



CMZoom. II. Catalog of Compact Submillimeter Dust Continuum Sources in the Milky Way's Central Molecular Zone

H Perry Hatchfield¹ , Cara Battersby^{1,2} , Eric Keto², Daniel Walker^{1,3,4}, Ashley Barnes⁵ , Daniel Callanan^{2,6}, Adam Ginsburg⁷ , Jonathan D. Henshaw⁸, Jens Kauffmann⁹, J. M. Diederik Kruijssen¹⁰ , Steve N. Longmore⁶ , Xing Lu⁴ , Elisabeth A. C. Mills¹¹ , Thushara Pillai¹² , Qizhou Zhang¹³ , John Bally¹⁴ , Natalie Butterfield¹⁵ , Yanett A. Contreras¹⁶ , Luis C. Ho^{17,18} , Jürgen Ott¹⁹ , Nimesh Patel² , and Volker Tolls²

¹ University of Connecticut, Department of Physics, 196A Auditorium Road, Unit 3046, Storrs, CT 06269, USA; h.hatchfield@uconn.edu

² Center for Astrophysics, Harvard & Smithsonian, MS-78, 60 Garden Street, Cambridge, MA 02138, USA

³ Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile

⁴ National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo, 181-8588, Japan

⁵ Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121, Bonn, DE, Germany

⁶ Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

⁷ University of Florida Department of Astronomy, Bryant Space Science Center, Gainesville, FL, 32611, USA

⁸ Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

⁹ Haystack Observatory, Massachusetts Institute of Technology, 99 Millstone Road, Westford, MA 01886, USA

¹⁰ Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany

¹¹ Department of Physics and Astronomy, University of Kansas, 1251 Wescoe Hall Drive, Lawrence, KS 66045, USA

¹² Boston University Astronomy Department, 725 Commonwealth Avenue, Boston, MA 02215, USA

¹³ Center for Astrophysics, Harvard & Smithsonian, MS-42, 60 Garden Street, Cambridge, MA 02138, USA

¹⁴ CASA, University of Colorado, 389-UCB, Boulder, CO 80309, USA

¹⁵ Green Bank Observatory, 155 Observatory Road, P.O. Box 2, Green Bank, WV 24944, USA

¹⁶ Leiden Observatory, Leiden University, P.O. Box 9513, NL 2300 RA Leiden, The Netherlands

¹⁷ Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China

¹⁸ Department of Astronomy, School of Physics, Peking University, Beijing 100871, People's Republic of China

¹⁹ National Radio Astronomy Observatory, 1003 Lopezville Road, Socorro, NM 87801, USA

Received 2020 July 17; revised 2020 August 31; accepted 2020 September 2; published 2020 November 9

Abstract

In this paper we present the CMZoom survey's catalog of compact sources ($<10''$, ~ 0.4 pc) within the central molecular zone (CMZ). CMZoom is a Submillimeter Array large program designed to provide a complete and unbiased map of all high column density gas ($N(\text{H}_2) \geq 10^{23} \text{ cm}^{-2}$) of the innermost 500 pc of the Galaxy in the 1.3 mm dust continuum. We generate both a robust catalog designed to reduce spurious source detections, and a second catalog with higher completeness, both generated using a pruned dendrogram. In the robust catalog, we report 285 compact sources, or 816 in the high-completeness catalog. These sources have effective radii between 0.04 and 0.4 pc, and are the potential progenitors of star clusters. The masses for both catalogs are dominated by the Sagittarius B2 cloud complex, where masses are likely unreliable due to free-free contamination, uncertain dust temperatures, and line-of-sight confusion. Given the survey selection and completeness, we predict that our robust catalog accounts for more than $\sim 99\%$ of compact substructure capable of forming high-mass stars in the CMZ. This catalog provides a crucial foundation for future studies of high-mass star formation in the Galactic Center.

Unified Astronomy Thesaurus concepts: [Star formation \(1569\)](#); [Galactic center \(565\)](#); [Submillimeter astronomy \(1647\)](#); [Star forming regions \(1565\)](#); [Dense interstellar clouds \(371\)](#); [Infrared dark clouds \(787\)](#); [Protostars \(1302\)](#); [Molecular clouds \(1072\)](#); [Interferometry \(808\)](#); [Catalogs \(205\)](#)

1. Introduction

The Milky Way's innermost 500 pc, the central molecular zone (CMZ), houses a vast complex of $\sim (3-5) \times 10^7 M_\odot$ of molecular gas (Morris & Serabyn 1996; Dahmen et al. 1998; Pierce-Price et al. 2000). The molecular clouds in this environment exist at high densities and temperatures relative to the disk, exhibiting intense pressures, magnetic fields, and turbulence (e.g., Mills & Morris 2013; Rathborne et al. 2014; Pillai et al. 2015; Federrath et al. 2016; Ginsburg et al. 2016; Henshaw et al. 2016b; Walker et al. 2018) as well as high cosmic-ray ionization rates (Goto et al. 2013; Harada et al. 2015; Le Petit et al. 2016; Oka et al. 2019) and UV background radiation (Lis et al. 2001). Star formation within this complex environment is an area of extensive previous research and current study (e.g., Downes & Maxwell 1966; Guesten & Downes 1982; Morris & Serabyn 1996; Yusef-Zadeh & Wardle 2008; Longmore et al. 2012, 2013a; Kauffmann et al. 2013b, 2017a, 2017b; Kruijssen et al. 2014; Lu et al. 2015, 2019a, 2019b). The

CMZ is a unique laboratory for studying star formation, as these conditions, rarely observed in the rest of the Milky Way, exhibit some similarities to properties of high redshift galaxies, allowing us an indirect glimpse into the cosmic history of star formation for which comparably detailed extragalactic observations are not currently possible (Kruijssen & Longmore 2013).

The birthplaces of stars are understood to be the over-densities in giant molecular clouds (GMCs), “clumps,” of a characteristic size of $\sim 0.3-3$ pc. These clumps fragment further into gravitationally bound “cores” of a characteristic size of $\sim 0.03-0.2$ pc, which in turn will form individual stellar systems (Bergin & Tafalla 2007; Kennicutt & Evans 2012). The formation of massive stars ($M > 8 M_\odot$) is relatively rare due to the fragmentation of dense clumps leading to a greater number of low-mass stars for which the mass distribution is described by an initial mass function (IMF; Bastian et al. 2010; Offner et al. 2014). Studying the nature of transitions between

these evolutionary stages, leading from turbulent GMCs on scales >3 pc down to individual young stellar systems demands large surveys of both star-forming and non-star-forming structures across the hierarchical continuum of relevant scales (e.g., Kauffmann et al. 2017a, 2017b; Walker et al. 2018; Lu et al. 2019a, 2019b).

There are several indications of recent and active star formation in the Galactic Center. The central few hundred parsecs of the Galaxy house two young massive clusters, the Arches Cluster and the Quintuplet Cluster, suggesting a recent burst of star formation activity in the CMZ within the last ~ 5 Myr (Figer et al. 2002; Habibi et al. 2013; Lu et al. 2013; Hosek et al. 2019). At present, there are several known stellar nurseries actively forming stars in the CMZ, most notably the cloud complexes Sagittarius (Sgr B2), Sgr C, and the Dust Ridge (Gordon et al. 1993; Yusef-Zadeh & Wardle 2008; Yusef-Zadeh et al. 2009; Schmiedeke et al. 2016; Ginsburg et al. 2018; Walker et al. 2018; Barnes et al. 2019; Lu et al. 2019a). The mean gas densities observed in the CMZ are much higher than those observed in the Galactic disk, leading conventional theories of star formation to predict a correspondingly high star formation rate (SFR) per molecular gas mass relative to the disk (Lada et al. 2012; Longmore et al. 2013a; Barnes et al. 2017). Such a high SFR is not observed, with studies revealing SFRs a factor of 10–100 lower than predicted by current theories of star formation relative to the dense gas content (Immer et al. 2012; Longmore et al. 2013a; Kruijssen et al. 2014; Koepferl et al. 2015; Barnes et al. 2017). In particular, Barnes et al. (2017) show how the observed low SFR cannot be accounted for by observational bias alone, demonstrating how independent methods of measuring the SFR agree with one another to within a factor of 2. This implies an environmental cause for the deficiency of star formation. It has been proposed that this low SFR and high dense gas content indicate an upcoming starburst in the Galactic Center, possibly part of a dynamical episodic cycle of star formation and quiescence (Kruijssen et al. 2014; Krumholz & Kruijssen 2015; Krumholz et al. 2017, see also Sormani & Li 2020). Recent simulations suggest that variation in the SFR might be driven by variations in the CMZ’s total gas mass, perhaps modulated by uneven accretion driven by the Galactic bar (Sormani et al. 2018, 2020; Tress et al. 2020). In order to effectively test theoretical explanations for this present lack of star formation, the CMZoom team has carried out a survey of all potential sites of massive star formation at sufficient resolution to resolve and characterize compact substructure.

In this paper, we present the catalog of compact sources constructed from the CMZoom survey’s 1.3 mm continuum data (Battersby et al. 2020, hereafter Paper I). We refer to the entries in this catalog as either compact sources or (dendrogram) leaves as they span spatial scales between both clump-like emission and core-like emission, between ~ 0.05 and 0.36 pc (this decision is explained further in Section 3.1). This catalog of compact sources is designed to ultimately help understand the anomalous SFR and hierarchy of dense gas structures in the CMZ by providing a complete survey of all compact sources embedded in gas above a column density of 10^{23} cm^{-2} .

Several of the regions selected in the CMZoom survey have already been the topic of more detailed single cloud studies. In particular the cloud G0.253+0.016, colloquially known as “the Brick” or “the Lima Bean Cloud,” has been closely studied as a

flagship example of the star formation deficiency in the CMZ, where observations detect an excess of dense gas tracers relative to the expected abundance of star formation signatures (Lis et al. 1994; Longmore et al. 2012; Kauffmann et al. 2013b; Rathborne et al. 2014; Lu et al. 2015; Mills et al. 2015; Henshaw et al. 2019). There is theoretical framework (Kruijssen et al. 2015; Jeffreson et al. 2018; Dale et al. 2019; Kruijssen et al. 2019) and evidence (Immer et al. 2012; Longmore et al. 2013b; Rathborne et al. 2014; Walker et al. 2016, 2018; Barnes et al. 2017, 2019) suggesting that the clouds composing the Dust Ridge, connecting the Brick to Sgr B2, together exhibit a time-sequence evolution of transition from quiescence to active star formation triggered by dynamical effects. While individual clouds display nonmonotonic properties along this stream (Kauffmann et al. 2017a, 2017b), it has been suggested these variations could possibly be explained by using different initial conditions in simulated versions of such a stream (Dale et al. 2019; Kruijssen et al. 2019).

Other notable objects in the region include the massive star-forming complexes around Sgr B2 and Sgr A, and Sgr C, which have been studied extensively since the discovery of their bright emission in the radio continuum (e.g., Downes & Maxwell 1966; Lo & Claussen 1983). Many other regions in the CMZoom survey have not been the subject of any previous study at comparable resolution and sensitivity. For a more complete description of the source selection and pointings, we refer the reader to Paper I.

The structure of this paper is as follows. In Section 2 we present a brief overview of the data, summarizing key points from Paper I. In Section 3 we provide a description of the cataloging algorithm and the simulated observational procedure used for source recovery experiments and the calculation of our catalogs’ completeness. Section 4 describes resultant contents and properties of the “robust” and “high-completeness” versions of the catalog that we have produced. Section 5 highlights the initial analysis of the catalog distributions and the role of Sgr B2 in the catalog, as well as placing limits on the catalog’s high-mass stellar precursor completeness and the star formation potential of the CMZ. In Section 6 we summarize the key points from this work. Appendix A recreates the figures from Section 4 for the high-completeness catalog. Additionally, Appendix B provides a parameter study for the catalog algorithm, showing the effects of varying the input parameters of the pruned dendrogram algorithm on the final catalog contents and statistical distributions. Appendix C is a gallery of zoomed in regions of the CMZoom dust continuum map with the catalog leaves over-plotted as contours. Lastly, Appendix D compares several key regions with corresponding data at higher spatial resolutions from the Atacama Large Millimeter/submillimeter Array (ALMA).

2. Description of the Data

The CMZoom survey is a 550 hr project using the Submillimeter Array²⁰ (SMA), mapping the highest column density gas in the innermost $5^\circ 0$ in longitude and $0^\circ 5$ in latitude of the Galaxy at 1.3 mm, as well as selected additional regions detailed in Paper I. The survey is designed to include all CMZ clouds with column densities above 10^{23} cm^{-2} as

²⁰ The Submillimeter Array is a joint project between the Smithsonian Astrophysical Observatory and the Academia Sinica Institute of Astronomy and Astrophysics, and is funded by the Smithsonian Institution and the Academia Sinica.

derived from the Herschel cold dust continuum (by the spectral energy distribution, SED, fitting procedure outlined in Battersby et al. 2011, explored in Mills & Battersby 2017), with the exception of the cloud to the southeast of Sgr B2 and isolated bright pixels in the Herschel continuum. The CMZoom pointings also contain some lower column density regions of interest across the CMZ, such as far-side cloud candidates, the circumnuclear disk, the Arches filaments, and a bridge of emission that, in projection, appears to connect the dust ridge and 50 km s^{-1} clouds (see Figure 1 in Paper I).

The survey uses the SMA in both its compact and subcompact configurations in order to effectively probe the range of spatial scales between $3''$ and about $45''$ (corresponding to a physical range between 0.12 and 1.8 pc at a Galactic Center distance of 8.178 ± 0.013 (stat.) ± 0.022 (sys.) kpc; The GRAVITY Collaboration et al. 2019). With this setup, we achieve a typical spatial resolution of $3''/2$ (0.13 pc) in the 230 GHz dust continuum wideband (8+ GHz) from which these catalogs are constructed. Single dish data from the Bolocam Galactic Plane survey have been feathered with the SMA data in order to supplement our sensitivity to larger-scale structure. For the catalogs created in this work, we used the SMA-only data, as we are interested only in the most compact emission structures. All data were calibrated in MIR IDL using standard SMA calibration procedures, and imaging and deconvolution were completed in a combination of MIR IDL and CASA. All of the CMZoom pointings, SMA configurations, and imaging pipeline details are described in more detail in Paper I, and the data products have been made available on *Dataverse* at <https://dataverse.harvard.edu/dataverse/cmzoom>.

3. Catalog Design and Methodology

3.1. Pruned Dendrogram Algorithm

The 1.3 mm dust continuum emission observed in the CMZoom survey contains structures widely varying in shape, spatial scale, intensity, and local noise (see Paper I for more details). We developed a cataloging algorithm in order to consistently catalog the compact emission in the CMZoom continuum data despite this diversity. To this end, the catalog was developed using an implementation of the dendrogram algorithm, with additional pruning for local noise.

The dendrogram algorithm is a hierarchical clustering algorithm, representing the significant structures in a data set as the branches and leaves of a tree-like data structure. The highest level structures in the cluster hierarchy are called “leaves,” which represent local maxima of the flux. These leaves are, therefore, the highest column density and most compact sources in the CMZoom maps. The parent “branches” of these leaves represent the less dense regions immediately surrounding the compact sources. The parent branches of those branches represent lower and lower level emission, and so on down to some flux floor.

The dendrogram tree for a given data set is uniquely determined by the choice of three parameters: a minimum structure value, a minimum number of pixels, and a minimum significance value. The minimum structure value is the lowest peak pixel value that will be considered as a structure in the tree. The minimum number of pixels determines the minimum size of a structure included in the tree. Finally, the minimum significance parameter is a measure of how high the pixel value of a new peak must be relative to nearby structures to be considered independent and added to the catalog, typically

chosen relative to the noise (e.g., Rosolowsky et al. 2008; Henshaw et al. 2016a).²¹

The continuum map was translated to units of megajanskys per steradian to allow comparisons between data sets despite variations in the beam size across the survey (Paper I, Section 4.2 provides a complete description of the beam size variation). The initial dendrogram parameters were chosen relative to the lowest rms noise estimates for the entire surveyed region, 3 MJy Sr^{-1} . The minimum pixel value is chosen to be a factor of 3 greater than this global noise estimate. The minimum pixel number was selected to be roughly half of the average beam size over the survey. The minimum significance parameter was chosen to be equivalent to the global low noise estimate, which allows detection of substructure within low flux regions while not significantly altering the distribution of cataloged structure in higher flux regions (this is justified in Appendices A and B).

This initial dendrogram is highly populated due to the fact that the algorithm considers the significance of leaves as compared the low global noise estimate without any regard for local noise, which can be more important in regions with irregular noise properties. We have chosen to narrow the catalog by “pruning” the list of dendrogram leaves relative to their local rms noise. This pruning process amounts to the removal of each dendrogram leaf that does not meet a chosen set of requirements on both the peak leaf flux and the mean leaf flux. In choosing these parameters, we have opted to construct two catalogs: the former prioritizing only the most robust leaves with a peak flux at least 6σ above the local rms estimate and the latter prioritizing a higher catalog completeness, which includes leaves with a peak flux at least 4σ above the local rms estimate. Both versions of the catalog require a mean flux of at least 2σ above the local rms. Such high thresholds are required since the UV coverage is not perfect, and residual non-Gaussian imaging artifacts can strongly effect source selection unless sufficiently strict local noise restrictions are implemented.

A map of the local rms noise estimates across the entire survey was constructed from the residual images generated in the cleaning process described in detail in Paper I. In order to generate an estimate of the rms for each pixel in the survey, we find the spatially averaged standard deviation of these residual maps, calculated as

$$\sigma_{x,y} = \sqrt{(R_{x,y} - (R * G)_{x,y})^2 * G_{x,y}} \quad (1)$$

where $R_{x,y}$ corresponds to the residual value at pixel coordinates (x,y) , and $*G$ denotes convolution with a Gaussian kernel with $\sigma = 14$ pixels, or $\sim 7''$. This effectively recovers an rms value for each pixel by taking the difference between the pixel’s value in a smoothed and unsmoothed version of the residuals. We find a median rms over all surveyed regions is 13 MJy sr^{-1} . For more information about the rms calculation and properties, we refer the reader to Section 4.2 of Paper I. The smoothing kernel size of 14 pixels is justified with a parameter study in Appendix B. The choices of the pruning thresholds for local noise pruning and the resulting completeness for each catalog are explored using the simulated source recovery experiment described in the following section.

²¹ A more detailed description and the documentation for the dendrogram algorithm implementation can be found at <https://dendrograms.readthedocs.io/>.

3.2. Source Recovery and Simulated Observations

There are many potential sources of error in the cataloging process including, but not limited to, the emergence of imaging artifacts in the cleaning process that might erroneously be identified as independent leaves, variations in measurement conditions such as the changes of antenna configuration or weather conditions, as well as the inherent biases of the dendrogram algorithm (Rosolowsky et al. 2008). To probe the effects of these sources of error, we estimate a completeness percentage for the catalog, derived from an algorithm built around the NRAO CASA simulated observation functionality (McMullin et al. 2007).

The scripts to generate this completeness percentage have been made available on Github²² in a generalized form such that it can be used to test the accuracy of any cataloging algorithm, with particular ease if the data were observed using one of the observatories compatible with the current version of CASA, though other interferometers can be included with relative ease. The procedure for determining our completeness using simulated observations is as follows:

Step (1) An artificial image is randomly populated with 10 point sources at a fixed intensity representing the emission of a field of unresolved core-like objects.

Step (2) A simulated observation of the artificial skymap is generated using the `CASA simobserve` function. This function allows us to mimic the true SMA observations by generating visibilities from a sample antenna configuration from the CMZoom observations, and a typical length of observation used for pointings in the survey. These fake data can be combined with noise in order to simulate the effects of thermal noise and atmospheric attenuation using the `AATM` package,²³ an atmospheric modeling package designed for use with ALMA (Pardo et al. 2001). The completeness results for our catalogs constructed from the simulated observations are resilient to variations typical in atmospheric conditions for CMZoom observations.

Step (3) The UV data generated by the simulated observations are then imaged using the `CASA tclean` task. This procedure is identical to the imaging process for a typical low rms noise region in the CMZoom survey. The `tclean` task was called with a multi-scale parameter set to scales [0, 3, 9, and 27]. Briggs weighting with a robust parameter of 0.5 was used throughout. This choice is described in more detail in Paper I. We used the `auto-multithresh`²⁴ feature (with identical input parameters to those detailed in Paper I), and a continuum threshold of 1 mJy/beam. An example of the resulting simulated image is shown in Figure 1.

Step (4) The residual images from `tclean` are used to produce an rms map, just as described in Section 3.1. An initial dendrogram catalog is run on the images resulting from the `tclean` process, using the same dendrogram parameters as in the catalog procedure, scaled by a global noise estimate corresponding to a low value noise estimate in the artificial image. The leaves of this dendrogram are then pruned using a noise map created identically to the catalog procedure in Section 3.1 so that the remaining leaves have a peak flux above at least 6σ and average flux above at least 2σ for the robust

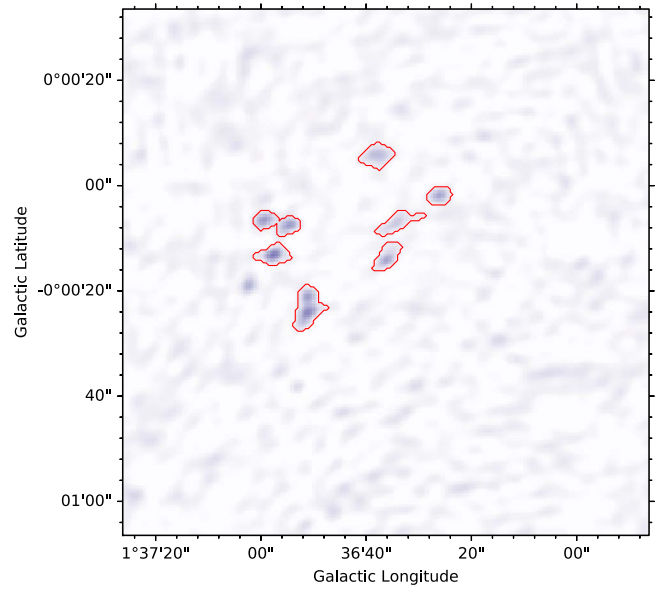


Figure 1. A typical simulated observation used to calculate the completeness of the CMZoom catalog algorithm. For this example, the skymap was populated with 10 point sources. The red contours are the leaf outlines or the simulated catalog, generated by applying the same version of the pruned dendrogram algorithm to the simulated image. In this case, nine of the 10 point sources are recovered, two nearby point sources are grouped together in one leaf, and there are no false-positive source detections.

cataloging, or a peak flux above at least 4σ and average flux above at least 2σ for the high-completeness catalog.

Step (5) Finally, the pruned dendrogram leaves are compared to the locations of the injected point sources. A given point source is considered to be recovered if its coordinates are located anywhere within a pruned leaf.

A plot of this completeness percentage for clumps as a function of leaf mass is depicted in Figure 2. We found that point sources above 80 solar masses, under the typical observational conditions for the CMZoom survey, are detected with more than 95% completeness by the cataloging algorithm. The detection rate rapidly drops off, leaving only $\sim 80\%$ of structures detected at a ~ 70 solar mass threshold, and about 50% at ~ 60 solar masses. To calculate these masses from the fluxes in the images, we assume a constant dust temperature of 20 K, a typical value from Herschel temperature estimates in the CMZ (Mills & Battersby 2017). These mass completeness estimates account only for the observational biases in observing point-source-like emission, and do not take into account the unrelated biases arising from more extended high-level emission structures. The bright, dense objects included in our catalog represent only the highest levels of this complex structure, and can be modeled by point-source-like emission for the purposes of this simplistic completeness estimate. We work on the assumption that the point-source analysis is sufficient for modeling the bright emission that we are prioritizing in our cataloging procedure. We emphasize that some of the cataloged emission is in more extended sources and does not closely match this point-source-like emission model, and we advise caution in interpreting these completeness estimates for more spatially intricate sources. We expect that some catalog leaves may include many unresolved or extremely nearby point sources, particularly in extreme regions like the Sgr B2 complex.

²² https://github.com/CMZoom/core_catalogue

²³ <https://www.mrao.cam.ac.uk/bn204/alma/atmomodel.html>

²⁴ <https://casa.nrao.edu/casadocs/casa-5.3.0/synthesis-imaging/masks-for-deconvolution>

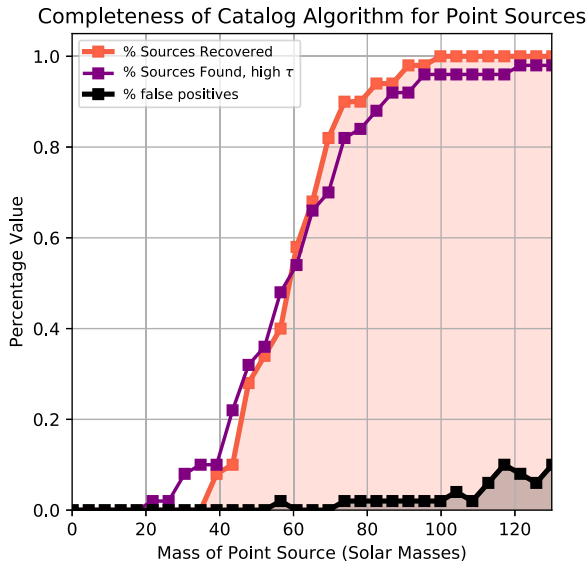


Figure 2. The completeness percentage for source recovery using the procedure outlined in Section 3.2. The red line shows the number of point sources recovered from the artificial sky map for varying point-source masses (assuming a dust temperature $T_d = 20$ K), while the purple line shows a sample recovery for weather conditions with a high zenith opacity ($\tau_0 = 0.2$). Our cataloging algorithm recovers 95% of objects with a mass of $\sim 80 M_\odot$ for an assumed dust temperature of $T_d = 20$ K. This completeness mass is very sensitive to the assumed dust temperature. For example, when we instead assume a dust temperature $T_d = 50$ K, we find a much lower 95% completeness mass of $\sim 25 M_\odot$. Higher dust temperatures are possible on scales smaller than those resolved by Herschel ($36''$) in areas of active star formation such as Sgr B2 and Dust Ridge cloud C.

A similar test was constructed to provide a simplistic measure of the rate of false-positive identifications, by measuring the number of leaves not associated with any of the injected point sources. The results of this completeness test are also displayed in Figure 2. This is an idealized measure of the false-positive value for objects in the catalog, since our technique considers only point-source emission, and is not necessarily applicable to the component of extended emission detected in the CMZoom survey. While the real interstellar medium is indisputably more complex than our simple model, this analysis provides a basic measure of the completeness of the survey, based on the effects of weather, imaging artifacts, and the cataloging scheme.

4. Catalog Results

Here we present the initial results of the catalog, initially focusing on the robust version of the catalog. Corresponding figures for the high complete catalog can be found in Appendix A. A complete gallery of all surveyed regions with cataloged leaves shown as contours is provided in Appendix C. We wish to emphasize two notes before continuing.

1. The nature of the objects cataloged in this work occupy a range of spatial sizes that do not correspond unanimously to either cores (~ 0.03 – 0.2 pc) or clumps (~ 0.3 – 3 pc). Some of sources appear more filamentary and extended, while other sources are more compact and resemble large cores that likely are fragmented below our resolution limit. We briefly compare by eye the nature of the SMA emission to that of higher spatial resolution maps from ALMA in several key regions in Appendix D. The mass–radius relationship of our cataloged sources is compared

to previous studies of the Galactic Center and disk in Section 4.4.

2. We refer to the contents of our catalog as “dendrogram leaves” and “compact sources” interchangeably. Typically we use “leaves” when discussing the nature of the objects in the context of the cataloging algorithm, and “compact sources” when describing the physical nature or properties of these objects. They are effectively interchangeable in either instance.

4.1. Leaf Distribution in l – b Space

The robust catalog contains 285 leaves, all located within the Galactic latitude range of $-0^\circ.25$ to $0^\circ.07$ and the Galactic longitude range $-0^\circ.87$ to $1^\circ.69$. Their distribution in l – b space is shown in Figure 3. There is a high density of leaves associated with the massive star-forming complex Sgr B2, many of which are the most massive objects contained in the catalog. See Appendix D for a direct comparison between the Sgr B2 leaves in this work and the dust continuum emission observed at 3 mm using ALMA by Ginsburg et al. (2018). Outside the Sgr B2 cloud complex, the leaf distribution peaks in the dust ridge and around Sgr A*/circumnuclear disk (CND). The CND is a highly time variable source of synchrotron emission in the submillimeter regime (e.g., Serabyn et al. 1997), and the flux detected in its immediate surroundings appears to be dominated by imaging artifacts. For this reason we have chosen to exclude all leaves in this region from the following analysis, though they are included in the version of the catalog released with this paper.

The distribution of leaves, in Galactic longitude and latitude, is displayed in Figures 3 and 4. Figure 4 shows the mass of sources per unit area surveyed as a function of latitude and longitude as well. While this distribution is highly influenced by the mapping strategy, it still highlights some of the high column density regions that lack compact substructure, such as the $1^\circ.1$ cloud complex. The leaf distribution is largely dominated by the Sgr B2 cloud complex, the Dust Ridge, and the 20 and 50 km s^{-1} clouds. The lack of compact substructure in many of the other surveyed clouds, such as the $1^\circ.1$ cloud complex, is analyzed in detail in Section 5 of Paper I.

4.2. Comparison with Cloud Scale Column Densities

To supplement this catalog, we measure the local column density and temperature of the SMA compact sources and their surroundings from the Herschel property maps created using the dust SED fitting technique detailed in Battersby et al. (2011), to be presented in C. Battersby et al. (2020, in preparation). The values are extracted from these maps at the position of the centroid of the catalog sources. These are cloud scale column density measurements at a resolution of $36''$ (1.5 pc).

Figure 5 shows the histogram of Herschel column density for pixels within the area observed by CMZoom. Overlaid on this is a subset histogram of Herschel column density for pixels associated with our cataloged SMA compact sources. We determine pixels as associated if they lie within one Herschel beam FWHM (i.e., $36''$) of an SMA catalog leaf centroid. The percentage of Herschel pixels that contain an SMA catalog source appears to sharply increase around a Herschel column density of $(1\text{--}2) \times 10^{23} \text{ cm}^{-2}$. The amount of compact

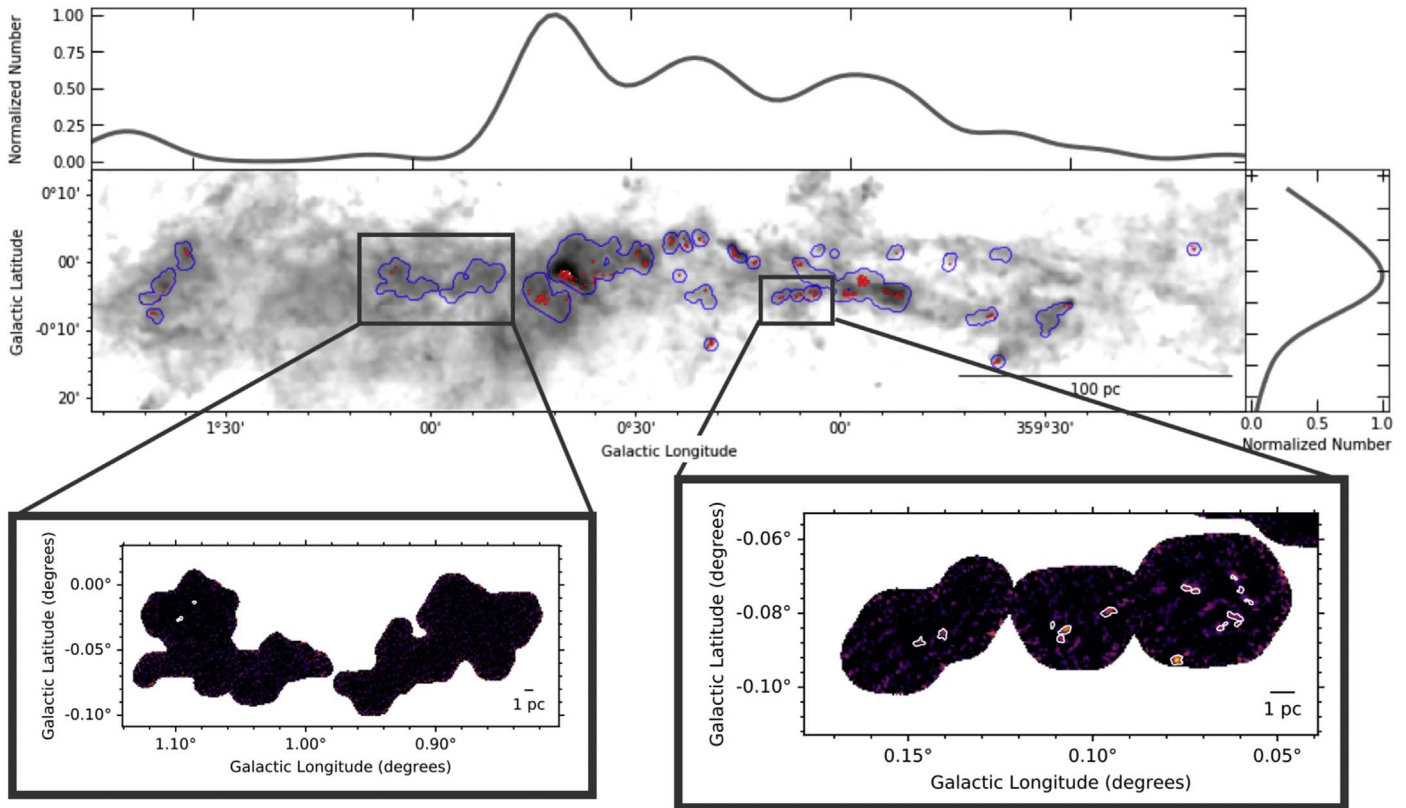


Figure 3. The kernel density estimate of the normalized number of compact sources from the robust catalog as a function of Galactic longitude (top) and latitude (right) with leaf contours over a map of Herschel (center). The highest concentrations of detected compact sources are in Sgr B2, the Dust Ridge, and the 50 and 20 km s⁻¹ clouds. Insets are shown of several regions with white catalog leaf contours over the SMA 1.3 mm dust continuum. The regions shown are G1.085-0.027, G1.038-0.074, and G0.891-0.048 on the bottom left, and G0.068-0.075, G0.106-0.082, and G0.145-0.086 (the three little pigs) on the bottom right.

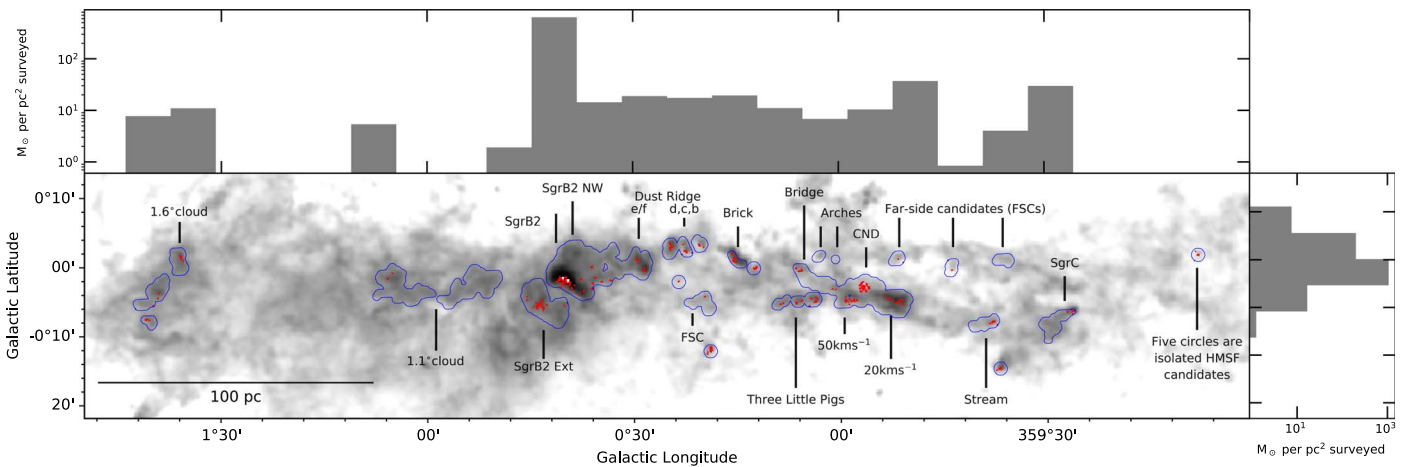


Figure 4. The average mass of compact sources per unit area surveyed from the robust catalog as a function of Galactic longitude (top) and latitude (right) with leaf contours (in red) and the surveyed area footprint (blue) over the Herschel $N(\text{H}_2)$ map from C. Battersby et al. (in preparation; bottom). Clouds/regions are also labeled by their colloquial names. We detect high concentrations of mass in compact sources per area surveyed in Sgr B2, Sgr C, the Dust Ridge, the 20 and 50 km s⁻¹ clouds, the Brick, and the three little pigs, while other regions of high column density lack a similar degree of compact substructure. This is quantified in Paper I (Section 5.4) as a compact dense gas fraction, calculated as both SMA flux divided by Bolocam Galactic Plane survey flux or SMA mass divided by Herschel mass. For an in-depth look at the compact dense gas fraction of these clouds, we refer the reader to Paper I.

substructure as measured by dendrogram leaf number seems highly dependent on the measured cloud scale column density.

While the catalog presented in this work does not include information about the star-forming properties of the leaves, the sharp increase in leaf occurrence at this threshold and its relationship to the star-forming properties of leaves will be further investigated in upcoming work. We note that the Herschel-derived

column densities saturate in some pixels toward the center of Sgr B2 and may not be representative of true column densities in that region, and so all pixels and leaves in the Sgr B2 region have been removed from this analysis, and are not included in the histogram in Figure 5. Because these leaves are all highly concentrated at high column densities, they do not significantly influence the trend described in this section.

Table 1
Small Subset of the Leaf Properties Table

Leaf ID	Area (arcsec ²)	<i>l</i> (deg)	<i>b</i> (deg)	<i>R</i> _{eff} (pc)	Integrated Flux (Jy)	Peak Cont. Flux (Jy Sr ⁻¹)	Mean Cont. Flux (Jy Sr ⁻¹)	Rms (Jy Sr ⁻¹)
G359.611+0.018c	45.00	-0.38	-0.25	0.15	6.71e-02	1.39e+08	6.35e+07	1.68e+07
G359.611+0.018a	259.25	-0.39	-0.24	0.36	2.09e+00	1.73e+09	3.43e+08	9.88e+06
G359.611+0.018b	34.75	-0.38	-0.24	0.13	7.52e-02	1.70e+08	9.20e+07	1.25e+07
G0.316-0.201c	38.75	0.32	-0.21	0.14	7.12e-02	1.28e+08	7.82e+07	1.91e+07
G0.316-0.201e	19.00	0.32	-0.20	0.10	3.88e-02	1.22e+08	8.68e+07	1.67e+07
G0.316-0.201g	15.50	0.31	-0.20	0.09	3.13e-02	1.27e+08	8.60e+07	1.35e+07
G0.316-0.201a	115.25	0.32	-0.20	0.24	9.12e-01	1.09e+09	3.37e+08	9.56e+06
G0.316-0.201k	8.00	0.31	-0.20	0.06	1.37e-02	8.82e+07	7.26e+07	1.36e+07
G0.316-0.201i	7.50	0.32	-0.20	0.06	2.28e-02	1.42e+08	1.29e+08	1.01e+07
G0.316-0.201l	4.75	0.31	-0.20	0.05	7.58e-03	7.40e+07	6.79e+07	1.07e+07

Note. The leaf properties table is available in full at doi:[10.7910/DVN/RDE1CH](https://doi.org/10.7910/DVN/RDE1CH).

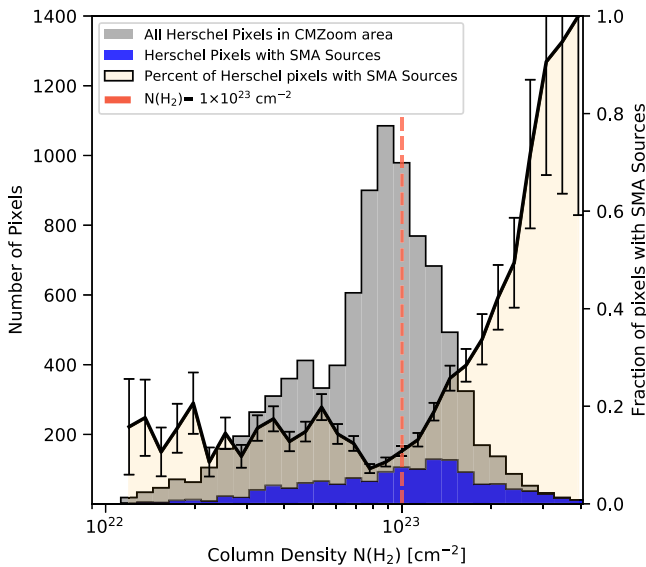


Figure 5. A histogram of the Herschel column density (derived using the procedure in Battersby et al. 2011, from C. Battersby et al. 2020, in preparation) associated with every pixel in the CMZoom survey. The gray histogram represents the column density associated with every pixel covered by the SMA survey, while the blue represents only the pixels within one Herschel beam (36'') of a catalog leaf centroid in the robust version of the catalog. The ratio of pixels within one Herschel beam of a source in the SMA catalog to the total number of pixels in the SMA's map for a given column density is shown as a solid line, which experiences a sharp uptick around a column density of $(1-2) \times 10^{23}$ (highlighted with the red dotted line at 10^{23}). The error bars represent the Poisson uncertainty from the number of Herschel pixels in each column density bin. Pixels in the Sgr B2 region have been excluded from this analysis due to their much higher column densities resulting from potentially unreliable fluxes (see Section 5.1). A similar analysis is available for Sgr B2 sources in Figure 15 of Ginsburg et al. (2018), looking instead at the fraction of material at a given column density that is associated with young stellar objects.

4.3. Physical Properties

The robust catalog is presented in Tables 1–3. We extract a number of physical properties of our cataloged leaves directly from the dendrogram structure using the `PP_Statistic` package in the `Astrodendro` package, including the minimum flux value, maximum flux value, integrated flux, structure area, and center position. The structure area is converted into an effective projected radius for each leaf using $R_{\text{eff}} \equiv (N_{\text{pix}} A_{\text{pix}} / \pi)^{1/2}$ (where N_{pix} is the number of pixels

associated with a given dendrogram leaf, and A_{pix} is the area of a single pixel). For each leaf, `Astrodendro` also extracts an effective ellipse major and minor axis, and position angle using moment analysis. For more information on the exact implementation of how these features are extracted, please refer to the `Astrodendro` documentation.²⁵ These approximate ellipse values are included in the catalog, but we use the effective leaf radius calculated above instead for the following analysis.

Assuming optically thin dust continuum emission, the beam-averaged column density at the location of peak emission, N_{H_2} is estimated for each dendrogram leaf using

$$N_{\text{H}_2} = \frac{F_{\nu}^{\text{peak}} R_{\text{gd}}}{\mu_{\text{H}_2} m_{\text{H}} \kappa_{\nu} B_{\nu}(T_{\text{d}})}, \quad (2)$$

where F_{ν}^{peak} is the peak flux density over the leaf (in janskys beam⁻¹), R_{gd} is the gas to dust mass ratio (for which we assume a value of 100; e.g., Battersby et al. 2011), mean atomic weight $\mu_{\text{H}_2} = 2.8$ (e.g., Kauffmann et al. 2008), and m_{H} is the mass of atomic hydrogen, κ_{ν} is the dust opacity per unit mass at a frequency ν , and $B_{\nu}(T_{\text{d}})$ is the Planck function at a dust temperature, T_{d} .

The dust opacity per unit mass is determined from $\kappa_{\nu} = \kappa_0 (\nu / \nu_0)^{\beta}$, assuming a dust emissivity index, β , where κ_0 is based on the moderately coagulated thin ice mantle dust model of Ossenkopf & Henning (1994) at a frequency, ν_0 . At a frequency of ~ 226 GHz, we adopt a value of $\kappa_{\nu} \approx 0.867 \text{ cm}^2 \text{ g}^{-1}$ (extrapolating from $\kappa_0 = 0.899 \text{ cm}^2 \text{ g}^{-1}$ at $\nu_0 = 230$ GHz with $\beta = 1.75$; e.g., Battersby et al. 2011). While we do not have a measure of dust temperature (T_{d}) at the resolution of our SMA observations, we use corresponding dust temperatures derived from Herschel measurements. This temperature map and the Herschel column density map were generated using the dust SED fitting procedure outlined in Battersby et al. (2011), and further detailed in Mills & Battersby (2017). These local temperature estimates have a resolution of 36'', while higher-resolution temperature maps derived using the point process mapping (PPMAP) procedure (Marsh et al. 2017) differ only by a few degrees kelvin for the set of pixels considered in this work. This small change in temperature on smaller scales affects the derived physical properties significantly less than the systematic errors described below.

²⁵ <https://dendrograms.readthedocs.io/en/stable/>

Table 2
Small Subset of the Leaf Properties Table (Continued)

Leaf ID	N_{Herschel} (cm^{-2})	N_{SMA} (cm^{-2})	Mass ($M_{\odot}$)	Mass (bg. sub.) ($M_{\odot}$)	n (cm^{-3})	ρ (g cm^{-3})	t_{ff} (yr)
G359.611+0.018c	6.81e+22	1.42e+23	1.03e+02	4.16e+01	1.05e+05	4.92e-16	9.49e+04
G359.611+0.018a	1.46e+23	1.73e+24	3.15e+03	2.62e+03	2.33e+05	1.09e-15	6.37e+04
G359.611+0.018b	8.59e+22	1.73e+23	1.14e+02	4.28e+01	1.72e+05	8.07e-16	7.41e+04
G0.316-0.201c	3.86e+22	1.24e+23	1.04e+02	4.97e+01	1.32e+05	6.21e-16	8.45e+04
G0.316-0.201e	4.65e+22	1.18e+23	5.62e+01	1.57e+01	2.09e+05	9.80e-16	6.73e+04
G0.316-0.201g	4.01e+22	1.03e+23	3.82e+01	1.10e+01	1.93e+05	9.03e-16	7.01e+04
G0.316-0.201a	5.74e+22	9.00e+23	1.13e+03	7.31e+02	2.82e+05	1.32e-15	5.79e+04
G0.316-0.201k	5.94e+22	6.94e+22	1.61e+01	2.08e+00	2.19e+05	1.03e-15	6.57e+04
G0.316-0.201i	7.08e+22	1.21e+23	2.91e+01	2.22e+00	4.36e+05	2.04e-15	4.66e+04
G0.316-0.201l	5.94e+22	5.82e+22	8.94e+00	6.08e-01	2.66e+05	1.25e-15	5.96e+04

Note. The leaf properties table is available in full at doi:[10.7910/DVN/RDE1CH](https://doi.org/10.7910/DVN/RDE1CH).

Table 3
Small Subset of Leaf Property Uncertainties

Leaf ID	Mass Unc. ($M_{\odot}$)	N_{SMA} Unc. (cm^{-2})	n Unc. (cm^{-3})	ρ Unc. (g cm^{-3} yr)	t_{ff} Unc. (yr)
G359.611+0.018c	3.15e+01	2.85e+23	4.47e+04	2.09e-16	4.04e+04
G359.611+0.018a	4.92e+02	3.48e+24	7.85e+04	3.68e-16	2.15e+04
G359.611+0.018b	2.34e+01	3.47e+23	6.20e+04	2.90e-16	2.67e+04
G0.316-0.201c	2.99e+01	2.49e+23	5.76e+04	2.70e-16	3.67e+04
G0.316-0.201e	1.38e+01	2.37e+23	8.48e+04	3.97e-16	2.73e+04
G0.316-0.201g	8.26e+00	2.06e+23	8.90e+04	4.17e-16	3.23e+04
G0.316-0.201a	1.72e+02	1.80e+24	1.15e+05	5.41e-16	2.37e+04
G0.316-0.201k	3.84e+00	1.39e+23	1.09e+05	5.09e-16	3.25e+04
G0.316-0.201i	4.91e+00	2.43e+23	1.78e+05	8.34e-16	1.90e+04
G0.316-0.201l	1.93e+00	1.17e+23	1.27e+05	5.95e-16	2.84e+04

Note. The leaf properties table is available in full at doi:[10.7910/DVN/RDE1CH](https://doi.org/10.7910/DVN/RDE1CH).

The SMA derived column densities range from $2.9 \times 10^{22} \text{ cm}^{-2} < N_{\text{H}_2} < 9.0 \times 10^{25} \text{ cm}^{-2}$, with a median value of $N_{\text{H}_2} \sim 1.2 \times 10^{23} \text{ cm}^{-2}$. The high end of this distribution is dominated by two leaves with column densities $6.2 \times 10^{25} \text{ cm}^{-2}$ and $9.0 \times 10^{25} \text{ cm}^{-2}$. These extreme sources are within the Sgr B2 complex, and are colloquially known as Sgr B2 main (M) and north (N), respectively.

Following the method described in Battersby et al. (2010), we find an isothermal estimate for the mass of each leaf using

$$M_{\text{leaf}} = \frac{d^2 S_{\nu} R_{\text{gd}}}{\kappa_{\nu} B_{\nu}(T_d)}, \quad (3)$$

where T_d is the local dust temperature estimate, d is the distance to the source (~ 8.178 kpc; The GRAVITY Collaboration et al. 2019), and S_{ν} is the integrated leaf flux (in janskys), κ_{ν} is the dust opacity per unit mass discussed above, and $B_{\nu}(T_d)$ is the Planck function at the local dust temperature, T_d . The resultant estimates for leaf masses range from $\sim 4.1 M_{\odot} < M_{\text{leaf}} < \sim 7.6 \times 10^4 M_{\odot}$, with a median value of $\sim 86 M_{\odot}$. As with the column densities, the two most massive leaves in the catalog, Sgr B2 north and main, lie well apart from the rest of the mass distribution of leaves (see Figure 6). Histograms of the physical properties are shown in Figure 7.

Two estimates for the leaf mass are included in the catalog: the upper-bound leaf mass (reported above) and the background-subtracted leaf mass. The upper-bound mass is calculated using

the total integrated flux from the pixel values for the whole leaf. This mass estimate assumes that all of the flux recovered in the leaf belongs to one, spatially cohesive clump, ignoring the more diffuse envelope surrounding the clump. Our alternative, lower-mass estimates are calculated by subtracting away the lowest flux value in the leaf from each pixel in the leaf, removing all flux associated with the envelope (similar to Henshaw et al. 2016a). This “background-subtracted” mass assumes that each leaf’s lowest value (the outer contour of the leaf) is a good approximation for the average flux for the background emission. As a consequence, this background-subtracted mass ignores any contribution from the lower dendrogram structures to the source mass, and therefore is likely to be an underestimate of the total mass.

It is also important to note that both masses calculated in this way are subject to a systematic uncertainty of a factor of approximately two due to uncertainty in the dust opacity as described in detail in Battersby et al. (2010). The uncertainties associated with each of the leaf properties (see Table 3) account for this dominant systematic uncertainty as well as fluctuations in the Herschel dust temperatures, uncertainty in the Galactic Center distance assumption, and the local noise estimates in the 1.3 mm continuum flux. We assume an uncertainty in the distance to any given catalog source be $\sim \pm 240$ pc, as this corresponds to the maximum longitudinal extent of sources identified in the survey, as the line-of-sight distance is not well known for many of the clouds hosting these sources.

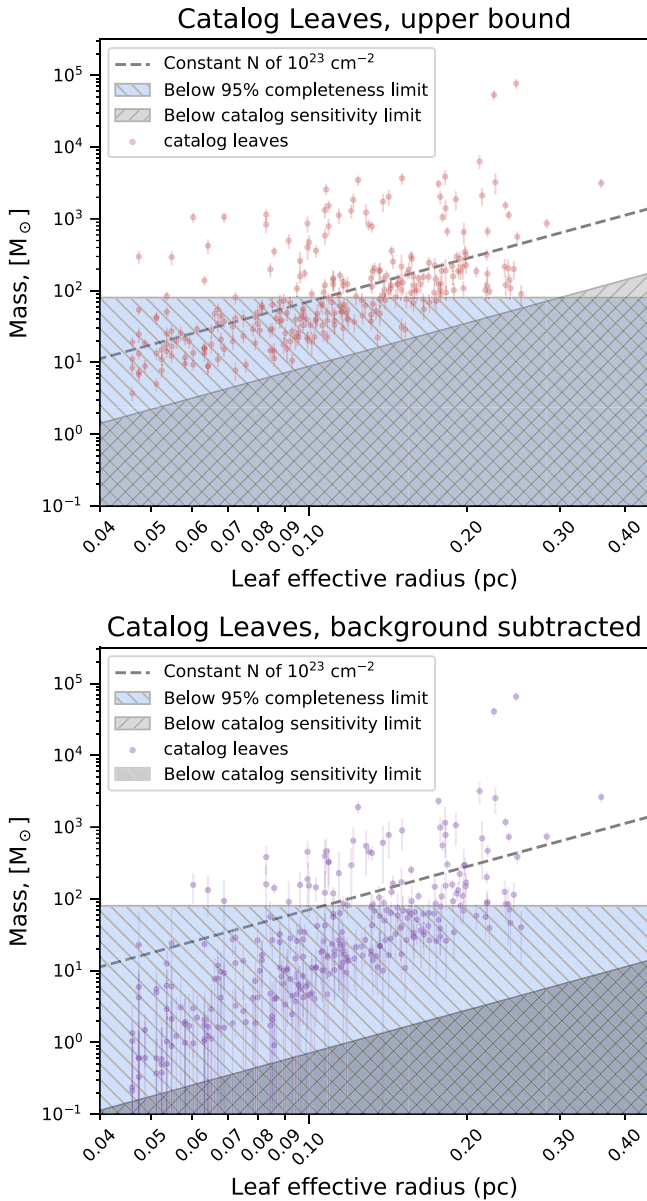


Figure 6. Two representations of the mass ($M_{\odot}$) vs. radius (parsecs) for each leaf in the robust version of the catalog. The upper panel shows the mass calculated using the flux achieved by integrating over the entire leaf structure, acting as an upper bound for the true structure mass. The bottom panel displays the background-subtracted leaf masses, calculated by subtracting the minimum value of the leaf and integrated over the entire leaf area. This background-subtracted mass is likely an underestimate for the true structure mass. Both panels show a line of constant column density $n = 10^{23} \text{ cm}^{-2}$ for reference. The range of masses below 95% completeness is hatched right and shaded blue, and the region below minimum flux possible for a cataloged leaf is hatched left and shaded gray.

The equivalent particle number density averaged over a leaf with radius R_{eff} , and mass, M , can be estimated using

$$n_{\text{H}_2} = \frac{M}{\frac{4}{3}\pi R_{\text{eq}}^3 \mu_{\text{H}_2} m_{\text{H}}}, \quad (4)$$

assuming the line-of-sight radius is equivalent to the radius on the plane of the sky. The range in particle number density is $2.3 \times 10^4 \text{ cm}^{-3} < n_{\text{H}_2} < 1.7 \times 10^7 \text{ cm}^{-3}$, with a median value of $n_{\text{H}_2} = 2.0 \times 10^5 \text{ cm}^{-3}$. The corresponding range in

the local freefall time is calculated by

$$t_{\text{ff}} = \left(\frac{3\pi}{32G\mu_{\text{H}_2} m_{\text{H}} n_{\text{H}_2}} \right)^{1/2}, \quad (5)$$

where G is the gravitational constant, and the remaining constants are the same as defined above. For the robust catalog, dendrogram leaves range between $7.4 \times 10^3 \text{ yr} < t_{\text{ff}} < 2.0 \times 10^5 \text{ yr}$, with a mean value of $t_{\text{ff}} = 6.9 \times 10^4 \text{ yr}$. Each of the above properties is displayed for a subset of leaves in Table 2. Uncertainties derived from propagating the dust opacity, assumed Galactic Center distance ($\pm 240 \text{ pc}$, as discussed above), typical fluctuations in Herschel dust temperature, and the local noise estimate for our 1.3 mm dust continuum flux measurements are presented in Table 3.

4.4. Mass–Radius Comparison to Previous Studies of the CMZ and Galactic Disk

As previously mentioned, our cataloged sources occupy a spatial scale ranging between ~ 0.04 and 0.4 pc , meaning they largely constitute objects intermediate to clumps and cores. We briefly compare the mass–radius distribution of our robust catalog sources to clumps and cores identified by previous observational catalogs in Figure 8. Our sources largely agree with the mass–radius distribution of previous studies of the CMZ. Many of these previous studies (e.g., Walker et al. 2018; Lu et al. 2019a) aim to identify star-forming objects and their properties, and CMZ sources with known star formation sites seem to have a higher mass per radius than the bulk of our catalog. This trend will be investigated further in upcoming work (H P. Hatchfield et al. 2020, in preparation).

4.5. High-completeness Catalog

As an alternative to the primary, robust catalog released with this paper, we also present a more lenient catalog that prioritizes greater completeness over a low false-positive rate. We use the same cataloging algorithm described in Section 3, while adjusting the parameters to suit the new design philosophy. Specifically we lower the pruning parameters from the high-reliability version,

$$\text{Flux}_{1.3 \text{ mm}}^{\text{peak}} > 6\sigma_{\text{RMS}},$$

where $\text{Flux}_{1.3 \text{ mm}}^{\text{peak}}$ is the peak pixel flux, and σ_{rms} is the local rms noise estimate, to

$$\text{Flux}_{1.3 \text{ mm}}^{\text{peak}} > 4\sigma_{\text{RMS}}.$$

The requirement that each leaf’s mean flux be $>2\sigma$ above the local rms applied to both catalogs. This reduction in the maximum required per pixel flux results in a significant increase in the number cataloged leaves in the high-completeness catalog compared to the high-reliability catalog. The change produces a catalog of 816 leaves as compared to the robust catalog with its 285 leaves. Despite this large change in the number of leaves, most of the leaves exclusive to the high-completeness catalog have relatively low mass and together constitute an additional $2.8 \times 10^4 M_{\odot}$, which is only 13% of the robust catalog’s total mass. The vast majority of mass is cataloged in both versions.

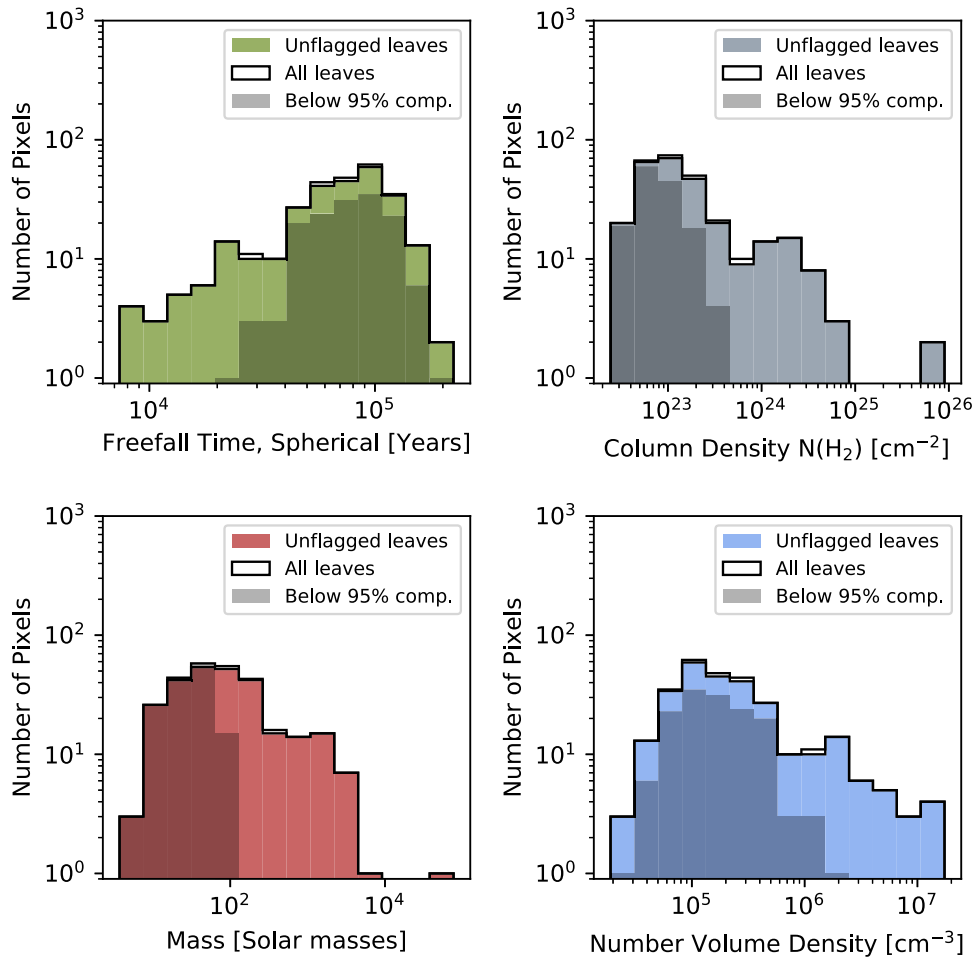


Figure 7. Histograms of the physical properties for the leaves contained in the high-reliability version of the catalog. These properties are calculated according to the methods described in Section 4.3. The upper left panel shows the distribution of freefall times calculated by assuming a spherical distribution of gas with the leaf’s effective radius. The upper right panel shows the distribution of column densities for each leaf. The distribution of leaf masses is shown in the bottom left panel, and the volume density distribution is shown in the bottom right panel, which is calculated again by assuming a spherical volume distribution according to the leaf’s effective radius. The darker shaded regions correspond to the portion of leaves below the catalog’s 95% mass completeness limit. Leaves flagged as suspicious for being within 15 pixels of the map’s edge are included as a non-filled histogram, though they largely follow the same trends as the more robust leaves. A roughly bimodal distribution is apparent, and its origin is explored in Section 5.1.

The more diffuse nature of emission in the CMZoom field as well as non-Gaussian components in the noise resulting from primary beam correction and cleaning artifacts together make it difficult to measure the trade-off of completeness versus false-positive rate using the same simulated observation procedure outlined in Section 3.2. The assumptions inherent in this procedure, namely the cataloging of a random spatial distribution of uniform mass point sources, might be ill suited for characterizing the lower flux component of emission in the CMZoom field. The simulated observations result in a steep increase in the number of false positives and a small increase in completeness for changes in the pruning, with a 95% completeness of $\sim 50 M_{\odot}$ (assuming $T_d = 20$ K).

Additionally, a large number of new leaves are included in the catalog that are in immediate proximity to the edge of the CMZoom maps, which is suspicious due to more extreme fluctuations in the noise at the map edges resulting from the primary beam correction process. To mitigate this, we have flagged all leaves with centroids within 15 pixels of the map’s edge, and confirmed that the removal of these leaves does not seem to change the overall distribution in the physical property

statistics. We include recreated versions of the physical property figures in Appendix A.

Qualitatively, the distribution of the leaves’ physical properties is similar in both catalogs, with the catalog design changes mainly permitting the detection of a larger number of low-mass leaves.

5. Discussion

5.1. The Elephant in the Room: Sagittarius B2

The distribution of leaves in Figure 6 reveals a striking characteristic of the mass–radius relation of the catalog. There appears to be a population set apart from the majority in the mass–radius distribution of the leaves. This feature is resilient to changes in the catalog algorithm parameters, as detailed in Appendix B. Almost all of the points in the “higher-mass” mode are located in the region surrounding Sgr B2, the most intense site of star formation in the Galactic Center. In Figure 9 we compare the catalog divided into leaves associated with Sgr B2 and all others. It is clear from the histograms of physical properties that objects in Sgr B2 are detected with higher flux per leaf area than elsewhere in the surveyed region, leading to

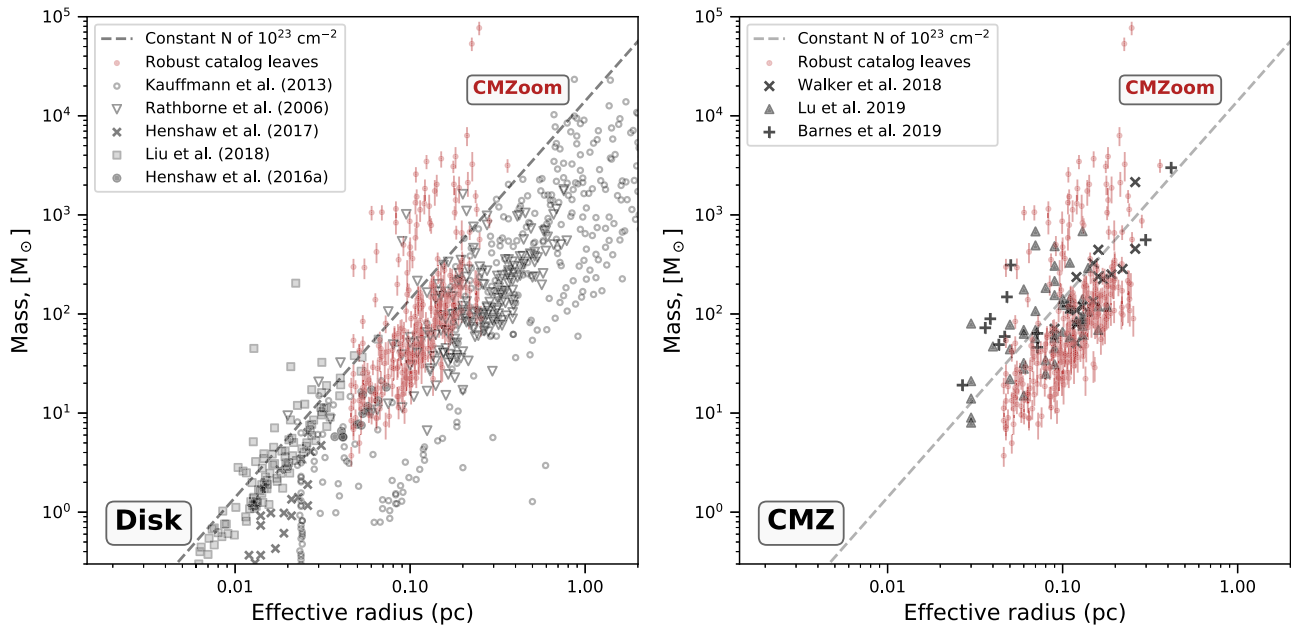


Figure 8. A mass–radius comparison of our catalog leaves (from the robust catalog, shown as red dots with mass uncertainties) with cataloged objects from previous work studying cores in the Milky Way’s disk (left panel) and in the Galactic Center (right panel). A line of constant column density 10^{23} cm^{-2} is shown plotted over the data for reference. For the Galactic disk, data are plotted from Rathborne et al. (2006), Kauffmann et al. (2013a), Henshaw et al. (2016a), Henshaw et al. (2017), and Liu et al. (2018). For the Galactic Center, data are plotted from Walker et al. (2018), Barnes et al. (2019), and Lu et al. (2019a).

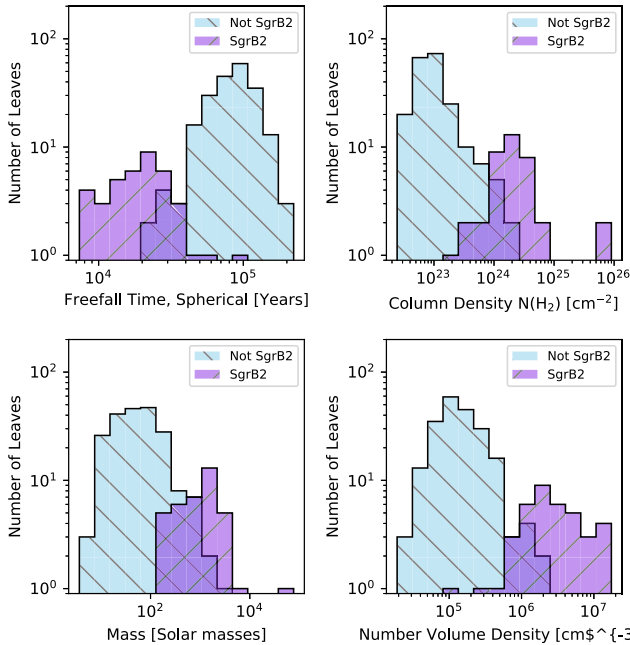


Figure 9. Separated histograms of the physical properties for the leaves within and apart from Sgr B2, for the robust version of catalog. These properties are calculated according to the methods described in Section 4.3, and separated according to the mask released with this work at <https://dataverse.harvard.edu/dataverse/cmzoom>.

higher masses, column and volume densities, and shorter freefall times. As measured, sources associated with Sgr B2 have more mass than the rest of the CMZ combined by nearly a factor of 10, assuming the validity of the Herschel dust temperatures used in the mass calculation from Section 4.3.

There are a few possible explanations for this extreme population. First, it is possible that there is significant contamination in the 1.3 mm flux from nondust emission. Of the possible sources of nondust continuum emission in Sgr B2,

the most likely culprit is free–free emission coming from H II regions. Previous observations have identified numerous H II regions distributed throughout the Sgr B2 complex (Mehringer et al. 1995; De Pree et al. 1996, 2015; Ginsburg et al. 2018; Lu et al. 2019b). Ginsburg et al. (2018) find the Sgr B2 complex’s 3 mm continuum emission to be well populated by both extended and compact H II regions. In particular, Sgr B2 main (identified in our catalog as leaf ID G0.699-0.028b, the second most massive source in the catalog) is dominated by H II region continuum flux at 3 mm (see Figure 3 and Table 4 of Ginsburg et al. 2018). While we expect free–free emission to be somewhat weaker at 1.3 mm, it is still likely the free–free emission accounts for some of the observed flux these regions.

Second, temperature has a strong effect on the inferred mass of our objects, so it is possible that the Sgr B2 area hosts higher dust temperatures than the Herschel estimates used in this work. In calculating the mass for each leaf, we assume the validity of the cloud scale dust temperature, so the difference between the two modes could be explained by a systematic difference in dust temperature on smaller scales. In Figure 10 we demonstrate how the mass of each leaf would vary with a temperature between the observed Herschel dust temperature (close to 20 K for most leaves) and 50 K, and then from 50 K to 150 K. If we assume $T_{\text{dust}} = \sim 120\text{--}150 \text{ K}$, the Sgr B2 source masses become consistent with the broader CMZ leaves’ population. The Herschel temperatures reported in this catalog agree with previous measurements of CMZ dust temperatures, typically ranging between 20 and 30 K (Pierce-Price et al. 2000; Etaluzze et al. 2013). The SMA data presented in this work probe smaller spatial scales than the Herschel maps (smallest beam FWHM = $5''.9$), so it is possible that different dust temperatures would be resolved on smaller spatial scales. Marsh et al. (2017) used the PPMAP procedure to derive temperature maps at a spatial resolution of $12''$, but these temperatures differ from the lower-resolution data by only a few degrees kelvin for pixels in the CMZZoom survey footprint.

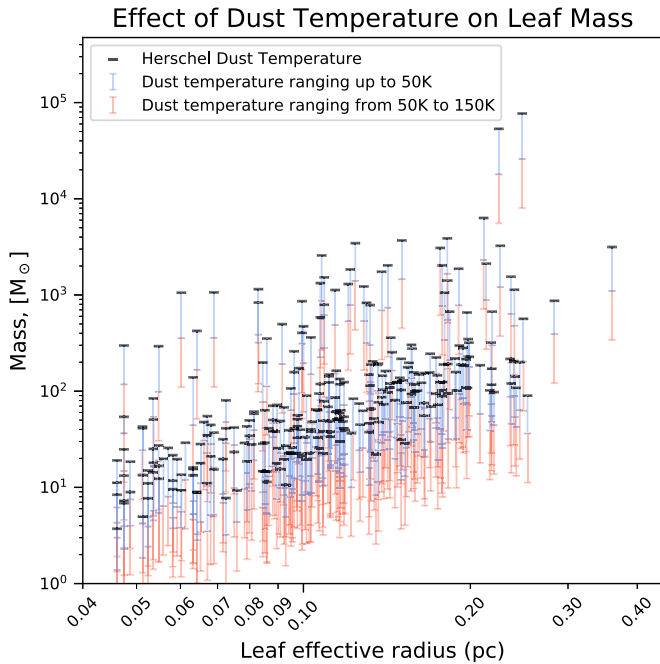


Figure 10. The mass ($M_{\odot}$) vs. radius (parsecs) for each leaf in the high-reliability version of the catalog for variations in the assumed dust temperatures. The black cap represents leaf mass for an assumed temperature of 20 K. The blue bar shows how much the mass of each leaf decreases for an assumed dust temperature of up to 50 K, and the red bar shows how much the mass decreases for a dust temperature of 150 K, which is higher than typically expected for dust in the CMZ (Pierce-Price et al. 2000; Etaluze et al. 2013). The temperature assumed would have to be close to 150 K to explain the separation of the Sgr B2 leaves’ distribution in the mass–radius relation. This is unlikely, as the multiwavelength modeling from Schmiedeke et al. (2016) derived considerably lower dust temperatures for most of the Sgr B2 complex.

While it is possible that even smaller spatial scales could reveal higher dust temperatures in areas of active or recent star formation such as Sgr B2, this is unlikely to have a widespread effect on our catalogs. It is also worth noting the Herschel temperatures in certain pixels in the Sgr B2 region are likely to be unreliable due to saturation and line contamination.

Lastly, it is possible that the measured masses and column densities are due to line-of-sight confusion. It has been suggested that the line-of-sight structure of the Sgr B2 complex could lead to complications in determining the mass of individual structures, with some evidence suggesting Sgr B2 north and main might have overlapping, elongated envelopes (Goldsmith et al. 1990; Schmiedeke et al. 2016). In this case, the background-subtracted masses calculated in Section 4.3 would be more appropriate for comparisons with the less-complicated regions in our catalog. These background-subtracted masses are in much better agreement with the rest of the catalog’s contents, as shown in Figure 11. This may suggest that the line-of-sight envelope effects contribute significantly to our reported masses in this region.

In reality, it is likely that a combination of these effects would impact the masses calculated in this work. Flux from H II regions certainly accounts for some of the flux in the two most massive leaves (leaf IDs G0.699-0.028a and G0.699-0.028b), but Ginsburg et al. (2018) found only 10 compact H II regions associated with sources other than Sgr B2 M, meaning H II regions are likely not solely responsible for the difference in mass–radius distribution, but may be more likely to explain the extreme mass we calculate for Sgr B2 N and M. As

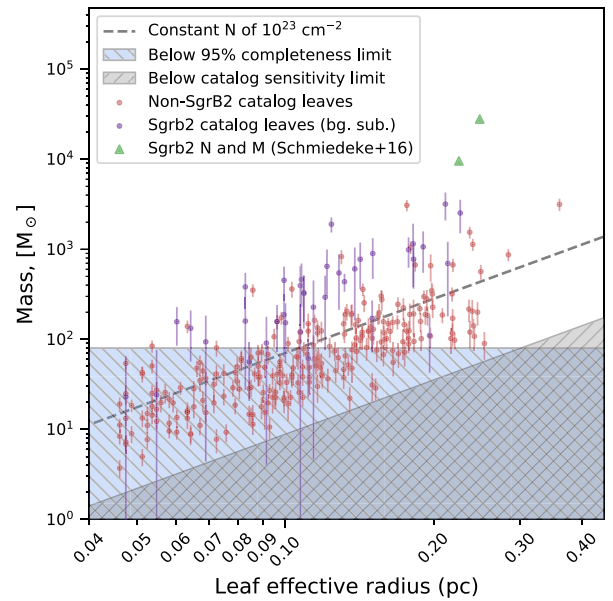


Figure 11. The mass ($M_{\odot}$) vs. radius (parsecs) for each leaf in the robust version of the catalog, using the background-subtracted mass values for leaves in the Sgr B2 complex (Sgr B2 leaves’ background-subtracted masses are shown in purple, the modified masses for Sgr B2 N and M from Schmiedeke et al. 2016 are shown as green triangles, while the rest of the catalog is shown in red). Using this background-subtracted mass mostly resolves the separation between the bimodal distribution of leaves in Figure 6, suggesting that the line-of-sight overlapping envelopes might be significantly affecting the mass of these leaves in particular (see Section 5.1). For comparison with Figure 6, the same line of constant column density $n = 10^{23} \text{ cm}^{-2}$ is shown. Again, the range of masses below 95% completeness is hatched right and shaded blue, and the region below minimum flux possible for a cataloged leaf is hatched left and shaded gray.

described in Figure 10, a dust temperature greater than 150 K would be required to fully explain the separation of these sources. While such high dust temperatures could exist for sources with extreme, ongoing star formation activity like Sgr B2 N and M, they are unlikely to be so high in this region unanimously. A combination of high local dust temperatures, line-of-sight envelope overlap, and non-continuum contamination together can readily explain the apparent divergence of these leaves from the more general distribution of CMZ compact structures in mass–radius space.

5.2. Maximum Star Formation Potential of the CMZ

The high-completeness catalog presented in this work constitutes more than 95% of compact sources with masses greater than $80 M_{\odot}$ (for the robust catalog, or $50 M_{\odot}$ for the high-completeness catalog) embedded in high column density gas ($\geq 10^{23} \text{ cm}^{-2}$) in CMZ clouds. Determining the star formation potential of any one of these objects is challenging due to the complexity of the environmental conditions in CMZ clouds. Some of these structures are well known to be actively star forming, while some others show tenuous signatures of star formation activity and many seem entirely quiescent. The connections between source properties and star formation signatures are investigated in upcoming work.

Because CMZzoom covers all high column density material in the CMZ, we suspect that these catalogs contain all possible sites of star cluster formation over a relevant timescale, up to our completeness limits. While it is not necessary that each of these compact sources will collapse down on a freefall time to

form stars, we can determine an upper limit on current star formation potential of the CMZ.

This star formation potential is the maximum possible SFR for the CMZ, calculated by assuming that each compact source in our catalog collapses on a freefall time and forms stars with an assumed star formation efficiency. Here this is done for the robust catalog, though the SFR estimated using the high-completeness catalog is similar, albeit with larger uncertainty. We choose a range of star formation efficiencies accounting for the wide range of proposed efficiencies on scales of clumps ($0.1 \leq \epsilon_{\text{clump}} \leq 0.3$) to the possible efficiencies of individual cores ($0.25 \leq \epsilon_{\text{core}} \leq 0.75$) since our compact sources are intermediate to these two hierarchical categories (Lada & Lada 2003; McKee & Ostriker 2007). We calculate our SFR limit for both the upper and lower bounds of this efficiency range, and report both. This choice of star formation efficiency is consistent with the results of Lu et al. (2019b), in which star formation efficiencies for gravitationally bound cores in Galactic Center clouds are measured to be ~ 0.3 , with a systematic error of about a factor of three. We calculate the SFR in two ways. First, we find an “individual” SFR for each compact source in the catalog using

$$\text{SFR}_{\text{indiv.}} = \sum_i \frac{\epsilon m_i}{t_{\text{ff},i}}, \quad (6)$$

where the sum is over each source in the catalog, ϵ is the star formation efficiency, m_i is the mass of the source, and $t_{\text{ff},i}$ is the freefall time of the source, calculated using Equation (5). Alternatively, we calculate the “mean” SFR for all sources over the mean freefall time, using

$$\text{SFR}_{\text{mean}} = \frac{\epsilon}{t_{\text{ff,mean}}} \sum_i m_i, \quad (7)$$

where $t_{\text{ff},i}$ is the mean freefall time of sources in the catalog (for the robust catalog, this value is 7.48×10^4 Myr).

To arrive at a more accurate upper value for the star formation potential, we use the background-subtracted values for masses of objects in the Sgr B2 complex. Additionally, we replace the extreme masses measured in this work (described in Section 4.3) for Sgr B2 N and M with more reliable gas masses derived in Schmiedeke et al. (2016) to arrive at the above estimates for the star formation potential. Without these modifications, Sgr B2 would completely dominate the SFR potential, exceeding the combined SFR of the entire Milky Way, which is not plausible. For an assumed efficiency of $\epsilon = 0.1$, we find

$$\text{SFR}_{\text{indiv.}} = \begin{cases} 0.29 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.1 \\ 2.2 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.75 \end{cases}$$

or using the mean freefall time,

$$\text{SFR}_{\text{mean}} = \begin{cases} 0.08 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.1 \\ 0.64 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.75 \end{cases}$$

The vast majority of mass generating this SFR potential is associated with leaves in the Sgr B2 complex, where we suspect our masses to be most unreliable for reasons discussed in Section 5.1.

If instead we consider the star formation potential of all cataloged sources outside the Sgr B2 complex, we find

$$\text{SFR}_{\text{indiv.}} = \begin{cases} 0.06 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.1 \\ 0.47 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.75 \end{cases}$$

or using the mean freefall time,

$$\text{SFR}_{\text{mean}} = \begin{cases} 0.04 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.1 \\ 0.30 M_{\odot} \text{ yr}^{-1} & \text{for } \epsilon = 0.75. \end{cases}$$

We conclude that the maximum star formation potential of the CMZ is between 0.08 and $2.20 M_{\odot} \text{ yr}^{-1}$, or, excluding Sgr B2, between 0.04 and $0.47 M_{\odot} \text{ yr}^{-1}$. We wish to emphasize that these SFR potentials may not be accurate predictions of the CMZ’s future SFR, as some of these objects may not ultimately collapse to form stars. Because the CMZ appears to be a highly complex environment, the longevity and boundedness of any given compact structure might be compromised by turbulence, dynamical interactions like inflow and shear, as well as other disrupting mechanisms. Therefore, these SFR potential estimates should be understood as an upper limit on the possible future SFR for the assumed star formation efficiency. Of the entire SFR of the Milky Way, estimated to be around $1.9 \pm 0.4 M_{\odot} \text{ yr}^{-1}$ (Chomiuk & Povich 2011), present tracers of star formation in the CMZ indicates that $0.06 M_{\odot} \text{ yr}^{-1}$ of this star formation activity occurs within the central 500 pc of the Galaxy (e.g., Longmore et al. 2013a; Barnes et al. 2017). The CMZ might roughly maintain the current SFR if we assume a low star formation efficiency for these objects, or alternatively if the dense structures required longer timescales than their freefall times to collapse and form stars (e.g., Lu et al. 2019b), and if some fraction of compact substructure remains unbound and never collapses to form stars.

5.3. High-mass Star Precursor Completeness

One of the key goals of the CMZoom survey is to compile potential sites of massive star formation as completely as possible. There is substantial evidence suggesting that massive stars largely form in high column density material (e.g., Kauffmann & Pillai 2010; Lada et al. 2010), and CMZoom covers all of the CMZ’s high column density material. The simulated observations performed in Section 3.2 are used to estimate how many leaves in the CMZoom survey area might be excluded from the catalog as a function of the leaf mass (see Figure 2). If we can estimate the number of missing leaves as a function of mass, and make the assumption that each of these leaves is destined to form stars according to some IMF, we can estimate an upper limit on the missing number of massive star precursor objects missed by the catalog, thus deriving our completeness for high-mass stellar precursors in the CMZ.

The curve in Figure 2 gives an estimate for the detected percentage of leaves of a given mass. The fraction of leaves missing at any given mass is equal to $1 - P(m)$ where $P(m)$ is the completeness as a function of leaf mass m . Because $P(m)$ is defined discretely for a subset of leaf masses, we bin the catalog leaf masses together with the nearest defined value for $P(m)$. This is a poor approximation for small leaf masses, where the completeness percentage changes rapidly as a function of mass. However, these lower-mass objects are also unlikely to produce many high-mass stars, so this approximation does not significantly affect our results for the high-mass stellar precursor completeness.

By sampling the Kroupa IMF (from Equation (2) of Kroupa 2001), we can convert these suspected missing leaf counts to an upper limit on the number of missed massive star precursors. We calculate the number of missing massive stellar precursors, denoted N_{miss} , as

$$N_{\text{miss}} = \sum_m (1 - P(m)) N_{\text{leaves}}(m) N_{\geq 8 M_{\odot}}(\epsilon_{\text{SF}} m), \quad (8)$$

where $P(m)$ is the completeness percentage for catalog leaves with mass m , $N_{\text{leaves}}(m)$ is the number of leaves binned to the nearest mass m with a defined value of $P(m)$, and $N_{\geq 8 M_{\odot}}(\epsilon_{\text{SF}} m)$ is the number of massive stars in a cluster with a final star mass defined by assumed star formation efficiency ϵ_{SF} and leaf mass m . To determine $N_{\geq 8 M_{\odot}}(\epsilon_{\text{SF}} m)$, we draw samples from the Kroupa IMF distribution until we reach the target cluster mass, and then either keep or discard the final draw to produce a total mass closest to the target cluster mass (this is identical to the “STOP_NEAREST” method in Appendix A of Krumholz et al. 2015). We then average the number of massive ($\geq 8 M_{\odot}$) stars for 1000 iterations of each leaf mass bin. This is performed for mass bins corresponding to each of the masses used in the simulated observation procedure (i.e., the masses in Figure 2), and star formation efficiencies ranging between 0.1 and 0.75, as in Section 5.2.

We choose the same range of star formation efficiencies as in Section 5.2. Multiplying this average by the estimated number of missing leaves at each mass bin, we arrive at an estimated number of missing massive stellar precursors, which we find to be $\lesssim 1\%$ (if we assume a star formation efficiency of 0.75) or $\approx 0\%$ (if we assume a star formation efficiency of 0.1). In summary, we estimate that CMZoom is complete to more than 99% of sites capable of forming massive stars embedded in high column density gas within the innermost ~ 500 pc of the Galaxy.

6. Summary

In this work, we produce the most complete catalogs to date of compact sources in the cold dust continuum across the Milky Way’s innermost 500 pc. These compact sources are identified in the 1.3 mm CMZoom survey as the leaves produced by a dendrogram algorithm, pruned according to local noise conditions. We present two catalogs:

1. a “robust” catalog designed for high accuracy, with sources required to have a peak 1.3 mm dust continuum flux 6σ above the local rms noise estimate. The robust catalog is $>95\%$ complete for compact sources with mass $\geq 80 M_{\odot}$.
2. a “high-completeness” catalog, designed to include more of the lower flux compact structures in the CMZoom field, which demands a larger tolerance for false-positive detections. The high-completeness catalog requires a peak 1.3 mm dust continuum flux 4σ above the same local rms noise estimate, and is $>95\%$ complete for compact sources with mass $\geq 50 M_{\odot}$.

The above mass completeness estimates assume a dust temperature that we take to be a typical value for Herschel-derived dust temperatures, 20 K. Since temperatures on smaller scales could be greater, these 95% mass completeness estimates should be understood as upper limits on our true completeness masses. By treating each of these leaves as an eventual site of

star formation (an intentionally generous assumption), along with an IMF for the resulting stars, we place an upper limit on the number of possible high-mass star precursors missing from the robust catalog, which we find to be $\lesssim 1\%$.

For objects included in both the robust and complete catalogs, we find a bimodal distribution in a number of physical properties, separating the population of sources associated with Sgr B2 from the sources elsewhere in the CMZ. This difference is also resilient to changes in the cataloging procedure (e.g., Figure 15). This separation is likely caused by a combination of several factors: increased dust temperatures on scales smaller than the Herschel measurements used in this work, overlapping envelopes and line-of-sight degeneracy of Sgr B2 clouds, and nondust continuum contamination from extended and compact H II regions distributed throughout the cloud complex.

We calculate a limit on the maximum possible star formation potential of the Milky Way’s CMZ, assuming that stars are not formed in this environment in isolation or in gas with a column density lower than $10^{23} N(\text{H}_2) \text{ cm}^{-2}$. Accounting for a range of possible star formation efficiencies for our cataloged sources, we find a maximum star formation potential of $0.08\text{--}2.2 M_{\odot} \text{ yr}^{-1}$. This SFR is dominated by the Sgr B2 complex, where more reliable SFR estimates are available (e.g., Schmiedeke et al. 2016; Ginsburg et al. 2018), and if we exclude Sgr B2, we find that the maximum SFR potential for the rest of the CMZ is $0.04\text{--}0.47 M_{\odot} \text{ yr}^{-1}$. This SFR potential ranges from near current estimates for active star formation in the CMZ to much higher values rivaling the entire Milky Way’s SFR, if a high star formation efficiency is expected for these compact structures.

Many of these sources are known to be sites of active star formation, and the relationship between star formation tracer activity and the line emission properties for these catalogs will be explored in upcoming work (H P. Hatchfield et al. 2020, in preparation). Future studies detailing the nature of these sources, their star formation activity, chemistry, outflow properties, gas temperatures, and kinematics will serve as an important stepping stone toward understanding star formation processes from the Milky Way GMCs to the great diversity of extragalactic environments.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. We acknowledge and thank the staff of the SMA for their assistance, and we thank Ray Blundell for his ongoing support of the CMZoom project. We thank Ken Young (Taco), Charlie Qi, and Mark Gurwell for their help working with the SMA data. We thank Glen Petitpas and Ryan Howie for scheduling these observations and Shelbi Holster and the SMA operators for making them happen. We thank Margaret Simonini and Anna Huang for their logistical assistance with this project and associated travel. We thank Jimmy Castaño, Liz Gehret, Mark Graham, Elizabeth Gutiérrez, Dennis Lee, and Irene Vargas-Salazar for their early investigations into these data that helped to guide its progress and improve its quality. This research has made use of NASA’s Astrophysics Data System. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This paper makes use of the following ALMA data: ADS/JAO. ALMA#2013.1.00617.S, ALMA#2013.1.00269.,

ALMA#2017.1.00687.S, and ALMA#2018.1.00850.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ.

HP.H. gratefully acknowledges support from the National Science Foundation under award No. 1816715. HP.H. thanks the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining grant No. 1829740, the Brinson Foundation, and the Moore Foundation; his participation in the program has benefited this work. HP.H. also acknowledges many helpful conversations and valuable advice from Imad Pasha and Matthew Phelps. C.B. gratefully acknowledges support from the National Science Foundation under award Nos. 1602583 and 1816715. D.W. also gratefully acknowledges support from the National Science Foundation under award No. 1816715. A.T.B. would like to acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 726384/Empire). A.G. acknowledges support from the NSF under grant 2008101. J. M.D.K. gratefully acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through an Emmy Noether Research Group (grant No. KR4801/1-1) and the DFG Sachbeihilfe (grant No. KR4801/2-1), as well as from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program via the ERC Starting grant MUSTANG (grant agreement No. 714907). X.L. acknowledges support by JSPS KAKENHI grant Nos. 18K13589 & 20K14528. E.A.C. M. gratefully acknowledges support by the National Science Foundation under grant No. AST-1813765. L.H. acknowledges

support by the National Science Foundation of China (11721303, 11991052) and the National Key R&D Program of China (2016YFA0400702).

Facilities: SMA, ALMA, Herschel, CSO.

Software: This research made use of APLpy, an open-source plotting package for Python (Robitaille & Bressert 2012; Robitaille 2019), astropy, a community-developed core Python package for astronomy (Astropy Collaboration et al. 2013, 2018), glueviz (Robitaille et al. 2017; Beaumont et al. 2015), CASA (McMullin et al. 2007), MIR, ²⁶ NumPy (Van Der Walt et al. 2011), and Miriad (Sault et al. 1995).

Appendix A High-completeness Catalog Figures and Tables

Here we present versions of Figures 5, 7, and 6 as well as Tables 1–3 for the high-completeness version of the catalog. Figure 12 presents the histogram of Herschel column densities and fraction of pixels with a leaf for the complete catalog. Similar to Figure 5, there is a sudden uptick in the frequency of Herschel pixels containing at least one leaf around $N = 10^{23} \text{ cm}^{-2}$, though the baseline frequency is higher, as would be expected from a more complete catalog with a greater number of leaves. Figure 13 shows the histograms of freefall time, column density, mass, and volume density. The distributions are largely similar to Figure 7, with a larger number of leaves flagged as suspicious due to their proximity to the CMZoom map edges. Figure 14 displays the mass–radius distribution of complete catalog leaves, for their upper-limit masses and background-subtracted masses, following largely the same trends as in Figures 6 and discussed in Section 5.1. Tables 4–6 have also been made available online at doi:10.7910/DVN/RDE1CH.

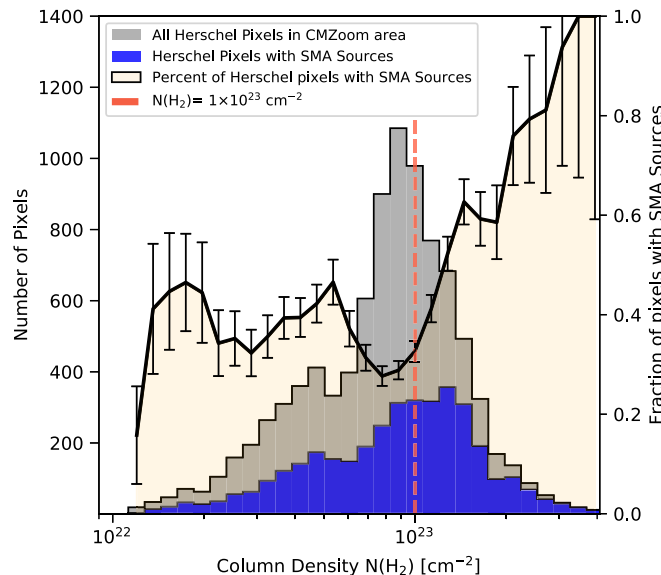


Figure 12. A histogram of the Herschel column density (derived using the procedure from Battersby et al. 2011 and C. Battersby et al. 2020, in preparation) associated with every pixel in the CMZoom survey. The gray histogram represents the column density associated with every pixel covered by the SMA survey, while the blue histogram represents only the pixels within one Herschel beam ($36''$) of a catalog leaf centroid in the high-completeness version of the catalog. The ratio of pixels within one Herschel beam of a source in the SMA catalog to the total number of pixels in the SMA’s map for a given column density is shown as a solid line, which experiences a sharp uptick around a column density of $2 \times 10^{23} \text{ cm}^{-2}$. Pixels in the Sgr B2 region have been excluded from this analysis due to their much higher column densities.

²⁶ <https://www.cfa.harvard.edu/cqi/mircook.html>

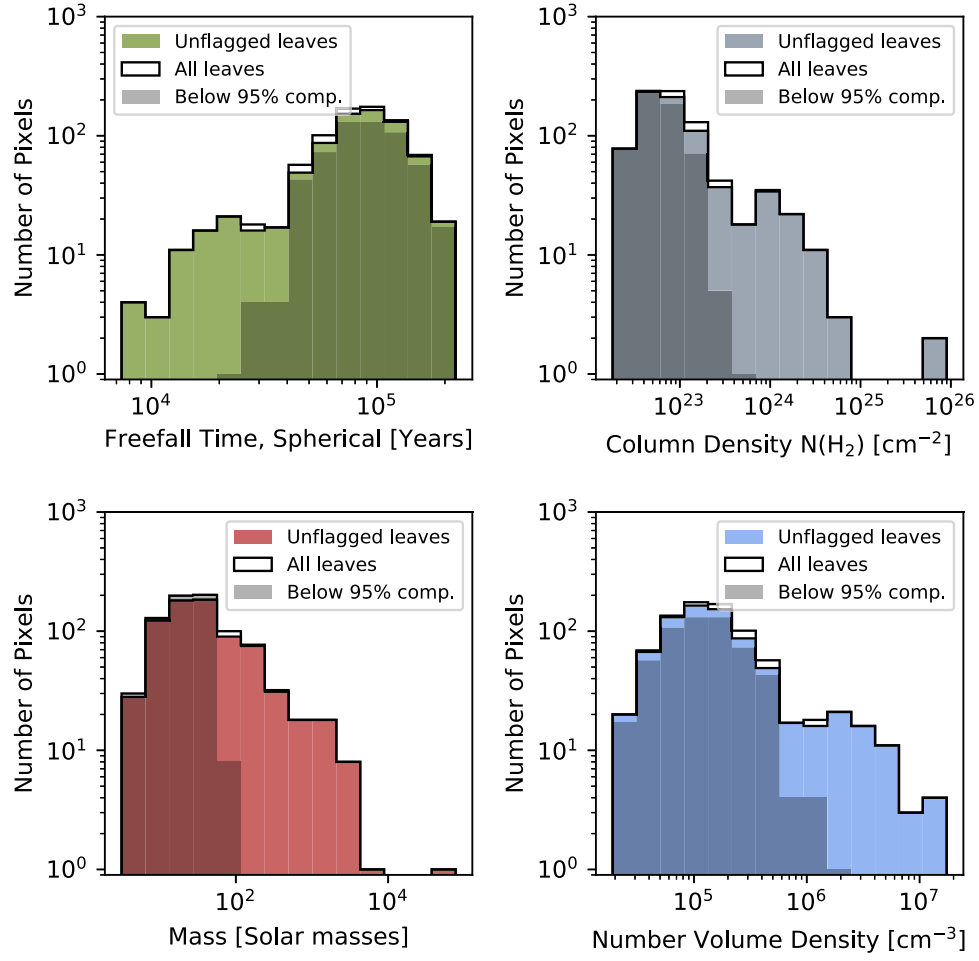


Figure 13. Histograms of the physical properties for the leaves contained in the high-completeness version of the catalog. These properties are calculated according to the methods described in Section 4.3. The upper left panel shows the distribution of freefall times calculated by assuming a spherical distribution of gas according to the leaf’s effective radius. The upper right panel shows the distribution of column densities for each leaf. The distribution of leaf masses is shown in the bottom left panel, and the volume density distribution is shown in the bottom right, which is calculated again by assuming a spherical volume distribution according to the leaf’s effective radius. The darker shaded regions correspond to the portion of leaves below the catalog’s 95% mass completeness limit of $60 M_{\odot}$. Leaves flagged as suspicious for being within 15 pixels of the map’s edge are included as a non-filled histogram, though they largely follow the same trends as the more robust leaves.

Table 4
Leaf Properties, High Completeness

Leaf ID	Area (arcsec ²)	l (deg)	b (deg)	R_{eff} (pc)	Integrated Flux (Jy)	Peak Cont. Flux (Jy Sr ⁻¹)	Mean Cont. Flux (Jy Sr ⁻¹)	Rms (Jy Sr ⁻¹)
G359.611+0.018c	45.00	−0.38	−0.25	0.15	6.71e−02	1.39e+08	6.35e+07	1.68e+07
G359.611+0.018g	14.00	−0.38	−0.25	0.08	1.64e−02	6.51e+07	4.97e+07	1.53e+07
G359.611+0.018d	43.75	−0.39	−0.25	0.15	3.35e−02	7.06e+07	3.26e+07	1.44e+07
G359.611+0.018a	259.25	−0.39	−0.24	0.36	2.09e+00	1.73e+09	3.43e+08	9.88e+06
G359.611+0.018i	4.75	−0.38	−0.25	0.05	6.50e−03	6.35e+07	5.82e+07	1.17e+07
G359.611+0.018b	34.75	−0.38	−0.24	0.13	7.52e−02	1.70e+08	9.20e+07	1.25e+07
G359.611+0.018h	15.50	−0.39	−0.24	0.09	1.43e−02	5.91e+07	3.92e+07	1.30e+07
G359.611+0.018f	6.75	−0.40	−0.24	0.06	1.95e−02	1.48e+08	1.23e+08	3.13e+07
G359.611+0.018e	8.00	−0.40	−0.24	0.06	2.24e−02	1.40e+08	1.19e+08	3.42e+07
G0.316−0.201f	27.25	0.33	−0.21	0.12	4.52e−02	1.12e+08	7.06e+07	2.45e+07

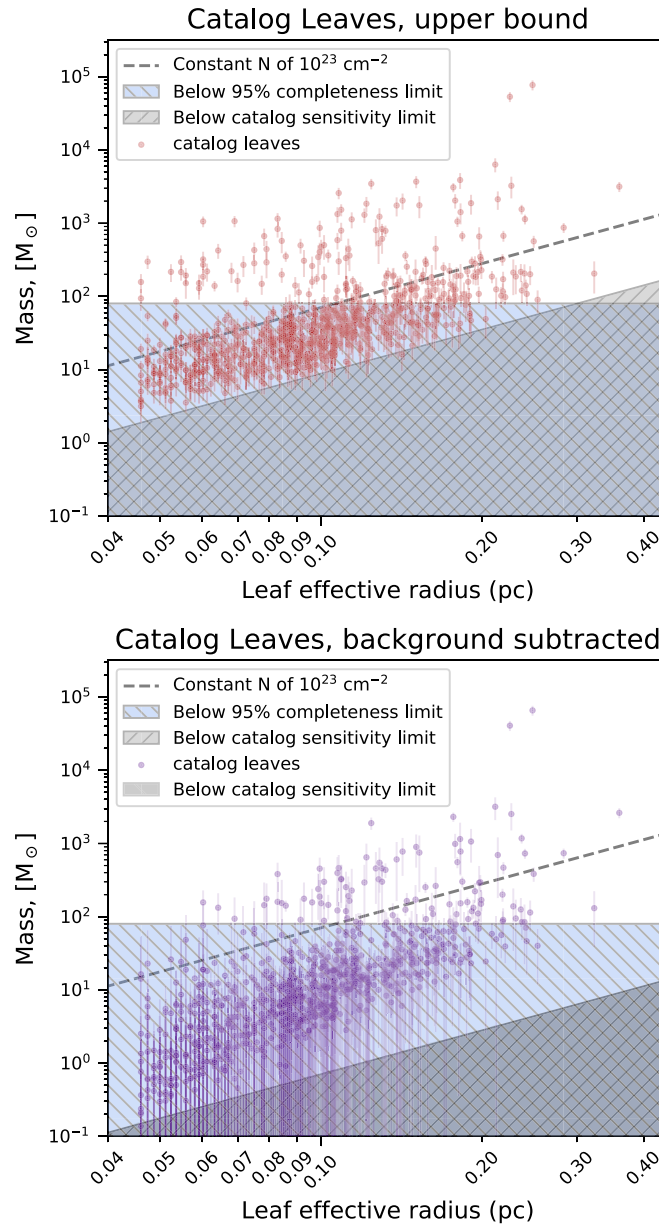


Figure 14. Two representations of the mass ($M_{\odot}$) vs. radius (parsecs) for each leaf in the high-completeness version of the catalog. The upper panel shows the mass calculated using the flux achieved by integrating over the entire leaf structure, acting as an upper bound for the true structure mass. The bottom panel displays the background-subtracted leaf masses, calculated by subtracting the minimum value of the leaf and integrated over the entire leaf area. This background-subtracted mass is likely an underestimate for the true structure mass. Both panels show a line of constant column density $n = 10^{23} \text{ cm}^{-2}$. The range of masses below 95% completeness is hatched right and shaded blue, and the region below minimum flux possible for a cataloged leaf is hatched left and shaded gray.

Table 5
Leaf Properties, High Completeness (Continued)

Leaf ID	N_{Herschel} (cm^{-2})	N_{SMA} (cm^{-2})	Mass ($M_{\odot}$)	Mass (bg. sub.) ($M_{\odot}$)	n (cm^{-3})	ρ (g cm^{-3})	t_{ff} (yr)
G359.611+0.018c	6.81e+22	1.42e+23	1.03e+02	4.16e+01	1.05e+05	4.92e-16	9.49e+04
G359.611+0.018g	6.81e+22	6.65e+22	2.51e+01	5.94e+00	1.48e+05	6.91e-16	8.01e+04
G359.611+0.018d	8.04e+22	6.36e+22	4.53e+01	2.36e+01	4.82e+04	2.26e-16	1.40e+05
G359.611+0.018a	1.46e+23	1.73e+24	3.15e+03	2.62e+03	2.33e+05	1.09e-15	6.37e+04
G359.611+0.018i	9.96e+22	6.20e+22	9.52e+00	8.45e-01	2.83e+05	1.33e-15	5.78e+04
G359.611+0.018b	8.59e+22	1.73e+23	1.14e+02	4.28e+01	1.72e+05	8.07e-16	7.41e+04
G359.611+0.018h	1.31e+23	5.57e+22	2.02e+01	7.52e+00	1.02e+05	4.79e-16	9.62e+04
G359.611+0.018f	5.33e+22	1.43e+23	2.83e+01	4.76e+00	4.98e+05	2.33e-15	4.36e+04
G359.611+0.018e	5.33e+22	1.36e+23	3.25e+01	4.31e+00	4.42e+05	2.07e-15	4.63e+04
G0.316-0.201f	3.67e+22	1.16e+23	7.00e+01	2.63e+01	1.52e+05	7.10e-16	7.90e+04

Table 6
Leaf Property Uncertainties, High Completeness

Leaf ID	Mass Unc. ($M_{\odot}$)	N_{SMA} Unc. (cm^{-2})	n Unc. (cm^{-3})	ρ Unc. ($\text{g cm}^{-3} \text{ yr}$)	t_{ff} Unc.
G359.611+0.018	3.15e+01	2.85e+23	4.47e+04	2.09e-16	4.04e+04
G359.611+0.018g	8.62e+00	1.34e+23	6.69e+04	3.14e-16	3.63e+04
G359.611+0.018d	2.11e+01	1.28e+23	3.04e+04	1.42e-16	8.82e+04
G359.611+0.018a	4.92e+02	3.48e+24	7.85e+04	3.68e-16	2.15e+04
G359.611+0.018i	2.40e+00	1.25e+23	1.15e+05	5.39e-16	2.35e+04
G359.611+0.018b	2.34e+01	3.47e+23	6.20e+04	2.90e-16	2.67e+04
G359.611+0.018h	7.37e+00	1.12e+23	5.21e+04	2.44e-16	4.90e+04
G359.611+0.018f	8.40e+00	2.88e+23	2.20e+05	1.03e-15	1.93e+04
G359.611+0.018e	1.06e+01	2.74e+23	2.06e+05	9.63e-16	2.15e+04
G0.316-0.201f	2.66e+01	2.34e+23	7.25e+04	3.40e-16	3.78e+04

Appendix B

Catalog Parameter Studies

The output tree of a dendrogram is highly dependent on the choice of parameters given to the dendrogram algorithm. Given that our cataloging algorithm is a collection of dendrogram leaves pruned according to local noise, it is necessary to understand how our catalog changes with the choices of both the parameters of the dendrogram, and also our method of estimating local noise. In this section, we perform simple parameter studies to explore the variability of the final catalog with small changes in the initial dendrogram parameters as well as the smoothing kernel for generating our rms noise maps.

Small changes in the minimum value parameter for the original dendrogram (see Section 3) have no effect on the total dendrogram, as all of the rms noise values are by design greater than the low noise estimate used to scale the initial dendrogram parameters. Therefore we neglect an experiment varying this quantity. The initial dendrogram, however, is sensitive to changes in the minimum significance parameter δ . Figure 15 shows three versions of the dendrogram leaves for three different values of δ , with key properties remaining largely similar. It is possible, by changing δ , to cause nearby leaves of similar flux distributions to merge together or split apart into separate leaves. This phenomenon is rare in the case of our catalogs, with $\lesssim 3\%$ of leaves contained in the robust catalog displaying this behavior. In Figure 16 we show the relationship between cataloged leaf masses before and after a leaf merger. The merged leaf mass and radius are larger than would be naively expected by summing the radii and masses of the

separate leaves, though this is not surprising given the clustering procedure at work in the dendrogram algorithm.

Similarly, leaves with particularly shallow distributions might change slightly in size with changes in δ . These changes can affect the inferred physical properties of the leaves in a nontrivial way. Figure 17 shows a selection of leaves identified in the robust catalog that vary with changes in δ . The leaf size and inferred mass appear to vary as a roughly consistent power law, with leaves varying in mass (radius) less than a factor of ~ 2 (~ 3). Even with these rare cases where leaves merge or change shape with varying δ , we find that the mass and radius distribution of leaf properties in the catalog is well preserved (see Figure 15), with the leaves associated with Sgr B2 and those elsewhere still clearly separated.

We also consider the effect of changes to the local noise map generation on the catalog. The only scalable parameter affecting the contents of the catalog is the convolution kernel size used in smoothing the rms noise map. We explore the sensitivity of the catalog to changes in the smoothing kernel size using a similar simulated observation method as Section 3.2. For each iteration of the simulated observations, and for a given kernel size, we record the false-positive percentage for point-source-like emission. In order to get a single number for the false-positive rate over a number of trials, we average together all simulated observation runs for source masses above the 95% completeness for that run. It is worth noting that the number of false positives is higher for larger-mass point sources, as the imaging artifacts have correspondingly higher fluxes and are less likely to be pruned away by the cataloging algorithm. The results from this

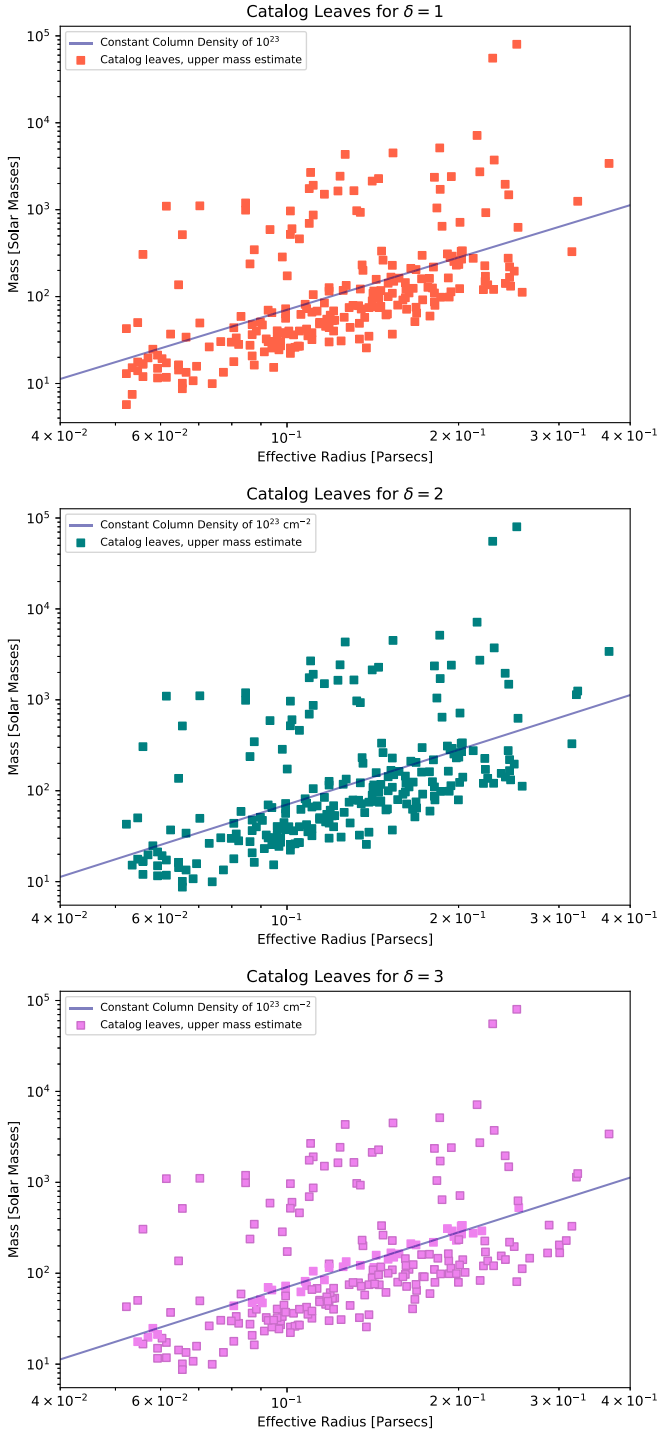


Figure 15. Three versions of the catalog, calculated for different choices of the initial dendrogram parameter for the minimum separation of structures, δ . From top to bottom, these catalogs were constructed using $\delta = 2$, $\delta = 3$, and $\delta = 4$. The middle panel is identical to the actual catalog described in the rest of this work. The majority of leaves are robust with respect to changes in δ , and the objects most affected seem to be those with a large effective radius and a small mass.

process are displayed in Figure 18. It appears that a kernel size of 14 pixels or 18 pixels minimizes the average number of false positives for objects above our 80% completeness limit. We chose to use the 14 pixel kernel as this choice also has a slightly lower average completeness mass. The variations

Variation of Leaf Mass, Effective Radius for a Merged Leaf

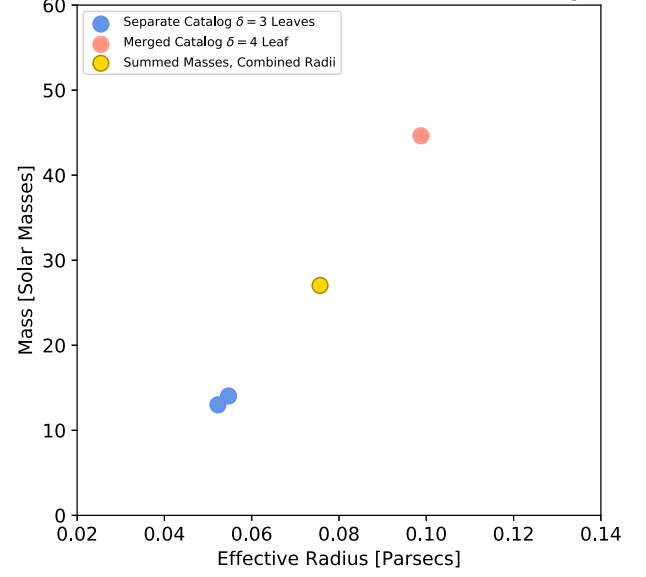


Figure 16. The behavior of a pair of nearby local flux peaks included as separate leaves in the catalog for a low $\delta = 1, 23$, but that merge together for $\delta = 3$. This behavior is uncommon in the catalog but has a significant effect on the total amount of flux included in the structure. The amount of flux enclosed by the merged structure is more than a factor of two larger than the sum of the individual leaves' mass.

Variation of Leaf Mass, Effective Radius for changes in δ

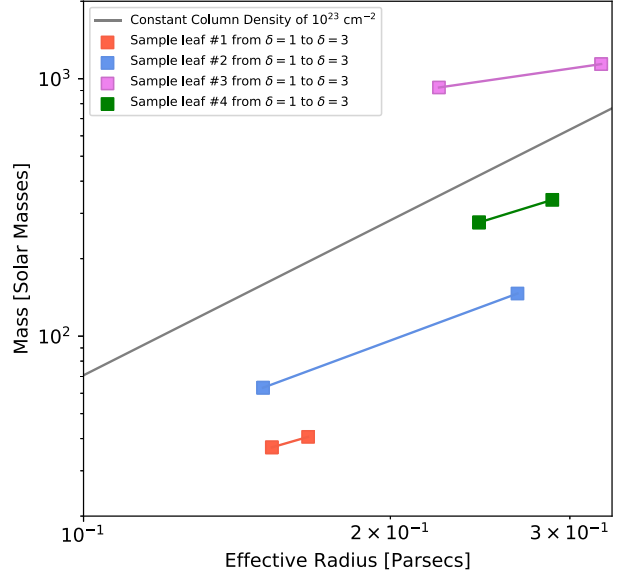


Figure 17. The variation of four leaves that change mass and effective radius for varying the initial dendrogram minimum difference parameter δ . A line of constant column density 10^{23} is shown for comparison to Figure 6.

from trial to trial of this experiment are large enough that there is no ideal choice of kernel for any given configuration of point sources, and we suspect the results would be further complicated by non-point-source-like emission. However the variation in completeness is typically smaller for choices in the range of 10–20 pixels, and resulting effects on the catalog varying between these choices are small relative to the overall trends in completeness in Figure 2.

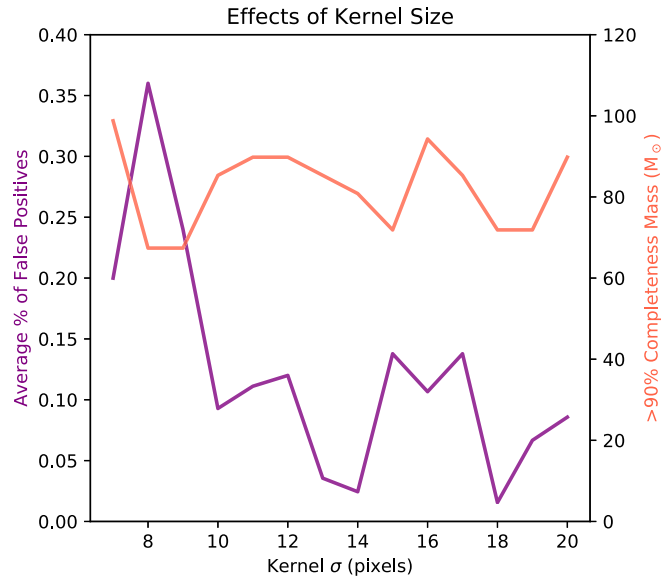


Figure 18. The effects of altering the smoothing kernel used in the convolutions for generating the local noise map used for pruning. Certain choices of kernel size lead to a much higher percentage of false-positive source detections for point objects above the completeness threshold of $70 M_{\odot}$ used in the high-reliability version of the catalog (see Section 3.2).

Appendix C Inset Images of Key Regions

In this section, we show a gallery of the regions included in the CMZoom survey, with the high-reliability catalog leaf contours overplotted in white. These cutouts are presented in Figures 19–24.

The SMA 1.3 mm continuum maps, as well as the leaf contour mask used to generate this gallery, have been made available at <https://dataverse.harvard.edu/dataverse/cmzoom>. All images are displayed with a 1 pc scale bar and are on a log scale with limits between 1×10^7 Jy sr $^{-1}$ and 3×10^8 Jy sr $^{-1}$.

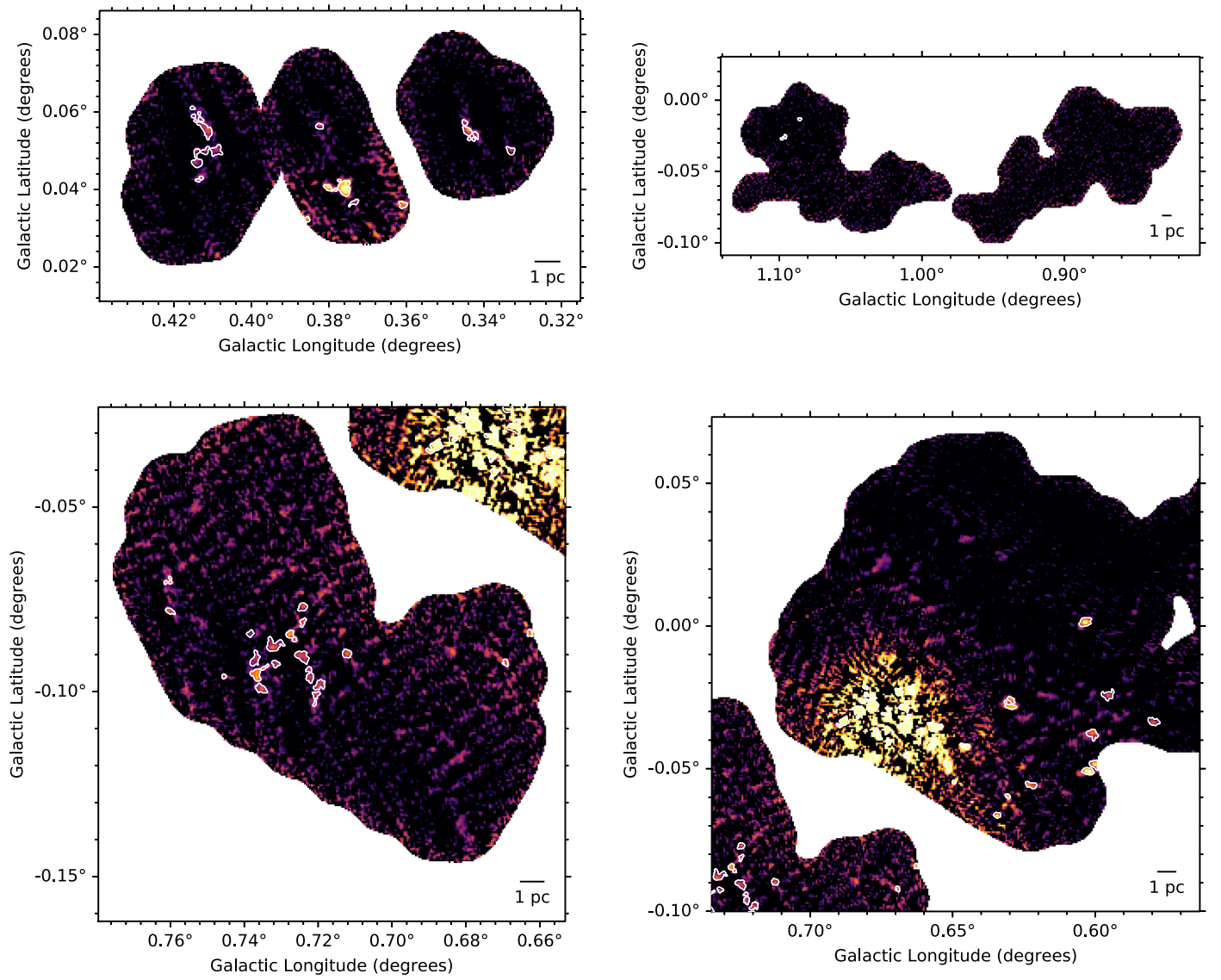


Figure 19. Robust catalog leaf contours superimposed over the 1.3 mm dust continuum from the CMZoom survey with a 1 pc scale bar. All images are displayed on a log scale with limits between $1 \times 10^7 \text{ Jy sr}^{-1}$ and $3 \times 10^8 \text{ Jy sr}^{-1}$.

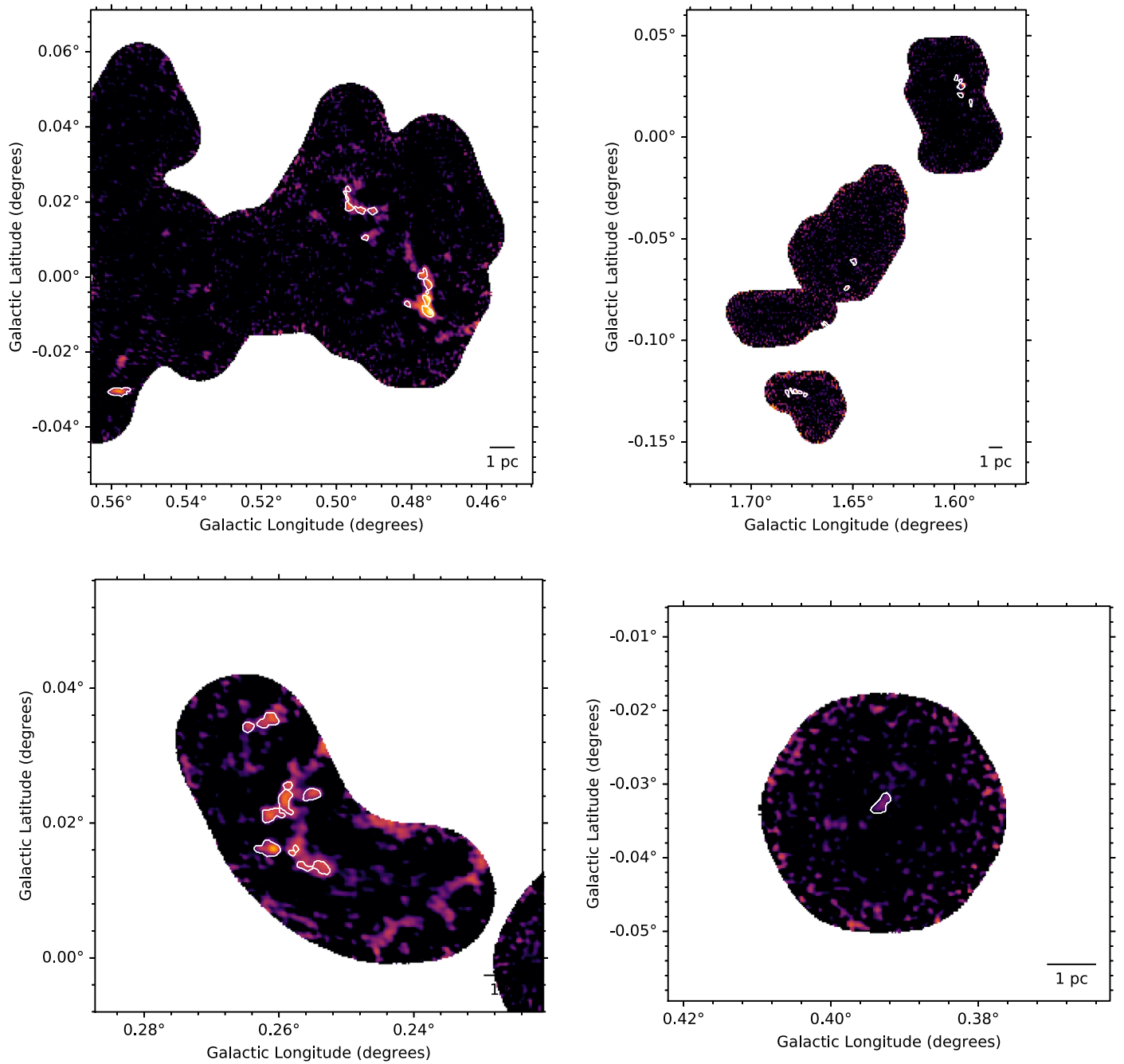


Figure 20. Robust catalog leaf contours superimposed over the 1.3 mm dust continuum from the CMZoom survey with a 1 pc scale bar. All images are displayed on a log scale with limits between $1 \times 10^7 \text{ Jy sr}^{-1}$ and $3 \times 10^8 \text{ Jy sr}^{-1}$.

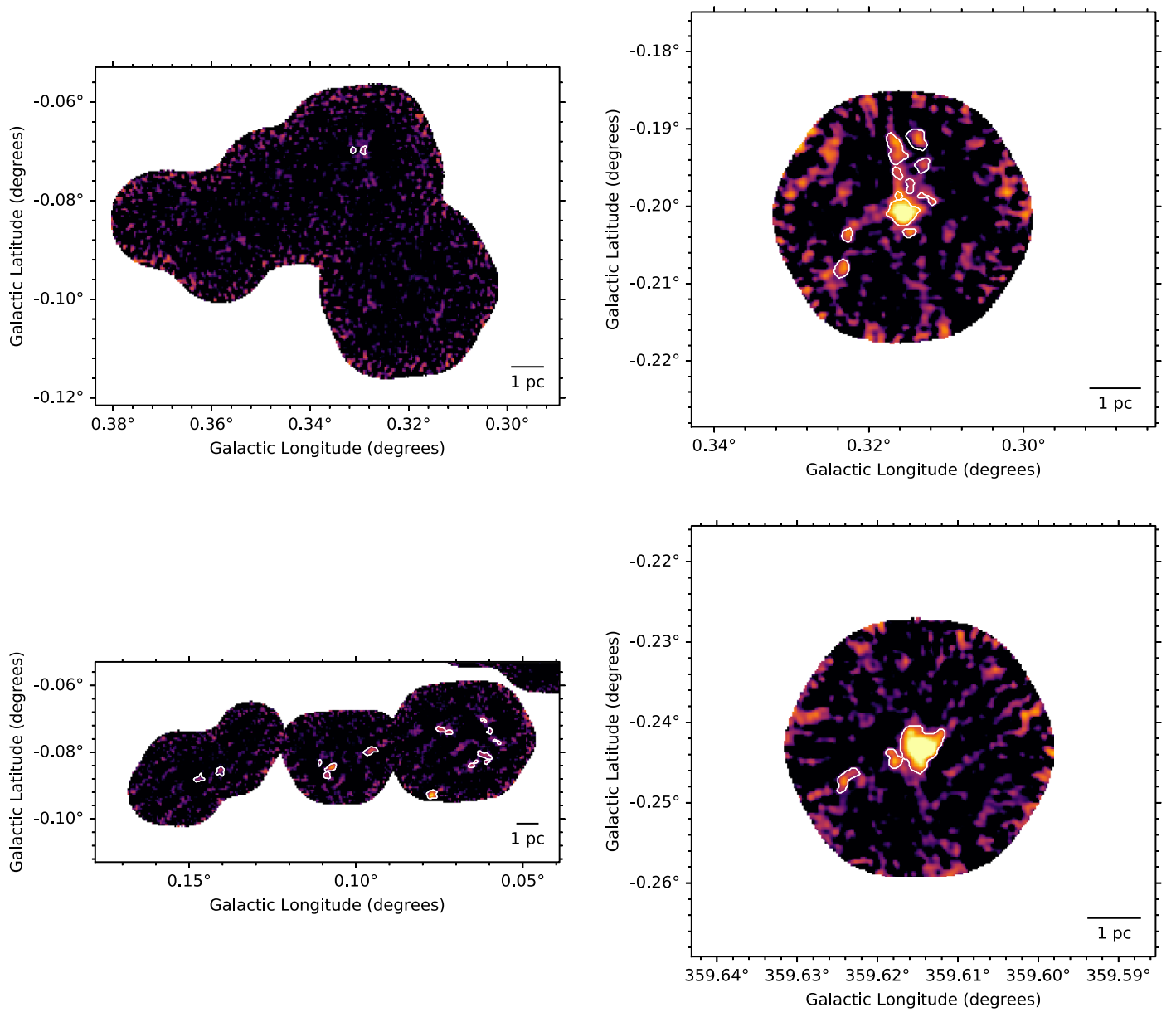


Figure 21. Robust catalog leaf contours superimposed over the 1.3 mm dust continuum from the CMZoom survey with a 1 pc scale bar. All images are displayed on a log scale with limits between $1 \times 10^7 \text{ Jy sr}^{-1}$ and 3×10^8 .

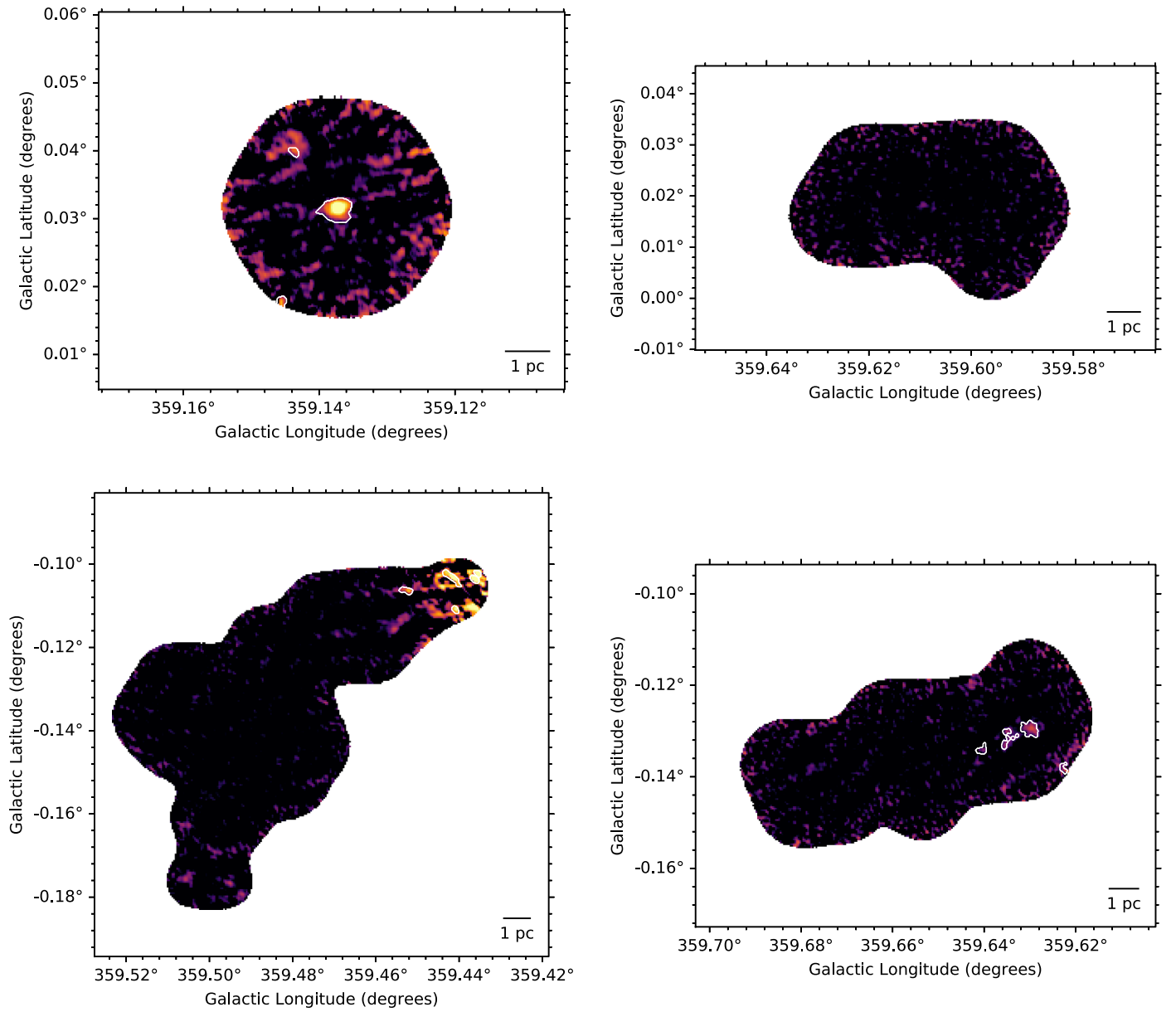


Figure 22. Robust catalog leaf contours superimposed over the 1.3 mm dust continuum from the CMZoom survey with a 1 pc scale bar. All images are displayed on a log scale with limits between $1 \times 10^7 \text{ Jy sr}^{-1}$ and $3 \times 10^8 \text{ Jy sr}^{-1}$.

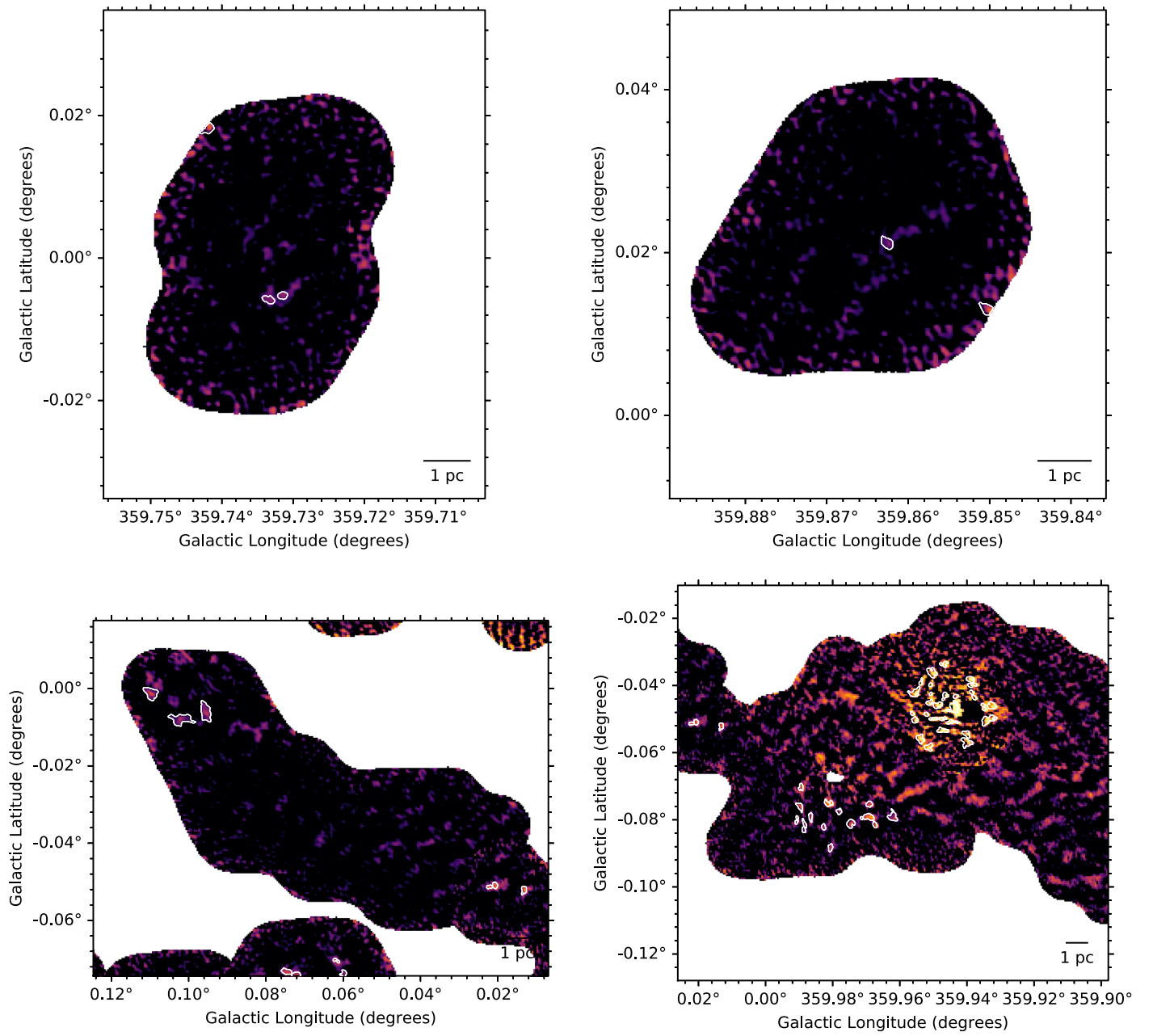


Figure 23. Robust catalog leaf contours superimposed over the 1.3 mm dust continuum from the CMZoom survey with a 1 pc scale bar. All images are displayed on a log scale with limits between $1 \times 10^7 \text{ Jy sr}^{-1}$ and $3 \times 10^8 \text{ Jy sr}^{-1}$.

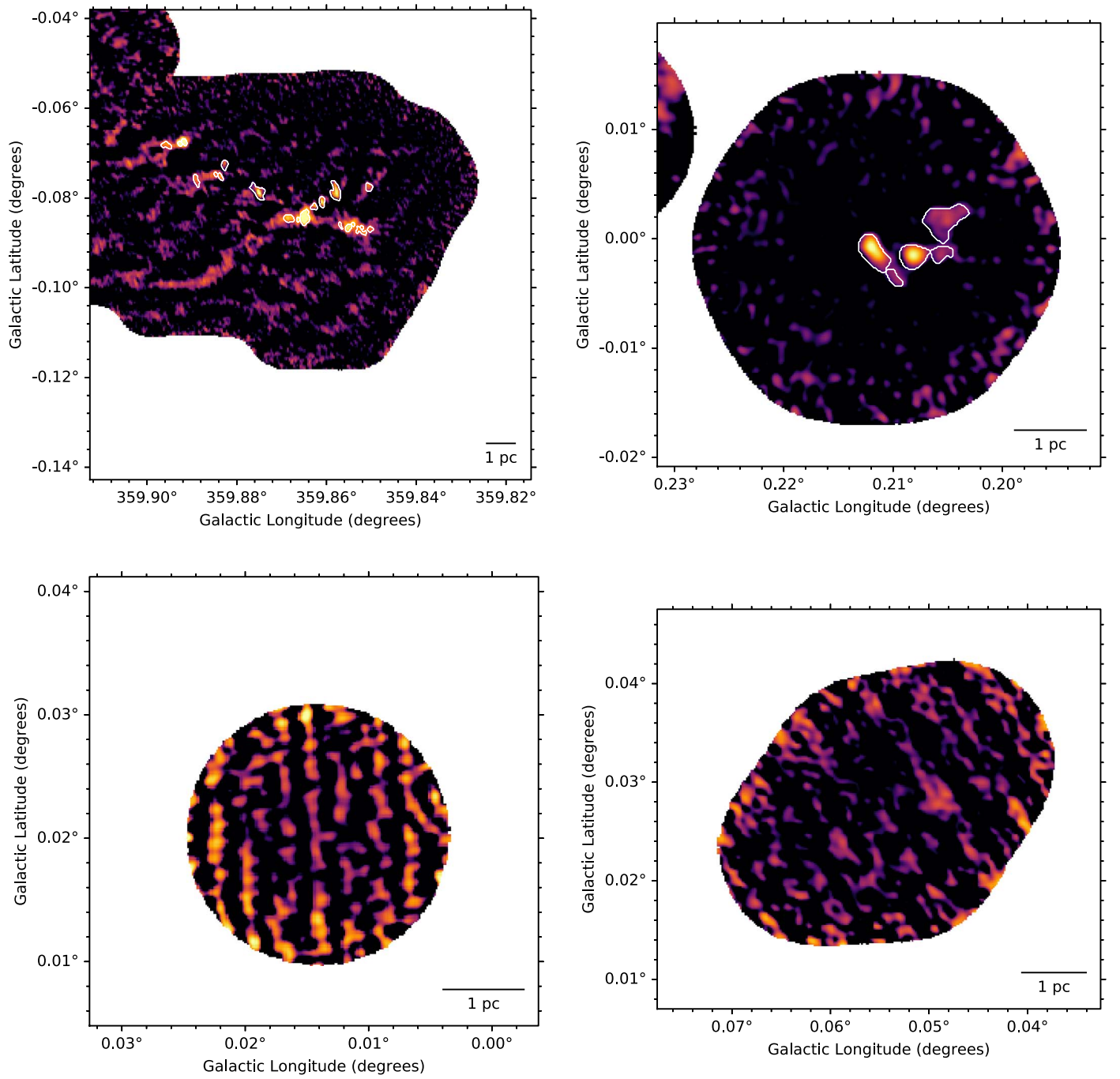


Figure 24. Robust catalog leaf contours superimposed over the 1.3 mm dust continuum from the CMZoom survey with a 1 pc scale bar. All images are displayed on a log scale with limits between $1 \times 10^7 \text{ Jy sr}^{-1}$ and $3 \times 10^8 \text{ Jy sr}^{-1}$.

Appendix D

Comparison with Sgr B2 ALMA Data

In order to further probe the trustworthiness of the catalog contents, the leaves in several key regions of our SMA continuum maps were compared with results from recent observations from ALMA of the same regions. In particular, we

used 1.1 mm continuum maps from Dust Ridge clouds D, E, and F (Barnes et al. 2019), as well as the Sgr B2 complex at 3 mm (Ginsburg et al. 2018). These maps from ALMA are compared side by side in Figures 25–27 below. Dust Ridge clouds D, E, and F appear by eye to have a very good agreement, with all catalog leaves having a clear counterpart in

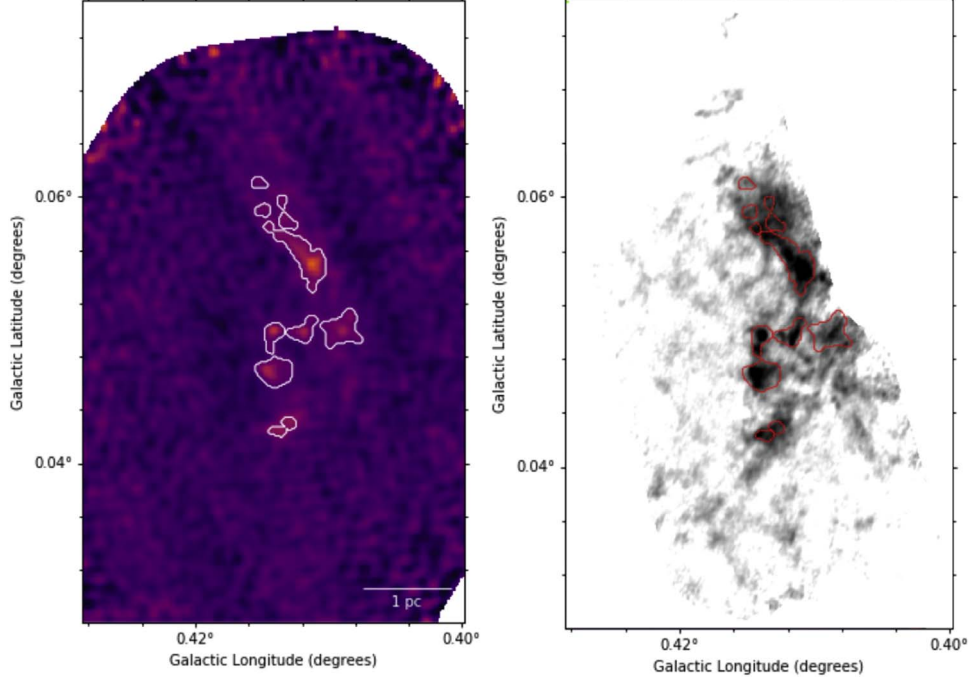


Figure 25. A side-by-side comparison of dust ridge cloud D in the ALMA 1.1 mm continuum combined with Bolocam single dish map (right) from Barnes et al. (2019), and the 1.3 mm SMA continuum maps from the CMZoom survey (left), both with contours of the catalog leaves in red. All of the catalog leaves have corresponding sources in the ALMA map. The SMA map is scaled between 10^7 and 3×10^8 Jy sr $^{-1}$. The ALMA map is scaled between -0.0026 and 1.1502 Jy beam $^{-1}$.

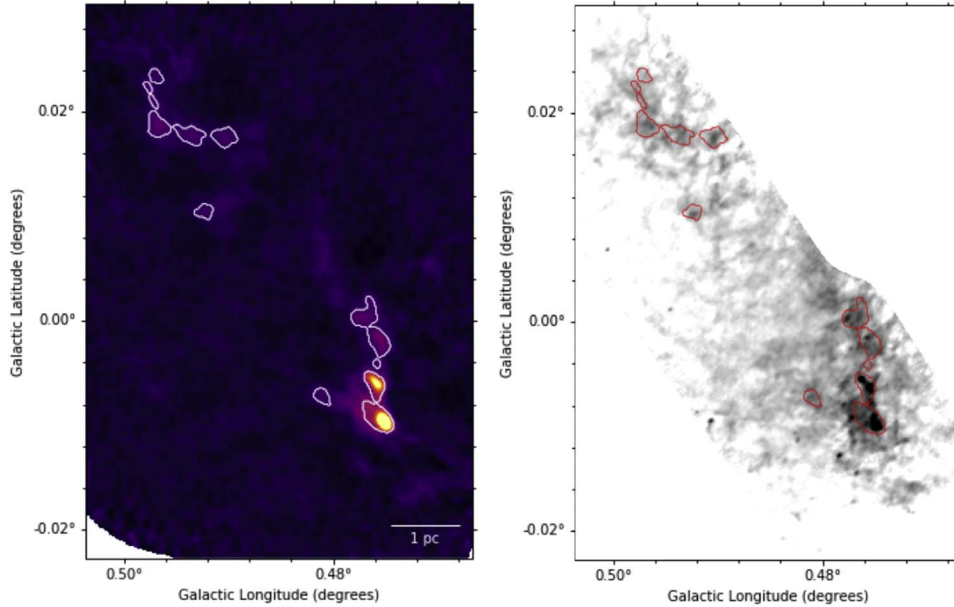


Figure 26. A side-by-side comparison of Dust Ridge Clouds E and F in the ALMA 1.1 mm continuum combined with Bolocam single dish map from Barnes et al. (2019; right), and the 1.3 mm SMA continuum maps from the CMZoom survey (left), both with contours of the catalog leaves in red and white, respectively. All of the catalog leaves have corresponding sources in the ALMA map. The 1.3 mm SMA map seems to more clearly resemble the filamentary structure of the emission of these clouds, which fragmented into more discrete substructure on the scales revealed by ALMA. The SMA map is scaled between 10^7 and 1.6×10^8 Jy sr $^{-1}$. The ALMA map is scaled between -0.0026 and 1.1502 Jy beam $^{-1}$.

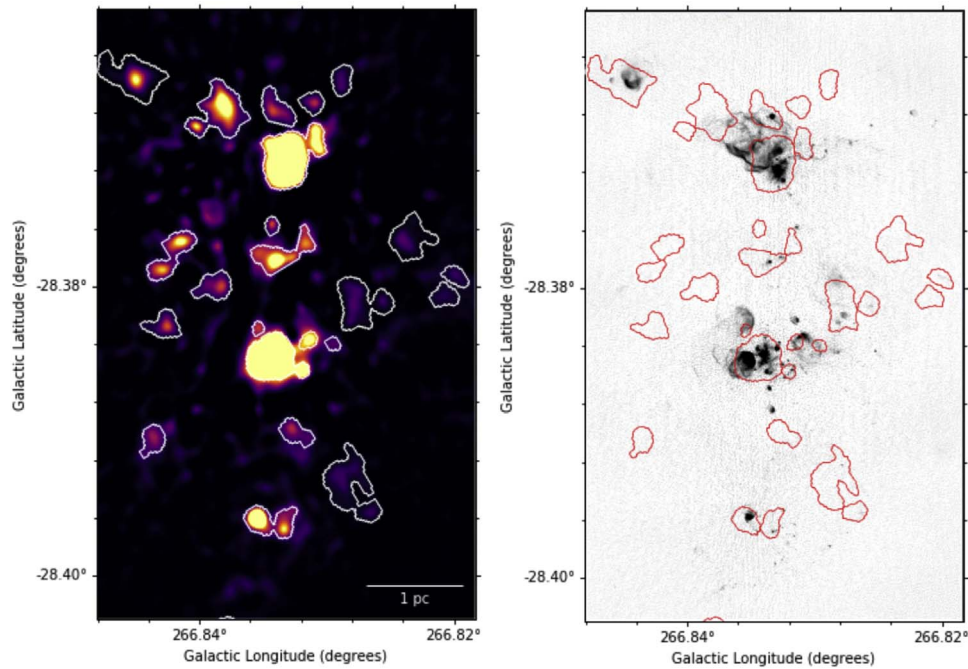


Figure 27. A side-by-side comparison of the Sgr B2 complex in the 1.3 mm SMA continuum map (left) from the CMZoom survey and the ALMA 3 mm continuum map (right) from Ginsburg et al. (2018), both with contours of the catalog leaves in white and red, respectively. This figure highlights several catalog leaves that do not appear to have corresponding emission in the ALMA map. We suspect that these sources and the ones that appear symmetrically on the other side of the SgrB2 complex might not be real emission, but instead imaging artifacts. Some of the extended emission visible in the ALMA 3 mm map is from optically thin H II regions, which would not be significant at 1.3 mm. The SMA map is scaled between 10^7 and 3×10^8 Jy sr $^{-1}$. The ALMA map is scaled between -0.0026 and 1.1502 Jy beam $^{-1}$.

the ALMA maps. Some of the ALMA sources do not appear to have been significant enough detections for inclusion in the catalog, but this is reasonable considering the higher sensitivity given by ALMA. The SMA Sgr B2 map, however, contains some compact sources that do not appear to be present in the ALMA map. While it is possible that this might be due to the fact that this region contains the highest noise and highest-intensity emission contained in the CMZoom survey coverage, it is also likely that differences in emission between the 1.3 mm and 3 mm maps, particularly given the widespread compact and extended H II regions known to populate the area, could also account for the discrepancy. We also note that the leaves, highlighted in Figure 27, lie next to a deep negative bowl in the SMA map, supporting the explanation that these leaves in particular are an artifact of the imaging process. These SMA artifact leaves do not appear to have properties remarkably different from nearby leaves, which are in agreement with the ALMA maps, and therefore do not significantly influence our interpretation of Sgr B2 leaves in Section 5.1.

ORCID iDs

H Perry Hatchfield <https://orcid.org/0000-0003-0946-4365>
 Cara Battersby <https://orcid.org/0000-0002-6073-9320>
 Ashley Barnes <https://orcid.org/0000-0003-0410-4504>
 Adam Ginsburg <https://orcid.org/0000-0001-6431-9633>
 J. M. Diederik Kruijssen <https://orcid.org/0000-0002-8804-0212>
 Steve N. Longmore <https://orcid.org/0000-0001-6353-0170>
 Xing Lu <https://orcid.org/0000-0003-2619-9305>
 Elisabeth A. C. Mills <https://orcid.org/0000-0001-8782-1992>
 Thushara Pillai <https://orcid.org/0000-0003-2133-4862>
 Qizhou Zhang <https://orcid.org/0000-0003-2384-6589>

John Bally <https://orcid.org/0000-0001-8135-6612>
 Natalie Butterfield <https://orcid.org/0000-0002-4013-6469>
 Yanett A. Contreras <https://orcid.org/0000-0002-6388-3635>
 Luis C. Ho <https://orcid.org/0000-0001-6947-5846>
 Jürgen Ott <https://orcid.org/0000-0001-8224-1956>
 Nimesh Patel <https://orcid.org/0000-0002-6021-9421>
 Volker Tolls <https://orcid.org/0000-0003-1841-2241>

References

- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Barnes, A. T., Longmore, S. N., Avison, A., et al. 2019, *MNRAS*, **486**, 283
- Barnes, A. T., Longmore, S. N., Battersby, C., et al. 2017, *MNRAS*, **469**, 2263
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, *ARA&A*, **48**, 339
- Battersby, C., Bally, J., Ginsburg, A., et al. 2011, *A&A*, **535**, A128
- Battersby, C., Bally, J., Jackson, J. M., et al. 2010, *ApJ*, **721**, 222
- Battersby, C., Keto, E., Walker, D., et al. 2020, *ApJS*, **249**, 35
- Beaumont, C., Goodman, A., & Greenfield, P. 2015, in ASP Conf. Ser. 495, *Astronomical Data Analysis Software and Systems XXIV (ADASS XXIV)*, ed. A. R. Taylor & E. Rosolowsky (San Francisco, CA: ASP), 101
- Bergin, E. A., & Tafalla, M. 2007, *ARA&A*, **45**, 339
- Chomiuk, L., & Povich, M. S. 2011, *ApJ*, **142**, 197
- Dahmen, G., Huttemeister, S., Wilson, T. L., & Mauersberger, R. 1998, *A&A*, **331**, 959
- Dale, J. E., Kruijssen, J. M. D., & Longmore, S. N. 2019, *MNRAS*, **486**, 3307
- De Pree, C. G., Gaume, R. A., Goss, W. M., & Claussen, M. J. 1996, *ApJ*, **464**, 788
- De Pree, C. G., Peters, T., Mac Low, M. M., et al. 2015, *ApJ*, **815**, 123
- Downes, D., & Maxwell, A. 1966, *ApJ*, **146**, 653
- Etzaluze, M., Goicoechea, J. R., Cernicharo, J., et al. 2013, *A&A*, **556**, A137
- Federrath, C., Rathborne, J. M., Longmore, S. N., et al. 2016, *ApJ*, **832**, 143
- Figer, D. F., Najarro, F., Gilmore, D., et al. 2002, *ApJ*, **581**, 258
- Ginsburg, A., Bally, J., Barnes, A., et al. 2018, *ApJ*, **853**, 171
- Ginsburg, A., Henkel, C., Ao, Y., et al. 2016, *A&A*, **586**, A50
- Goldsmith, P. F., Lis, D. C., Hills, R., & Lasenby, J. 1990, *ApJ*, **350**, 186

- Gordon, M. A., Berkemann, U., Mezger, P. G., et al. 1993, *A&A*, **280**, 208
- Goto, M., Indriolo, N., Geballe, T. R., & Usuda, T. 2013, *JPCA*, **117**, 9919
- Guesten, R., & Downes, D. 1982, *A&A*, **117**, 343
- Habibi, M., Stolte, A., Brandner, W., Hußmann, B., & Motohara, K. 2013, *A&A*, **556**, A26
- Harada, N., Riquelme, D., Viti, S., et al. 2015, *A&A*, **584**, A102
- Henshaw, J. D., Caselli, P., Fontani, F., et al. 2016a, *MNRAS*, **463**, 146
- Henshaw, J. D., Ginsburg, A., Haworth, T. J., et al. 2019, *MNRAS*, **485**, 2457
- Henshaw, J. D., Jiménez-Serra, I., Longmore, S. N., et al. 2017, *MNRAS*, **464**, L31
- Henshaw, J. D., Longmore, S. N., Kruijssen, J. M. D., et al. 2016b, *MNRAS*, **457**, 2675
- Hosek, M. W., Jr., Lu, J. R., Anderson, J., et al. 2019, *ApJ*, **870**, 44
- Immer, K., Schuller, F., Omont, A., & Menten, K. M. 2012, *A&A*, **537**, A121
- Jeffreson, S. M. R., Kruijssen, J. M. D., Krumholz, M. R., & Longmore, S. N. 2018, *MNRAS*, **478**, 3380
- Kauffmann, J., Bertoldi, F., Bourke, T. L., Evans, N. J., II, & Lee, C. W. 2008, *A&A*, **487**, 993
- Kauffmann, J., & Pillai, T. 2010, *ApJL*, **723**, L7
- Kauffmann, J., Pillai, T., & Goldsmith, P. F. 2013a, *ApJ*, **779**, 185
- Kauffmann, J., Pillai, T., & Zhang, Q. 2013b, *ApJL*, **765**, L35
- Kauffmann, J., Pillai, T., Zhang, Q., et al. 2017a, *A&A*, **603**, A89
- Kauffmann, J., Pillai, T., Zhang, Q., et al. 2017b, *A&A*, **603**, A90
- Kennicutt, R. C., Jr. & Evans, N. J., II 2012, *ARA&A*, **50**, 531
- Koepferl, C. M., Robitaille, T. P., Morales, E. F. E., & Johnston, K. G. 2015, *ApJ*, **799**, 53
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Kruijssen, J. M. D., Dale, J. E., & Longmore, S. N. 2015, *MNRAS*, **447**, 1059
- Kruijssen, J. M. D., Dale, J. E., Longmore, S. N., et al. 2019, *MNRAS*, **484**, 5734
- Kruijssen, J. M. D., & Longmore, S. N. 2013, *MNRAS*, **435**, 2598
- Kruijssen, J. M. D., Longmore, S. N., Elmegreen, B. G., et al. 2014, *MNRAS*, **440**, 3370
- Krumholz, M. R., Fumagalli, M., da Silva, R. L., Rendahl, T., & Parra, J. 2015, *MNRAS*, **452**, 1447
- Krumholz, M. R., & Kruijssen, J. M. D. 2015, *MNRAS*, **453**, 739
- Krumholz, M. R., Kruijssen, J. M. D., & Crocker, R. M. 2017, *MNRAS*, **466**, 1213
- Lada, C. J., Forbrich, J., Lombardi, M., & Alves, J. F. 2012, *ApJ*, **745**, 190
- Lada, C. J., & Lada, E. A. 2003, *ARA&A*, **41**, 57
- Lada, C. J., Lombardi, M., & Alves, J. F. 2010, *ApJ*, **724**, 687
- Le Petit, F., Ruaud, M., Bron, E., et al. 2016, *A&A*, **585**, A105
- Lis, D. C., Menten, K. M., Serabyn, E., & Zylka, R. 1994, *ApJL*, **423**, L39
- Lis, D. C., Serabyn, E., Zylka, R., & Li, Y. 2001, *ApJ*, **550**, 761
- Liu, H. B., Dunham, M. M., Pascucci, I., et al. 2018, *A&A*, **612**, A54
- Lo, K. Y., & Claussen, M. J. 1983, *Natur*, **306**, 647
- Longmore, S. N., Bally, J., Testi, L., et al. 2013a, *MNRAS*, **429**, 987
- Longmore, S. N., Kruijssen, J. M. D., Bally, J., et al. 2013b, *MNRAS*, **433**, L15
- Longmore, S. N., Rathborne, J., Bastian, N., et al. 2012, *ApJ*, **746**, 117
- Lu, J. R., Do, T., Ghez, A. M., et al. 2013, *ApJ*, **764**, 155
- Lu, X., Mills, E. A. C., Ginsburg, A., et al. 2019a, *ApJS*, **244**, 35
- Lu, X., Zhang, Q., Kauffmann, J., et al. 2015, *ApJL*, **814**, L18
- Lu, X., Zhang, Q., Kauffmann, J., et al. 2019b, *ApJ*, **872**, 171
- Marsh, K. A., Whitworth, A. P., Lomax, O., et al. 2017, *MNRAS*, **471**, 2730
- McKee, C. F., & Ostriker, E. C. 2007, *ARA&A*, **45**, 565
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in ASP Conf. Series 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw et al. (San Francisco, CA: ASP), 127
- Mehringer, D. M., de Pree, C. G., Gaume, R. A., Goss, W. M., & Claussen, M. J. 1995, *ApJL*, **442**, L29
- Mills, E. A. C., & Battersby, C. 2017, *ApJ*, **835**, 76
- Mills, E. A. C., Butterfield, N., Ludovici, D. A., et al. 2015, *ApJ*, **805**, 72
- Mills, E. A. C., & Morris, M. R. 2013, *ApJ*, **772**, 105
- Morris, M., & Serabyn, E. 1996, *ARA&A*, **34**, 645
- Offner, S. S. R., Clark, P. C., Hennebelle, P., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 53
- Oka, T., Geballe, T. R., Goto, M., et al. 2019, *ApJ*, **883**, 54
- Ossenkopf, V., & Henning, T. 1994, *A&A*, **291**, 943
- Pardo, J. R., Cernicharo, J., & Serabyn, E. 2001, *ITAP*, **49**, 1683
- Pierce-Price, D., Richer, J. S., Greaves, J. S., et al. 2000, *ApJL*, **545**, L121
- Pillai, T., Kauffmann, J., Tan, J. C., et al. 2015, *ApJ*, **799**, 74
- Rathborne, J. M., Jackson, J. M., & Simon, R. 2006, *ApJ*, **641**, 389
- Rathborne, J. M., Longmore, S. N., Jackson, J. M., et al. 2014, *ApJ*, **786**, 140
- Robitaille, T. 2019, APLpy v2.0: The Astronomical Plotting Library in Python, Zenodo, doi:10.5281/zenodo.2567476
- Robitaille, T., Beaumont, C., Qian, P., Borkin, M., & Goodman, A. 2017, glueviz v0.13.1: multidimensional data exploration, Zenodo, doi:10.5281/zenodo.1237692
- Robitaille, T., & Bressert, E. 2012, APLpy: Astronomical Plotting Library in Python, Astrophysics Source Code Library, ascl:1208.017
- Rosolowsky, E. W., Pineda, J. E., Kauffmann, J., & Goodman, A. A. 2008, *ApJ*, **679**, 1338
- Sault, R. J., Teuben, P. J., & Wright, M. C. H. 1995, in ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco, CA: ASP), 433
- Schmiedeke, A., Schilke, P., Möller, T., et al. 2016, *A&A*, **588**, A143
- Serabyn, E., Carlstrom, J., Lay, O., et al. 1997, *ApJL*, **490**, L77
- Sormani, M. C., & Li, Z. 2020, *MNRAS*, **494**, 6030
- Sormani, M. C., Sobacchi, E., Fragkoudi, F., et al. 2018, *MNRAS*, **481**, 2
- Sormani, M. C., Tress, R. G., Glover, S. C. O., et al. 2020, *MNRAS*, **497**, 5024
- The GRAVITY Collaboration, Abuter, R., & Amorim, A. 2019, *A&A*, **625**, L10
- Tress, R. G., Sormani, M. C., Glover, S. C. O., et al. 2020, *MNRAS*, in press (arXiv:2004.06724)
- Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Walker, D. L., Longmore, S. N., Bastian, N., et al. 2016, *MNRAS*, **457**, 4536
- Walker, D. L., Longmore, S. N., Zhang, Q., et al. 2018, *MNRAS*, **474**, 2373
- Yusef-Zadeh, F., Hewitt, J. W., Arendt, R. G., et al. 2009, *ApJ*, **702**, 178
- Yusef-Zadeh, F., & Wardle, M. 2008, in ASP Conf. Series 387, Massive Star Formation: Observations Confront Theory, ed. H. Beuther et al. (San Francisco, CA: ASP), 361