

Prestellar cores in Sagittarius B2

NAZAR BUDAIEV  ¹ AND ADAM GINSBURG  ¹

¹*Department of Astronomy, University of Florida, Gainesville, FL 32611*

ABSTRACT

We present ~ 500 AU resolution ALMA observations of pre- and proto-stellar cores in the giant molecular cloud Sagittarius B2 (Sgr B2) in 1 mm and 3 mm. Previous 3 mm observations of the covered area detected ~ 150 sources down to 8 solar masses. With a tenfold increase in resolution, we detect 314 sources at 3 mm and 163 sources in the smaller field of view at 1 mm. Some of the sources seen at low resolution are observed to fragment into multiple protostars, while a few are undetected at high resolution as they are resolved out. We find that most of the sources are resolved and are at least partially optically thick. We report the lower limit on the source masses.

Keywords: Star formation, prestellar cores, protostellar cores, molecular cloud, Sagittarius B2

1. INTRODUCTION

The high-density, turbulent, and overall extreme environment of the Central Molecular Zone (CMZ) can be used to probe disk scales in conditions similar to the star forming environment at cosmic noon ($z \sim 2$ Madau & Dickinson 2014) (Henshaw et al. 2016). In young star clusters such as the Arches and Quintuplet clusters, there is evidence of a top-heavy IMF (Hosek et al. 2019). Peering into early stages of star formation in the CMZ can help us understand if the IMF varies systematically with local environmental conditions and thus is shallower in the early universe.

The densest star forming cloud in our galaxy, Sagittarius B2 (Sgr B2), is a powerful laboratory for studying star formation and evolution in highest-density region where we can resolve disk scales. The region is split into several smaller, concentrated regions of star formation: North (N), Main (M), South (S), and Deep South (DS). A large fraction ($\sim 25 - 50\%$)

of the current star formation in this cloud occurs in the N and M regions, with many stars forming in bound clusters (Ginsburg et al. 2018; Barnes et al. 2017). Half of the total star formation in the CMZ occurs in Sgr B2, despite containing $\sim 10\%$ of the mass and $< 1\%$ of the volume (Ginsburg et al. 2018; Barnes et al. 2017). Sgr B2 appears to be forming stars preferentially in dense, bound clusters (Ginsburg et al. 2018; Ginsburg & Kruijssen 2018). Since other Galactic Center star clusters have excesses of high-mass stars (Hosek et al. 2019), Sgr B2 is likely also forming extra massive stars. If this excess is present, it indicates a systematic variation of the stellar IMF with gas conditions, with a shallowing IMF under conditions prevalent in galaxies with rapid star formation. However, the inferred cluster formation efficiency in previous work is drawn assuming a “normal” Kroupa IMF, so this story is not entirely self-consistent. Because of sensitivity and resolution limits, those studies were forced to extrapolate

Table 1. Five sigma RMS for `tclean` threshold parameter for each pointing. The noise was measured in a region close to the center of the pointing but free of visible emission.

	Band 3 mJy/beam	Band 6 mJy/beam	Coordinates RA, Dec (J2000)
M	0.05	0.5	17:47:20.174-28:23:4.788
N	0.035	0.5	17:47:19.925-28:22:18.608

from single high-mass YSO (HMYSO) candidates to the full stellar population.

We report follow-up observations of Sgr B2 regions M and N in 1 mm and 3 mm. We describe the mass distribution of pre- and protostellar cores in the attempt to understand whether the IMF is top-heavy in young cluster and varies with the surrounding conditions.

2. OBSERVATIONS AND DATA REDUCTION

This paper utilizes the data from ALMA project 2016.1.00550.S from Cycle 4. The observations consist of two pointings centered on Sagittarius B2 M and N. Both locations were observed in Band 3 and Band 6 with a target resolution of 0.05" for both bands, which corresponds to a spatial resolution of about 500 AU. To achieve such resolution the maximum available baseline of 16.2 km was used for Band 3. The data were collected in September 2017.

The data were imaged using CASA version 5.7.0-134.el7 `tclean` (McMullin et al. 2007). The initial cleaned images contained many imaging artifacts present, such as radial streaks, circular patterns, and parallel "wavy" streaks. Therefore, we decided to self-calibrate the data to improve the quality of the final image. The images produced prior to self-calibration were also used to determine the data sensitivity, which then were used in the `tclean` threshold parameter, and to define masks for future cleans. The observed regions have high source density at the central area and contain many extended structures in the field-of-view. The noise was determined based on a region that visually

contained no sources or extended structure in the inner third part of the image. The `tclean` threshold parameter was set to five times the noise. The 5 sigma noise values that were used as threshold parameters are shown in Table 1.

During the data reduction part of the project there were hundreds of versions of images and the reduction code was not properly version controlled at the beginning. Unfortunately, the images that have the lowest noise levels and are used to create the final catalog are not reproducible at this point. The reduction code for images with slightly higher noise is available on GitHub: https://github.com/bazarsen/SgrB2_ALMA_continuum.

Due to the nature of the interferometric images the noise level increases radially: the noise at the edge of the image is roughly five times higher than at the center.

During the self-calibration attempts we discovered that the data did not benefit from amplitude self-calibration and thus only two iterations of phase calibration were performed. Self-calibration decreased the overall image noise by $\sim 30\%$ in Band 3. Band 6 image was dominated by imaging artifacts to the point where no structure outside of the very center of the pointing was visible. Two iterations of phase calibration significantly improved the image, but with some streaks and waves present, as seen on Figure 2.

3. ANALYSIS

3.1. Source extraction

Automating source extraction presents numerous problems in a region like SgrB2. The

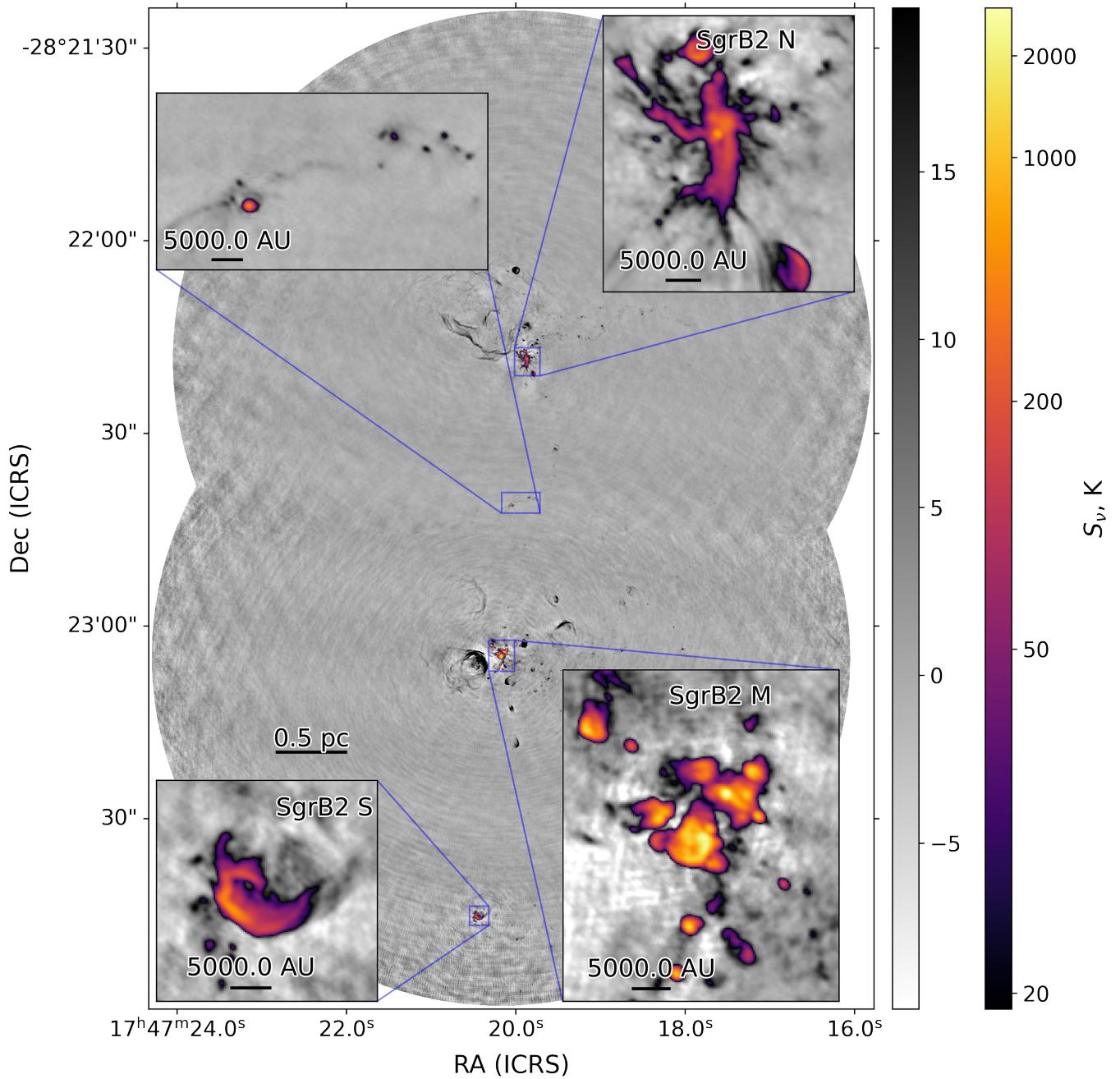


Figure 1. Band 3 (3 mm) continuum image of Sagittarius B2 M and N. The two pointings were imaged separately and then the final images were cut and overlaid to not overlap. The images were subsequently stitched together; the pixels in the central zoom-in on Sgr B2 z come from only the N pointing. The insets show the high-density regions of Sgr B2 N and M, as well as Sgr B2 S and a region around Sgr B2 z.

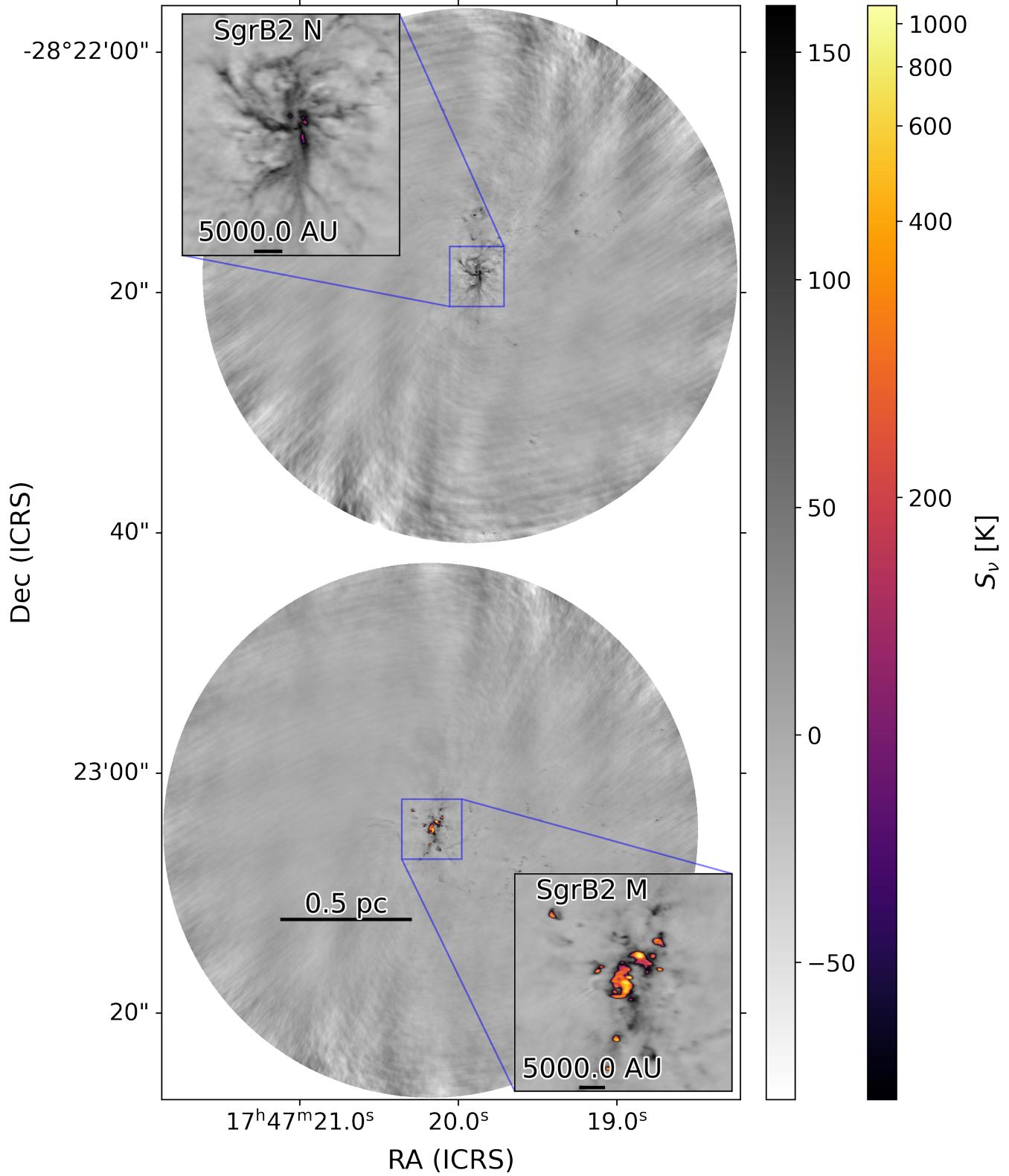


Figure 2. Band 6 (1 mm) continuum image of Sagittarius B2 Main and North. Sgr B2 M appears to be hotter and denser than Sgr B2 N. Due to the smaller field of view, Sgr B2 S and z that are visible in Band 3 are not present here.

images have high dynamic range, contain complex extended structures, and have radially varying noise. Combining this with the irregular structure of the observed objects makes fully automated source detection inefficient. Therefore, we utilize automated source finding algorithms augmented with a by-hand approach described below.

We visually inspect the images and identify regions in each image that contain what appear to be real sources. While somewhat arbitrary, the sources are identified based on their shape (round, slightly elongated, follow the relative size and position angle of the beam), spatial distribution (compact, brighter towards the center), intensity (bright), relative location (bright compact emission inside HII regions is ignored). If there are multiple sources present nearby, we include all of them in one region. For individual sources, we center the region on the visual peak of the source.

Due to the high dynamic range of the data we also create a separate set of regions for the central parts of clouds N and M.

Then, we perform source detection on the primary-beam-corrected images using the `astrodendro` source finding algorithm. We run the algorithm on both pointings twice using different input parameters for the central and outer regions, which produces four dendrogram catalogs. We take five times the RMS determined earlier as a minimum threshold value for the outer region. The inner region threshold was chosen based on visual inspection of the region, about 50 times the RMS. The process is identical for both frequencies.

Using a semi-automated source extraction process can be beneficial to use in unique regions such as Sgr B2. However, it is important to address the drawbacks of such methods. Here we use an automated approach augmented with by-hand source selection and review. The main drawback of the used method is that all sources

that are missed in the initial selection of regions will never be recovered. Thus, we select these regions very liberally, as there will be many opportunities to remove sources in the later stages.

Extended structure and higher noise at the edge of the images result in false positives. We removed all detections outside of the previously defined regions. We utilize the `dendrocat` package to extract further information about the sources: structure contours, locations, fluxes, and local (aperture) signal-to-noise. We also record the brightest pixel value inside each of the dendrogram contours.

We review each of the catalogued sources for both pointings and frequencies by hand. In Band 3 data there is some overlap in the two pointings. Thus, we "draw" a line in between the two images and only use one set of data in the overlapping region. We rank every source as a false detection (0), detection (2), or possible detection (1). To do so, we inspect close up cutouts of the sources. We remove all sources with $\text{SNR} < 5$ based on local noise determined by `dendrocat`. The noise is calculated by taking the source flux from the catalog (which is a summed flux over the dendrogram leaf) and dividing it by the standard deviation of the flux in a 6-pixel-thick annulus with an inner radius of the source major FWHM from the catalog plus 6 pixels. We inspect these rejections manually to make sure that the local noise value was not affected by the presence of sources inside the ring aperture. Every source that has length to width ratio over ~ 4 is rejected. All inconclusive detections are given a score of 1. Then, we inspect each of the detections again, now using the whole image to get context on the source morphology based on location. For example, if a source with a score of 1 that somewhat resembles filamentary structure is located in an area with extended emission, such as an HII region, we mark it as a false detection and adjust the score accordingly. If such a source is located in

a cluster of other sources with a score of 2 and is not visibly a part of any extended structure, we change its score to 2. We further investigate each source with a score of 1 by comparing their structure in 1 mm and 3 mm data. If a potential source is a part of a filamentary structure at 3 mm, but no local peak is observed at 1 mm, we score such a source as 0. The final catalog is available on GitHub: https://github.com/bazarsen/SgrB2_ALMA_continuum.

3.2. Photometry

We use flux within the dendrogram structure contour and the brightest pixel within the contour as the primary tools for analysis.

3.2.1. Photometry comparison

As a part of ALMA project 2013.1.00269.S [Ginsburg et al. \(2018\)](#) obtained Band 3 observations of Sagittarius B2 cloud. We compare our Band 3 data photometrically. [Ginsburg et al. \(2018\)](#) data had basebands centered at 89.48, 91.28, 101.37, and 103.23 GHz with bandwidth 1.875 GHz, while our observations had basebands centered at 85.47, 87.37, 97.42, and 99.42 GHz with the same bandwidth. Nevertheless, imaged continuum should be very similar photometrically. We apply an elliptical aperture representative of the [Ginsburg et al. \(2018\)](#) beam on the locations of [Ginsburg et al. \(2018\)](#) detections on our data. We record the flux within this aperture and apply an aperture correction factor. We compare this quantity with the brightest pixel of [Ginsburg et al. \(2018\)](#) sources. The ratio of flux at each location is shown in Fig. 3.

3.3. Source flux

For each source we recorded the total flux within the dendrogram object contour and the brightest pixel value within the contour. If a source is not resolved, then the brightest pixel should contain all the flux from the source. If a source is resolved, the brightest pixel will contain only part of the flux. Thus, we use a ratio

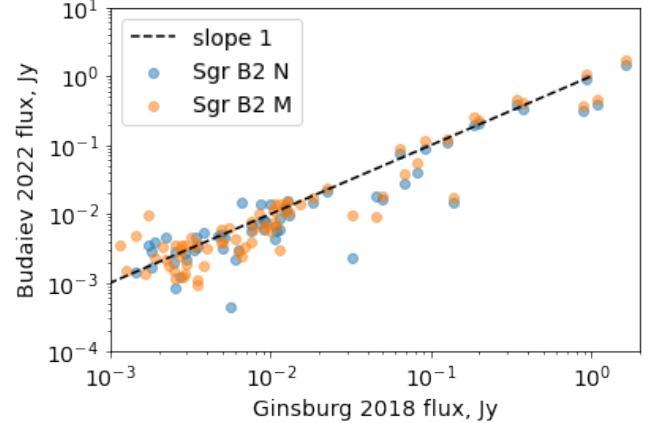


Figure 3. Aperture photometry comparison of our data and [Ginsburg et al. \(2018\)](#) data. Some scatter is present that can be explained by the difference in resolution and the presence of imaging artifacts. The data points from Sgr B2 N and M that are almost identical are caused by the overlap in the two pointings' field of view.

of brightest pixel flux and a summed flux within dendrogram structure to serve as a metric of whether a source is resolved or not.

3.4. Source masses

Using a number of assumptions about the properties of a core, including that the dust is optically thin, we can derive the mass of such a core. Starting from a modified blackbody equation:

$$F_\nu = \pi \frac{2h\nu^3}{c^2} (1 - e^{h\nu/k_B T})^{-1} (1 - e^{-\kappa_\nu \Sigma}), \quad (1)$$

where κ_ν is the dust opacity index and Σ is the surface density. Knowing that

$$M = \Sigma A \quad (2)$$

$$S_\nu = B_\nu \frac{\pi r^2}{d^2} \quad (3)$$

We can apply the optically thin dust assumption ($\kappa_\nu \Sigma \ll 1$) and solve for the mass:

$$M = \frac{S_\nu d^2}{B_\nu \kappa_\nu} \quad (4)$$

Or:

$$M = \frac{S_\nu d^2 c^2}{2\kappa_\nu \nu^2 k_B T}. \quad (5)$$

Using a simplistic uniform temperature assumption and extrapolating the dust opacity index we can derive the source mass of an optically thin dust core. However, as later discussed in Section 4.2 we can only get the lower limit on the source masses as most of the sources in our data are optically thick. To do so, we extrapolate the Ossenkopf & Henning (1994) opacity models for dust grains with ice mantles assuming $\beta_{\text{dust}} = 1.75$ to get the dust opacity index. Then, assuming a uniform core temperature of 50 K and a distance of 8.4 kpc we obtain source masses shown in Figure 7.

4. DISCUSSION

4.1. Source extraction

Since our data includes just two ALMA pointings and the data sets are not combined in the overlapping regions, the noise increases radially. The noise at the edges of the cleaned image is about five times larger than the noise close to the center. This has a significant effect when using only two dendrogram thresholds per pointing: if the threshold value is set to be five times the noise at the center of the image, then the `astrodendro` catalog will contain a lot of false detections at the edges of the image due to higher noise. If the threshold value is selected according to the noise at the edges of the image, then there will potentially be a lot of missed sources in the central regions of the image. This is something that we are planning to improve in the future: we will use several separate catalog extractions with different threshold values that will be used to detect sources withing concentric rings of similar noise.

4.1.1. Photometry

During testing we attempted to perform 2D Gaussian photometry. However, due to the

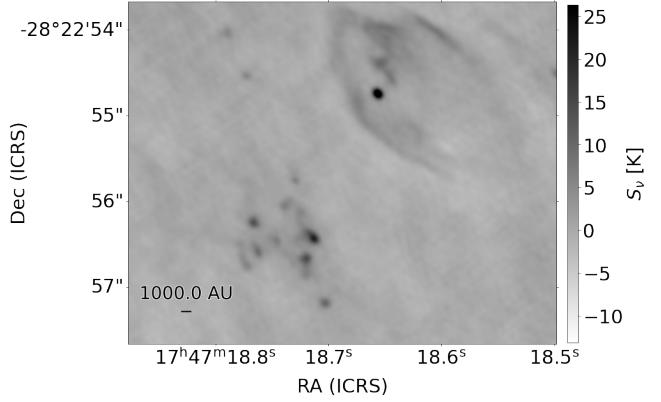


Figure 4. A zoom in on a region north-east from SgrB2 M. It features a highly fragmented source, HII region Y from Schmiedeke et al. (2016), and a core that is within the HII region on the sky. It is not clear whether this core is inside the HII region. The group of seven detected cores was previously thought to be one massive core.

somewhat resolved structure of sources the fitting would often fail or the residuals would still contain a lot of emission. Thus, we decided against Gaussian fitting and instead are using dendrogram contours to define sources.

We used `astrodendro` leaf contours to determine the flux as described in Section 3.3. The main drawback of this method is that if there is extended structure present above the threshold level around the source, all this background will be summed up in the output total flux. On the other hand, if there is a dim extended source with a peak just over the threshold value, the contour will cut the edges of the source off and will result in a lower value of total flux. As discussed above, this is further worsened by the use of one catalog extraction with one threshold value for most of the image. Since the threshold value was picked based on the noise of the inner part of the image, sources located at the edges where noise is higher will tend to include more of the surrounding area and their flux will be overestimated.

4.1.2. Photometry comparison

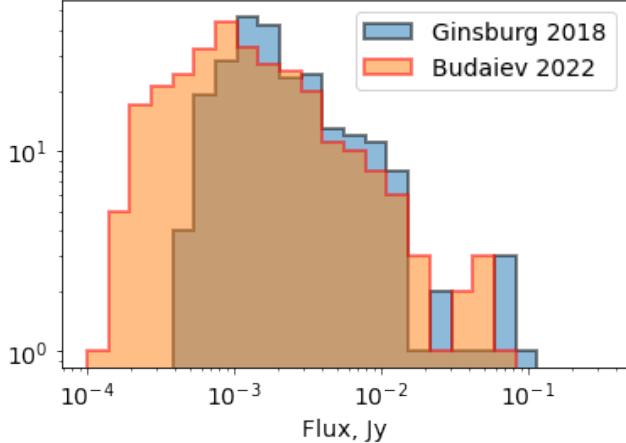


Figure 5. A comparison of source fluxes of our data and the full [Ginsburg et al. \(2018\)](#) catalog, which covers a larger 15 by 15 pc area, with known HII regions removed from both data sets. The general shape appears to be similar, but with the higher resolution data we are able to detect fainter sources down to 0.1 mJy. The turnover point is at about 1 mJy, which is just below the turn point in [Ginsburg et al. \(2018\)](#) catalog.

Due to the factor of 10 difference in spatial resolution, different sensitivity, and high number of clustered sources, direct comparison of cataloged source fluxes between our and [Ginsburg et al. \(2018\)](#) catalogs is problematic. Instead of comparing the flux values in the two catalogs, we perform aperture photometry as described in Section 3.2.1. Due to the larger minimum baseline length our data could have some flux resolved out. If that is the case, we would expect a consistently smaller flux in our data than in [Ginsburg et al. \(2018\)](#).

As shown on Figure 3, there is a slight trend for higher resolution sources having less flux. A smaller fraction of sources are above the line. Most of such sources are faint and are more affected by the noise. The somewhat large scatter can be partially explained by the presence of imaging artifacts in the images. If the aperture was just on the edge of a negative bowl or a bright point source we expect a noticeable change in flux ratio.

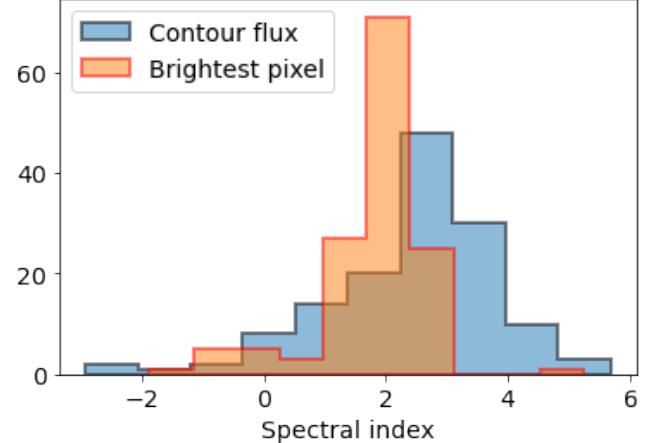


Figure 6. Spectral indexes of the sources in our catalog. We cross-matched 100 GHz and 225 GHz detections and calculated the spectral indexes using Equation 7. We used two approaches: using the brightest pixel of each source and using the summed up flux within the source contour.

4.2. Spectral indexes and source masses

[Ginsburg et al. \(2018\)](#) assumed an optically thin dust to approximate the source mass. We find that most of the sources are most likely optically thick. We measure the spectral indexes in two ways, using summed flux within dendrogram leaf contours and the brightest pixel value for each source. Utilizing our 100 GHz and 225 continuum flux values and spectral index equation we can calculate the spectral index for each source that has been cross-matched between the two frequencies. A histogram of spectral indexes are shown in Fig. 6.

$$S_\nu \propto \nu^\alpha \quad (6)$$

$$\alpha = \frac{\log(\frac{S_{100\text{GHz}}}{S_{225\text{GHz}}})}{\log(\frac{100\text{GHz}}{225\text{GHz}})} \quad (7)$$

The spectral indexes that are based on the source contours would work best for unresolved sources, which constitute a small part of our sample. Furthermore, as discussed previously some flux values measured in such a method might overestimate the flux.

The brightest pixel of each source contains information about at least the inner, most dense parts of the source in its 2D projection. Our assumed $\beta_{\text{dust}} = 1.75$ implies a spectral index of ~ 3.75 for an optically thin dusty core. Instead, we see that the spectral index based on the brightest pixel is ~ 2 for most of the source. This indicates that at least part of the source is optically thick and we cannot use the optically thin dust assumption to derive source masses. The sources with spectral index < 2 are most likely HII regions and the emission is free-free. While we cannot apply the optically thin assumption to our sources, we can still invoke it to estimate the lower mass limit for each source: other variables being constant, an optically thick source contains more mass than an optically thin source.

We assumed a uniform temperature of 50 K to obtain lower mass limits for each source shown in Figure 7. It is important to note that the temperature of observed cores is not uniform and can vary between < 20 K and > 150 K. Still, using such assumptions can give us some understanding of possible mass ranges and show the general shape of the core mass distribution. It is unclear whether the turnover point is due to intrinsic shape of the Core Mass Function (CMF) or because our mass sensitivity is close to that level.

We can perform some crude calculations to evaluate whether such core masses fall within an expected range. Assuming that about a third of a core mass will end up in a star we can describe the potential IMF of this cluster. If the turnover point is indeed caused by our data sensitivity, we can expect to be complete down to at least $8 M_{\odot}$. Knowing this, we can extrapolate Kroupa (2001) IMF below $8 M_{\odot}$ and measure the total mass of the cloud. Removing HII regions cross-matched with Schmiedeke et al. (2016), we measure the total stellar mass to be $\sim 250,000 M_{\odot}$ using 3 mm data and $\sim 45,000 M_{\odot}$ using 1

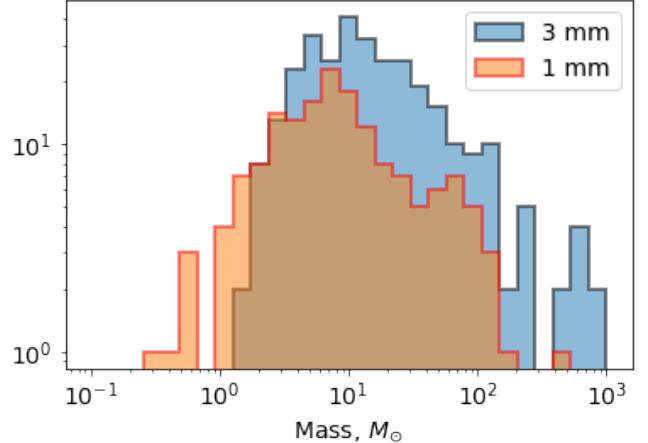


Figure 7. The inferred lower mass limit of the sources using the optically thin dust assumption for 1 mm and 3 mm data. 1 mm data has a higher mass sensitivity that is most likely caused by the dust becoming more optically thick at higher frequencies and thus lower mass sources can be detected.

mm data. The result from 1 mm observations is on par with Schmiedeke et al. (2016): they measured the mass of just the Sgr B2 N and M combined to be $\sim 25,000 M_{\odot}$. The $250,000 M_{\odot}$ total mass inferred from 3 mm data is less believable even considering the larger field of view that includes Sgr B2 S and z. Such huge mass is probably caused by a few HII regions being included in the calculation and contributing a large mass fraction to the sample.

4.3. Future work

As we continue working on this project we plan to improve the robustness of the current catalog and extract further scientific results from the data. Our first goal is to re-clean and re-calibrate the data in a consistent and reproducible manner. We expect that this will also cause a slight improvement in the overall image noise. The current reduction code released on GitHub will be updated appropriately. Then, we plan to refine the source extraction process to be less affected by the nonuniform noise of the interferometric images. This will be achieved by running the dendrogram algorithm with dif-

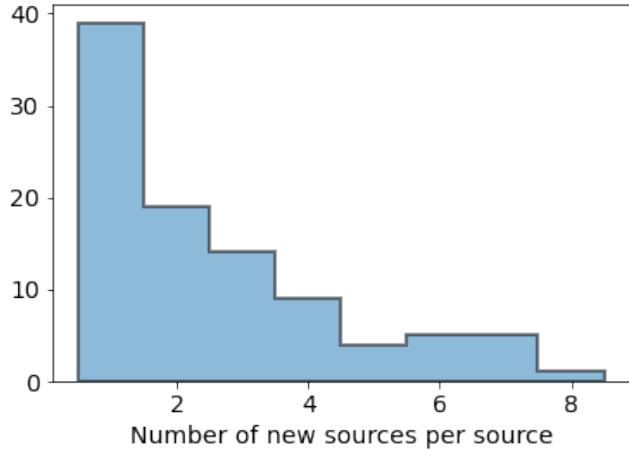


Figure 8. Number of sources from our data that fall within 2 FWHM of [Ginsburg et al. \(2018\)](#) sources. Over 50% of the sources contain two or more sources. This indicates an even higher Cluster Formation Efficiency than previously thought.

ferent threshold values on multiple concentric rings. A more consistent local signal-to-noise cutoff throughout the image will prevent background noise from attributing to the source flux and thus improve the accuracy of the measured flux.

With the improved source extraction process we expect to see an improvement of source contours defined by the dendrogram algorithm.

Based on the improved catalog we will measure the Cluster Formation Efficiency (CFE) and compare our results with a $\sim 35\%$ CFE result from [Ginsburg & Kruijssen \(2018\)](#). Just from visual comparison, we see that many previously thought to be massive sources have split up into multiple smaller sources. In fact, more than 50% of sources ended up being fragmented into two or more sources as shown on Figure 8. Thus, we expect more stars to form in gravitationally bound clusters than previously thought.

5. SUMMARY

We presented high-resolution follow-up observations of a massive molecular cloud in the CMZ

– Sagittarius B2. With a tenfold increase in resolution, we reported the detection of 314 3 mm and 163 1 mm sources, a factor of two increase in number of 3 mm sources. Most of the newly detected sources come from an increased sensitivity, while about a third are due to resolving fragmented sources. We find most of our sources to be spatially resolved and at least partially optically thick. We describe the source lower limit mass distribution. We find that the inferred stellar mass of our sample is similar to the previous measurements of stellar mass in the region.

ACKNOWLEDGMENTS

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2016.1.00550.S, ADS/JAO.ALMA#2013.1.00269.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

Facilities: ALMA

Software: Image reduction and final catalogs are available on GitHub: https://github.com/bazarsen/SgrB2_ALMA_continuum. This work utilized the following tools: astropy ([Astropy Collaboration et al. 2013, 2018](#)), CASA ([McMullin et al. 2007](#)), `astrodendro`, `dendrocat`, `radiobeam`.

REFERENCES

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- Barnes, A. T., Longmore, S. N., Battersby, C., et al. 2017, MNRAS, 469, 2263, doi: [10.1093/mnras/stx941](https://doi.org/10.1093/mnras/stx941)
- Ginsburg, A., & Kruijssen, J. M. D. 2018, ApJL, 864, L17, doi: [10.3847/2041-8213/aada89](https://doi.org/10.3847/2041-8213/aada89)
- Ginsburg, A., Bally, J., Barnes, A., et al. 2018, ApJ, 853, 171, doi: [10.3847/1538-4357/aaa6d4](https://doi.org/10.3847/1538-4357/aaa6d4)
- Henshaw, J. D., Longmore, S. N., Kruijssen, J. M. D., et al. 2016, MNRAS, 457, 2675, doi: [10.1093/mnras/stw121](https://doi.org/10.1093/mnras/stw121)
- Hosek, Matthew W., Lu, J. R., Anderson, J., et al. 2019, ApJ, 870, 44, doi: [10.3847/1538-4357/aaef90](https://doi.org/10.3847/1538-4357/aaef90)
- Kroupa, P. 2001, MNRAS, 322, 231, doi: [10.1046/j.1365-8711.2001.04022.x](https://doi.org/10.1046/j.1365-8711.2001.04022.x)
- Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: [10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615)
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127
- Ossenkopf, V., & Henning, T. 1994, A&A, 291, 943
- Schmiedeke, A., Schilke, P., Möller, T., et al. 2016, A&A, 588, A143, doi: [10.1051/0004-6361/201527311](https://doi.org/10.1051/0004-6361/201527311)