC is relatively isolated from the surrounding molecular material and there is no obvious connection with other ATLASGAL clumps, we classify it as I-type. An example of the classification for morphology of the surrounding of AS HMSCs is shown in Fig. B.1. A short summary is shown in Table B.1. The most common morphologies for the surroundings of AS, PA, and NA are S-type (54%), S-type (36%), and F-type (57%), respectively. I-type clumps are less common for AS ( $\approx 10\%$ ) compared to NA ( $\approx 30\%$ ), suggesting that HMSCs associated with H II regions are usually hosted in a larger molecular cloud complex. Further analyses of this classification and its meaning are ongoing (Zhang et al., in prep.).

**Table B.1.** Environmental morphology statistics.

Types	S-type	F-type	C-type	I-type
AS HMSCs (193)	104 (54%)	50 (26%)	23 (12%)	16 (8%)
PA HMSCs (64)	23 (36%)	18 (28%)	10 (16%)	13 (20%)
NA HMSCs (81)	–	46 (57%)	7 (9%)	28 (34%)



**Fig. B.1.** Example of a more detailed classification of the surroundings of HMSC. Four types of morphology are shown for AS. The RGB images are constructed in the same methods as in Fig. 1. The solid circles indicate the H II regions. Yellow crosses and blue contours represent the ATLASGAL clumps identified by Csengeri et al. (2014) and the ATLASGAL 870  $\mu\text{m}$  emission, respectively. The white crosses with purple squares indicate the HMSCs.

## Appendix C: Detailed descriptions of other properties

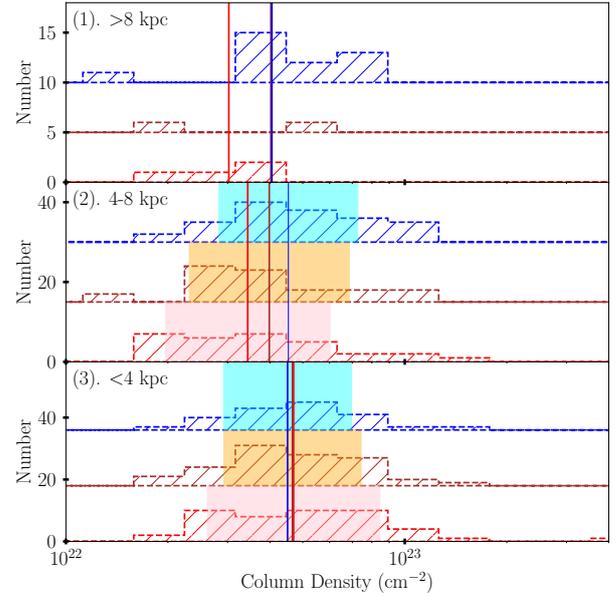
In this section, we describe in more detail the similar properties of the AS and NA HMSCs; these properties consist of  $N_{\text{H}_2}$ ,  $n_{\text{H}_2}$ ,  $M_{\text{clump}}$ , size, and eccentricity.

**Column density  $N_{\text{H}_2}$ .**  $N_{\text{H}_2}$  likely have a flat relation with distance in Figs. 10 and C.1. The K-S tests indicate that S-type AS have a distribution similar to NA and O-type AS. The difference between S-type AS and NA at nearby distance are smaller than  $0.1\sigma$  level (see Fig. 15). Normally, the uncertainties of  $N_{\text{H}_2}$  is around 15%, thus the uncertainties are around  $1 \times 10^{22} \text{ cm}^{-2}$  if we set  $N_{\text{H}_2}$  as an acceptable value of  $5 \times 10^{22} \text{ cm}^{-2}$ . The uncertainties are larger than or around the  $N_{\text{H}_2}$  difference between different types.

The corresponding surface density of the median  $N_{\text{H}_2}$  derived by Y17 is a magnitude lower than the minimum surface density of  $1 \text{ g cm}^{-2}$  for massive star formation derived by Krumholz & McKee (2008). The reason is the beam dilution effect (Yuan et al. 2017). Hi-GAL and ATLASGAL combined images with a resolution of about  $36''$  cannot resolve the inner cores.

**Density  $n_{\text{H}_2}$ .** There is a decreasing trend of  $n_{\text{H}_2}$  with distance. The distance distribution of NA is more concentrated at 4–8 kpc (intermediate distance) compared to that of AS, thus a larger number of NA tend to be in the range of intermediate  $n_{\text{H}_2}$  to create a distribution of  $n_{\text{H}_2}$  that is different from that of AS. Therefore, the K-S test to full sample is unreliable due to the distance bias. K-S tests to AS and NA in each distance bin are not performed because of the small size of sample. The  $\delta$  values indicate that there is a weak trend that HMSCs become denser (at  $0.2\text{--}0.4\sigma$  level) from NA to S-type AS at near distance range, but this trend is inverse at larger distance as shown in Figs. 15 and C.2. The uncertainty of  $n_{\text{H}_2}$  is larger than or similar to the difference between different HMSCs.

**Clump mass  $M_{\text{clump}}$ .** An increasing trend of  $M_{\text{clump}}$  with distance is clearly shown in Figs. 10 and C.3. The reason is that our sample of HMSCs has a physical size increasing with distance while  $N_{\text{H}_2}$  is similar. Mass distributions of clumps (clump mass function; CMF) are largely discussed by a number of research (Klessen 2001; Clark et al. 2007; Veltchev et al. 2013; Wiene et al. 2015), and their results show that the high-mass end of mass distribution could be described by a power law  $dN/dM_{\text{clump}} \propto M_{\text{clump}}^\alpha$ , where  $dN$  is the number of clumps within mass bin  $dM_{\text{clump}}$ . We could not directly perform such a fitting to our mass distribution owing to the bias that mass increases with distance and because the number of sources in different distance bins is limited. Rathborne et al. (2006) obtained a power-law index of  $\alpha \simeq -2.1 \pm 0.4$  for the cores ( $>100 M_\odot$ ) without  $8 \mu\text{m}$  emission in IRDC at 2 kpc to 8 kpc. A tentative fitting to all HMSCs and AS HMSCs ( $M_{\text{clump}} > 500 M_\odot$  and the mass completeness at 4 kpc is about  $450 M_\odot$ ) results in  $\alpha \simeq -2.38 \pm 0.49$  and  $-2.23 \pm 0.32$  (see Fig. C.6), which is similar to the results of Rathborne et al. (2006). This tentative fitting sheds light on the facts that the commonly existing distance bias could produce an unreal CMF. Including more distant clumps, which are generally more massive, could make mass distribution



**Fig. C.1.** Number distributions of  $N_{\text{H}_2}$  of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for  $N_{\text{H}_2}$ .

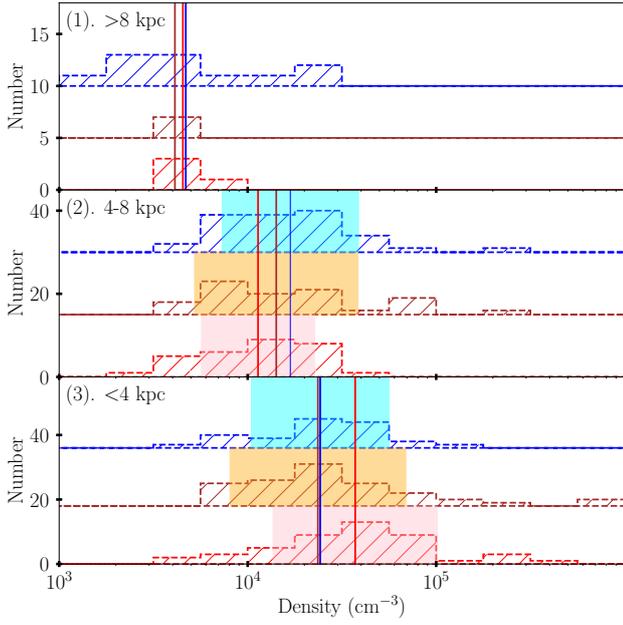
flatter at high-mass end. We propose that distance bias should be seriously considered when studying CMF.

The question arises whether environmental factors impact the mass function of massive clumps (Liu et al. 2018). Wiene et al. (2015) suggested that power-law index ( $-2.2 \pm 0.1$ ) is independent from the Galactic scale environment and evolutionary stages by studying the massive end ( $>1000 M_\odot$ ) of ATLASGAL clumps at different Galactocentric distances for various kinds of sources. The question regarding whether different types of HMSCs could have a similar or different CMF need to be confirmed in a larger sample with a similar distance (Olmi et al. 2018).

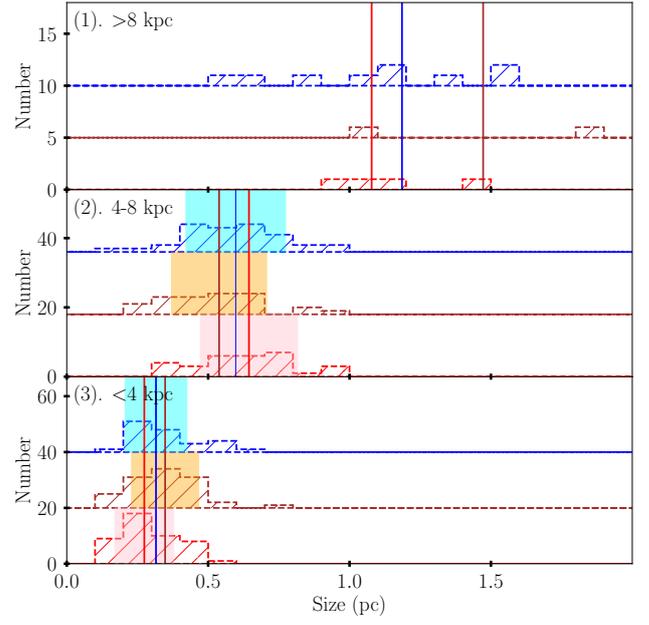
The  $\delta$  values show that NA are slight massive than S-type AS at  $0.3\text{--}0.5\sigma$  level in the same distance bins according to Fig. 15. The uncertainty of clump mass, which is about 20–40%, mainly comes from the uncertainty in distance and column density. The mass difference between different HMSCs is less than its uncertainty.

**Size.** The sizes of our HMSCs are in core-to-clump regime (0.1–1 pc) and increase with the distance as shown in Figs. 10 and C.4. The impact of H II regions on the size of HMSCs is complex. There is no consistent differences between AS and NA in different distance bins. S-type AS are smaller than NA at  $1\sigma$  level in  $<4$  kpc bin while they are larger at 4–8 kpc. Furthermore, there is no consistent changes from S-type AS to O-type AS and then NA.

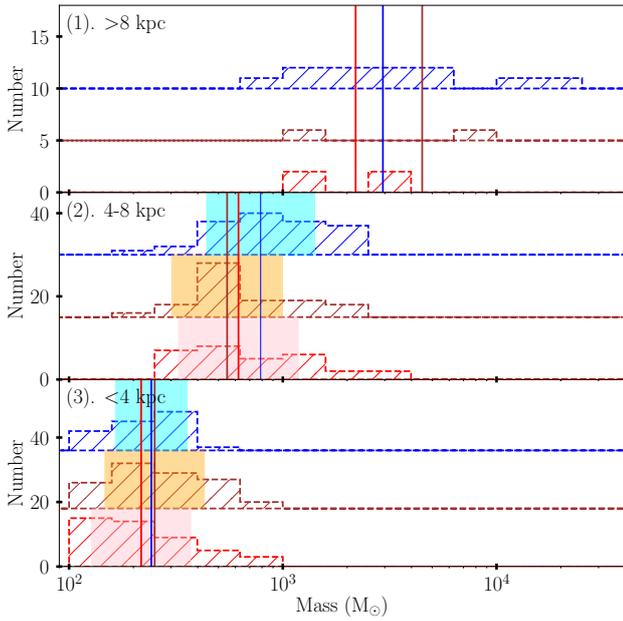
**Eccentricity.** The eccentricity of HMSCs is calculated with  $e_{\text{clump}} = \sqrt{a^2 - b^2}/a$ , where  $a$  and  $b$  are the semimajor axis and semiminor axis of the ellipse defining the shape of clump (Zhang et al. 2016). K-S tests show eccentricity distributions are similar between S-type AS, O-type AS, and NA. Furthermore,  $\delta$  values in different distance bins show an inconsistent trend.



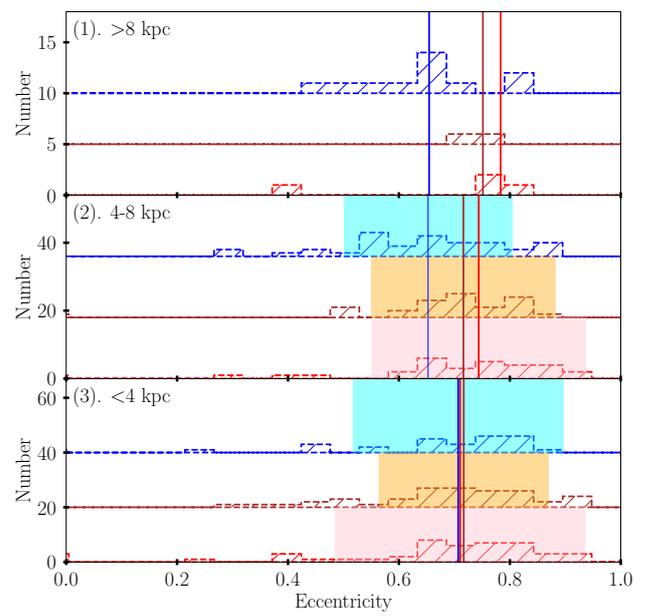
**Fig. C.2.** Number distributions of  $n_{\text{H}_2}$  of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for  $n_{\text{H}_2}$ .



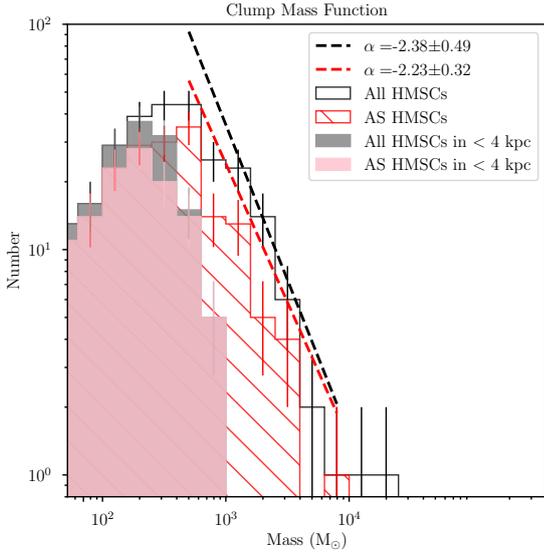
**Fig. C.4.** Number distributions of size of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for size.



**Fig. C.3.** Number distributions of  $M_{\text{clump}}$  of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for  $M_{\text{clump}}$ .



**Fig. C.5.** Number distributions of eccentricity of HMSCs. The meanings of histograms, lines, and color-filled regions are similar to Fig. 12 but for eccentricity.



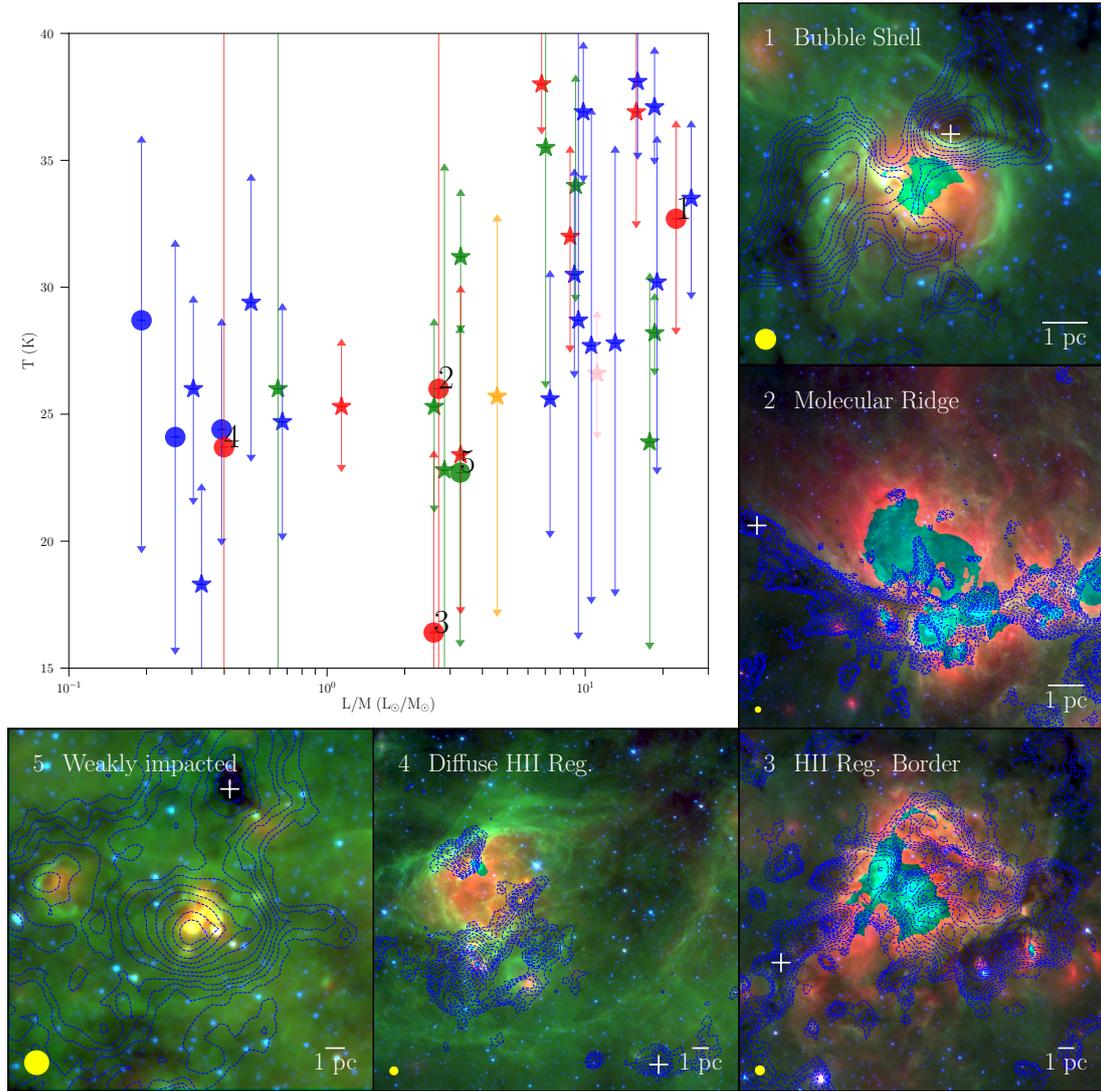
**Fig. C.6.** Clump mass function. The red hatched and black histograms show the number distributions of all selected AS HMSCs and all selected HMSCs, respectively. The black and red dashed lines indicate the fittings with power-law index of  $-2.38$  and  $-2.23$ , respectively. The gray and pink histograms show the number distributions of all HMSCs and AS HMSCs in  $<4$  kpc, respectively.

#### Appendix D: Independent check on $\text{CH}_3\text{C}_2\text{H}$ kinematic temperature and $L/M$ ratio relations

About 100 ATLASGAL massive clumps in the inner Galaxy are included in the survey of [Giannetti et al. \(2017\)](#). The selected

massive clumps are separated into four categories according to star formation indicators observed in IR and radio images: (1)  $70\ \mu\text{m}$  weak clumps: these clumps are weak at  $70\ \mu\text{m}$  thus they are considered as IR quiescent. They represent the clumps at the earliest evolutionary stages similar to our HMSCs. (2) IR weak clumps. These clumps are associated with  $70\ \mu\text{m}$  compact sources or weak  $24\ \mu\text{m}$  sources. (3) IR bright clumps. These clumps are associated with strong IR sources. (4) Compact H II regions. We use  $70\ \mu\text{m}$  weak massive clumps and IR weak massive clumps to carry out the check. We cross match the clumps with H II region in a method similar to that in Sect. 4. The clumps are classified as associated, possibly associated, and nonassociated clumps. The results are shown in Fig. D.1. Five IR quiescent massive clumps associated with H II regions are shown in Fig. D.1. Source 1 is located at the edge of a bubble; ATLASGAL  $870\ \mu\text{m}$  continuum emission contours indicate that the clump is being compressed. Source 2 is located in the famous molecular ridge of the NGC 6334 H II region;  $870\ \mu\text{m}$  continuum shows that the ridge is being compressed. Source 3 is located at the edge of the H II region with complex morphology and intense star formation. Source 4 is located at the edge of a very diffuse bubble. The surrounding diffuse  $24\ \mu\text{m}$  emission indicates that this clump is weakly impacted by the H II region. Its  $L/M$  ratio is quite similar to the source that is nonassociated with H II regions. Source 5 is possibly associated with H II region owing to the lack of H II region velocity.

It is significant that the  $L/M$  ratio and  $\text{CH}_3\text{C}_2\text{H}$  kinematic temperature of Source 1 are shifted to the level of IR weak clumps. Furthermore,  $L/M$  ratios of other sources associated with H II regions are also shifted to the  $L/M$  regime of IR weak clumps.



**Fig. D.1.**  $\text{CH}_3\text{C}_2\text{H}$  temperature,  $L/M$  ratio, and environments. *Top left plot:*  $L/M$  and  $\text{CH}_3\text{C}_2\text{H}$  gas temperature of massive clumps. The dots and stars represent the  $70\ \mu\text{m}$  weak clumps and IR weak clumps, respectively. The red, green, pink, orange, and blue labels represent S-type associated, O-type associated, S-type possibly associated, O-type possibly associated, and nonassociated clumps, respectively.  $L/M$  ratio,  $\text{CH}_3\text{C}_2\text{H}$  temperature and its upper and lower limits are taken from [Giannetti et al. \(2017\)](#). *Other plots:* environments of all  $70\ \mu\text{m}$  weak clumps associated (or possibly associated) with H II region by *Spitzer* RGB images (constructed in the same methods as Fig. 1). The contours and yellow dots in the bottom left of the figures show the ATLASGAL  $870\ \mu\text{m}$  continuum and its resolution.