







Characterizing the 3D Structure of Molecular Cloud Envelopes in the Cloud Factory Simulations

Elijah Mullens^{1,2} , Catherine Zucker^{1,3} , Claire E. Murray^{1,4} , and Rowan Smith^{5,6} ¹Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA²Department of Astronomy and Carl Sagan Institute, Cornell University, 122 Sciences Drive, Ithaca, NY 14853, USA³Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138 USA⁴Department of Physics & Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA⁵School of Physics and Astronomy, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9SS, UK⁶Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Received 2023 December 6; revised 2024 March 1; accepted 2024 March 4; published 2024 May 2

Abstract

We leverage recent numerical simulations of highly resolved star-forming regions in a Milky Way–like galaxy to explore the nature of extended gaseous envelopes around molecular clouds. We extract a sample of two dozen star-forming clouds from the feedback-dominated suite of Cloud Factory simulations. With the goal of exploring the 3D thermal and chemical structure of the gas, we measure and fit the clouds’ radial profiles with multiple tracers, including n_{H_1} , n_{H_2} , $n_{\text{H}_{\text{tot}}}$, n_{CO} , and gas temperature. We find that while solar neighborhood clouds recently detected via 3D dust mapping have radially symmetric, low-density envelopes that extend ~ 10 – 15 pc, the simulated cloud envelopes are primarily radially asymmetric with low-density envelopes that extend only ~ 2 – 3 pc. One potential explanation for the absence of extended envelopes in the simulated clouds may be the lack of magnetic fields, while a stronger local feedback prescription compared to solar neighborhood conditions may drive the radially asymmetric cloud morphologies. We make the pipeline used to extract and characterize the radial profiles of the clouds publicly available, which can be used in complementary and future simulations to shed additional light on the key physics shaping the formation and evolution of star-forming structures in the Milky Way.

Unified Astronomy Thesaurus concepts: [Diffuse molecular clouds \(381\)](#); [Molecular clouds \(1072\)](#); [Interstellar clouds \(834\)](#); [Hydrodynamical simulations \(767\)](#)


1. Introduction

An appreciable fraction of the mass budget of the interstellar medium (ISM) belongs to dense, cold formations termed molecular clouds (e.g., Blitz 1993). These molecular clouds are the sites of star formation in galaxies, and as a result play a key role in galaxy evolution. The structure of a molecular cloud is subject to a complicated interplay between galactic dynamics, self-gravity, feedback, and magnetic fields (e.g., Dobbs et al. 2014). Understanding the structure of molecular clouds is therefore integral not only to understanding these physical mechanisms, but also understanding where, when, and how stars form and interact with their broader galactic environment (e.g., Chevance et al. 2022).

Molecular cloud structure has been extensively studied using both observations and simulations. Traditionally, observations of molecular cloud structure have largely been limited to either 2D projected space (via integrated plane-of-the-sky dust emission or dust extinction maps; André et al. 2010; Froebrich & Rowles 2010; Lombardi et al. 2010; Planck Collaboration et al. 2011), or 3D “position–position–velocity” space obtained via carbon monoxide (CO) spectral-line mapping (Dame et al. 2001; Jackson et al. 2006; Ridge et al. 2006; Colombo et al. 2019; Duarte-Cabral et al. 2021), where the third axis is the radial velocity of the gas obtained from the Doppler effect, not distance.

However, with the launch of Gaia, true 3D spatial, “position–position–position” maps of the ISM can now be reconstructed with the distance resolution necessary to resolve the internal structure of molecular clouds, owing to a technique known as 3D dust mapping (Green et al. 2019; Rezaei & Kainulainen 2022; Vergely et al. 2022; Dharmawardena et al. 2023). Notably, Leike et al. (2020) produce a highly resolved ~ 2 pc resolution 3D dust map of the solar neighborhood (out to a distance ~ 400 pc from the Sun) using distance and extinction estimates to a large number of stars inferred from a combination of Gaia DR2 astrometric data and Gaia, Two Micron All Sky Survey, Pan-STARRS, and ALLWISE photometric data (see Anders et al. 2019).

Zucker et al. (2021) utilize the 3D dust map from Leike et al. (2020) to create 3D models of local molecular clouds and characterize their 3D spatial structure.⁷ Zucker et al. (2021) find that local molecular clouds are filamentary, and extract radial volume density profiles from the 3D dust map to study the extended structure of clouds as a function of distance from their filamentary “spines.” Zucker et al. (2021) find that the averaged radial volume density profiles of local molecular clouds are best fit with a two-component Gaussian function, where an inner Gaussian fits the higher-density peak near the core of the cloud and an outer Gaussian fits the lower-density extended tail. Zucker et al. (2021) propose that the two separate Gaussians either represent a chemical phase transition (between low-density atomic gas in the outskirts and dense molecular gas

 Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

⁷ See also Dharmawardena et al. (2023) for another work examining the 3D spatial structure of local molecular clouds, and Figure 6 in Cahlon et al. (2024) for a comparison between the Dharmawardena et al. (2023) and Leike et al. (2020) results.

near the core of the cloud) or an atomic gas thermal phase transition (between the unstable neutral medium (UNM) near the outskirts of the cloud and the cold neutral medium (CNM) near the core of the cloud). However, observational 3D dust data only probes differential extinction, which can only approximate the total hydrogen gas volume density within the cloud thanks to constraints on the wavelength-dependent extinction curve (Draine 2009). Any information on gas phase, composition, and temperature is not directly accessible from 3D dust maps alone. In order to test the hypothesis from Zucker et al. (2021), we instead turn to simulations of molecular clouds.

Simulations of molecular clouds over the last decade have progressed from isolated clouds with only one or two feedback mechanisms to simulations that follow molecular cloud lifetimes in full with both local and galactic contexts included. One goal of a specific subset of simulations is to elucidate the relationship between molecular cloud structure and its chemical composition as the cloud evolves over time and forms both dense gas and young stars. These simulations have tied together chemical processes that form molecular hydrogen (H_2) and CO with the large-scale hydrodynamic evolution of molecular clouds (e.g., Armillotta et al. 2020; Seifried et al. 2020; Hu et al. 2021, 2021; Jeffreson et al. 2021; Kim et al. 2023).

In particular, the Cloud Factory simulations of Smith et al. (2020) study the effects of the galactic potential, gravitational collapse, and supernova feedback on the formation and destruction of H_2 and CO. More broadly, 3D simulations like the Cloud Factory track a network of filamentary clouds and their chemical compositions, which provide a large set of simulated clouds to compare with multiwavelength observations probing the chemical and thermal structure of Milky Way clouds (see Chevance et al. 2022, for a more complete review of resolved cloud simulations).

With the advent of Gaia and 3D dust mapping, there are now sufficient observational data on the 3D structure of molecular clouds to facilitate a comparison to 3D simulated filamentary clouds. In this paper, we extend the methodology of Zucker et al. (2021) and develop a pipeline that efficiently compares the 3D structure detected in resolved numerical simulations of star-forming clouds to new 3D observational data. Leveraging this pipeline we aim to extract the radial profiles and morphologies of star-forming clouds in the Cloud Factory simulations to examine differences in cloud structure detected in observations and simulations quantitatively. After we show that the simulations quantitatively reproduce the observations, a pipeline is built to examine whether the two-component Gaussian structure of molecular clouds identified in previous 3D dust mapping observational work may be tracing a chemical or thermal phase transition.

In Section 2, we summarize the Cloud Factory suite of simulations utilized in this work. In Section 3, we detail the methods used to analyze the simulations. In Section 4, we summarize the results of the pipeline applied to the Cloud Factory simulations and highlight the range of structures detected in an ensemble of the clouds. In Section 5 we discuss the results and propose potential explanations for discrepancies between simulations and observations. Finally, we conclude in Section 6.

2. The Cloud Factory

The aim of the Cloud Factory simulations is to investigate the formation of cold dense structure in the ISM while

including the galactic-scale forces thought to be responsible for cloud assembly. Specifically the simulations contain an analytic galactic potential, and therefore naturally have disks with differential rotation and well-defined spiral arms. Full details can be found in Smith et al. (2020) but we summarize them here.

The models are performed using the AREPO code (Springel 2010) with custom physics modules to describe star formation and cold, dense gas. The chemical evolution of the gas is modeled using the hydrogen chemistry of Glover & Mac Low (2007a, 2007b), together with the highly simplified treatment of CO formation and destruction introduced in Nelson & Langer (1997). Our modeling of the hydrogen chemistry includes H_2 formation on grains, H_2 destruction by photodissociation, collisional dissociation of atomic hydrogen, H^+ recombination in the gas phase and on grain surfaces (see Table 1 of Glover & Mac Low 2007a), and cosmic-ray ionization. We assume that the strength and spectral shape of the UV portion of the interstellar radiation field (ISRF) are the same as the values for the solar neighborhood derived by Draine (1978; equivalent to $1.7\times$ the field strength derived by Habing 1968). To treat the attenuation of the ISRF due to H_2 self-shielding, CO self-shielding, the shielding of CO by H_2 , and by dust absorption, we use the TREECOL algorithm developed by Clark et al. (2012) assuming a shielding length of $L_{\text{sh}} = 30$ pc. We adopt a cosmic-ray ionization rate of $\xi_{\text{H}} = 3 \times 10^{-17} \text{ s}^{-1}$ for atomic hydrogen, and a rate twice this for molecular hydrogen. Finally we assume a solar metal abundance, and a 100:1 gas-to-dust ratio. Heating and cooling of the gas is computed simultaneously with the solution of the chemical rate equations.

Star formation is modeled via sink particles (Bate et al. 1995), which are nongaseous particles that represent collapsing regions of gas that will form small (sub)clusters of stars. These are formed by checking if regions of gas exceed a critical density and are unambiguously bound, collapsing, and accelerating inwards. Only if these criteria are met will the gas be replaced with a sink particle, which can then accrete additional mass that falls within a chosen accretion radius of the cell if it is gravitationally bound to it.

Using the model of Sormani et al. (2017), we sample the initial mass function and associate supernovae with the massive stars as described by Tress et al. (2020). For each supernova, we calculate an injection radius, which is the radius of the smallest sphere centered on the supernova that contains at least 40 grid cells. If the injection radius is smaller than the expected radius of a supernova remnant at the end of its Sedov–Taylor phase, we inject thermal energy from the supernova; otherwise, we inject momentum (e.g., Gatto et al. 2015). Mass is returned with each supernova explosion such that when the last supernova occurs the gaseous component of the sink is exhausted. The sink is then turned into a star particle. To account for Type Ia supernovae, we also randomly select a star particle every 250 yr and create a supernova event at its position.

The gravitational potential of nongaseous cells is determined using an analytic potential. For the axisymmetric part of the potential, we use the best-fitting model of McMillan (2017), which was created to be consistent with various observational and theoretical constraints for the Milky Way and consists of the sum of bulge, disk, and halo components. We then include a spiral perturbation to the potential, generated in the same way as in Smith et al. (2014). Briefly, we use a four-armed spiral

component from Cox & Gómez (2002) with a pitch angle $\alpha = 15^\circ$ and a pattern speed of $2 \times 10^{-8} \text{ rad yr}^{-1}$.

The initial condition is inspired by the Milky Way gas disk model of McMillan (2017), which is based on a combination of observational constraints and theoretical modeling. It consists of two density distributions for H I and H₂ that decline exponentially at large radii. As we focus on Milky Way–like clouds outside the central bar, we neglect galactic radii smaller than 4 kpc. The gas disk is given the initial rotation curve that arises from the analytic potential, which for our disk corresponds to a rotation curve of order 220 km s^{-1} .

The simulations initially have a base mass resolution of $1000 M_\odot$. However once a steady state has been reached (after 150 Myr) we turn on refinement for two spiral arm passages ($\sim 70 \text{ Myr}$) within a 3 kpc box that corotates with the gas centered on a galactic radius of 8 kpc. In this high-resolution region the gas has a target mass of initially $100 M_\odot$ for the first 60 Myr, but it is further lowered to $10 M_\odot$ for the final 10 Myr. In addition to the mass requirement we always require that the Jeans length is resolved by at least four cells up to our sink creation density. Within the 3 kpc box, sinks form above a minimum creation density of 574 cm^{-3} (typically about 10^4 cm^{-3} in practice once energy checks are satisfied) and have an accretion radius of 0.1 pc. To avoid discontinuous jumps in the cell size, particularly where the target resolution is changing at the boundaries of the high-resolution box, we require that the cell radius of adjacent AREPO cells can differ by no more than a factor of 2 at any time throughout the entire simulation volume.

The Cloud Factory simulations proceed to refine to even smaller scales in selected clouds, but for this study we are interested in the statistics of a large area and so we use only the 3 kpc box. The original models also had two versions of this 3 kpc box: a “potential-dominated” version and a “feedback-dominated” version. The potential-dominated version is where gas self-gravity was only turned on for the lifetimes of current molecular clouds (5 Myr), and supernova feedback is random. The feedback-dominated version is where gas self-gravity has been turned on for multiple molecular cloud generations (50 Myr) and supernova feedback is tied to the sink particles that represent star formation. The aim of the two versions is to focus on the differences arising between two extreme cases. In this paper we focus primarily on the feedback-dominated case, as it more faithfully reproduces the larger scale height of the disk in the solar neighborhood seen in recent 3D dust maps (Leike et al. 2020; Edenhofer et al. 2023). We emphasize, however, that the feedback-dominated case is likely an upper limit on the strength of supernova feedback experienced in the ISM. We additionally take a 0.5 kpc subset of the z -direction since the z -direction in the 3 kpc box is largely empty. We revisit the potential-dominated case in Section 5.

3. Methods

Here we outline our methodology for characterizing the structure of molecular clouds in the Cloud Factory simulations. We start by subdividing the Cloud Factory simulations into smaller subgrids and identifying subgrids with an appreciable fraction of molecular star-forming gas (Section 3.1). Then, building on the methodology of Zucker et al. (2021), we iteratively extract and “skeletonize” simulated molecular clouds in the simulations (Section 3.2), compute their radial profiles (Section 3.3), and fit the radial volume density profiles with one- and two-component Gaussian functions (Section 3.4). After

extracting the clouds and characterizing their radial profiles, we also sort clouds based on whether they contain sink particles, in order to correlate their observed cloud structure with their degree of star formation.

3.1. Converting and Sorting Grids

As discussed in Section 2, we focus primarily on the feedback-dominated suite of the Cloud Factory simulations. However, we repeat the analysis detailed in the following sections on one grid of the potential-dominated suite and report preliminary results in Section 5. We start by dividing the feedback-dominated box (a $3 \times 3 \times 0.5 \text{ kpc}$ slab) into subregions. Each subregion is a $500 \times 500 \times 500 \text{ pc}$ box that has been regrided to a uniform 1 pc^3 resolution. The Cloud Factory is built on the AREPO code (Springel 2010; Pakmor et al. 2016). The Cloud Factory implementation of AREPO tracks the total gas density (ρ_{gas}) and temperature, as well as the carbon monoxide abundance (x_{CO}), molecular hydrogen abundance (x_{H_2}), and ionized hydrogen abundance (x_{H^+}) with respect to the total gas density.

We convert AREPO units to physical units and derive the total volume density of hydrogen nuclei using the mean molecular weight of hydrogen ($n_{\text{H}_{\text{tot}}} = \frac{\rho_{\text{gas}}}{1.4 \times m_p}$, where 1.4 accounts for the helium abundance). Using the total gas density and abundance grids we then derive the volume density of carbon monoxide, molecular hydrogen, and ionized hydrogen ($n_{\text{CO,H}_2,\text{H}^+} = x_{\text{CO,H}_2,\text{H}^+} \times n_{\text{H}_{\text{tot}}}$). The atomic hydrogen (H) volume density is derived by subtracting the contributions of the molecular and ionized hydrogen from the total hydrogen gas density ($n_{\text{H}_I} = n_{\text{H}_{\text{tot}}} - 2n_{\text{H}_2} - n_{\text{H}^+}$). We then narrow down the grids based on the amount of molecular gas contained within them. We choose only to analyze grids containing at least 1000 voxels ($>1000 \text{ pc}^3$ volume, equivalent to a spherical cloud with a radius of $\approx 6 \text{ pc}$) with a relative abundance $x_{\text{H}_2} > 0.4$. Grids with appreciable fractions of molecular gas provide the best comparison to observed molecular clouds, and are utilized in the remainder of the analysis. We identify nine grids that fit this criterion, namely x1465y1390z1400, x1465y1400z1395, x1465y1400z1400, x1465y1405z1395, x1465y1400z1400, x1470y1400z1395, x1470y1400z1400, x1475y1390z1395, and x1475y1390z1400.

3.2. Simulated Cloud Identification and Skeletonization

In previous work molecular clouds and their substructure have been characterized by topological skeletons (Men’shchikov 2013; Koch & Rosolowsky 2015). Topological skeletons are 1 pixel-wide representations of 2D or 3D shapes that are equidistant to the boundary surface of the original shape. Topological skeletons are derived by starting from the boundary of a shape and subtracting off 1 pixel, moving inwards until only a single pixel-wide shape remains. For example, the skeletonization of a 2D circle will result in a single point in the center, and the skeletonization of a 3D cylinder will result in a straight line through the center of the cylinder.

Following the formalism of Zucker et al. (2021) we compute topological skeletons of simulated molecular clouds by employing a 3D version of the FilFinder package (Koch & Rosolowsky 2015). We first apply a density threshold mask of $n_{\text{H}_{\text{tot}}} > 35 \text{ cm}^{-3}$ to the entire simulation grid, following the same line of reasoning as Zucker et al. (2021). Individual features that survive masking are then iteratively identified. This is accomplished by looping over each voxel in the grid

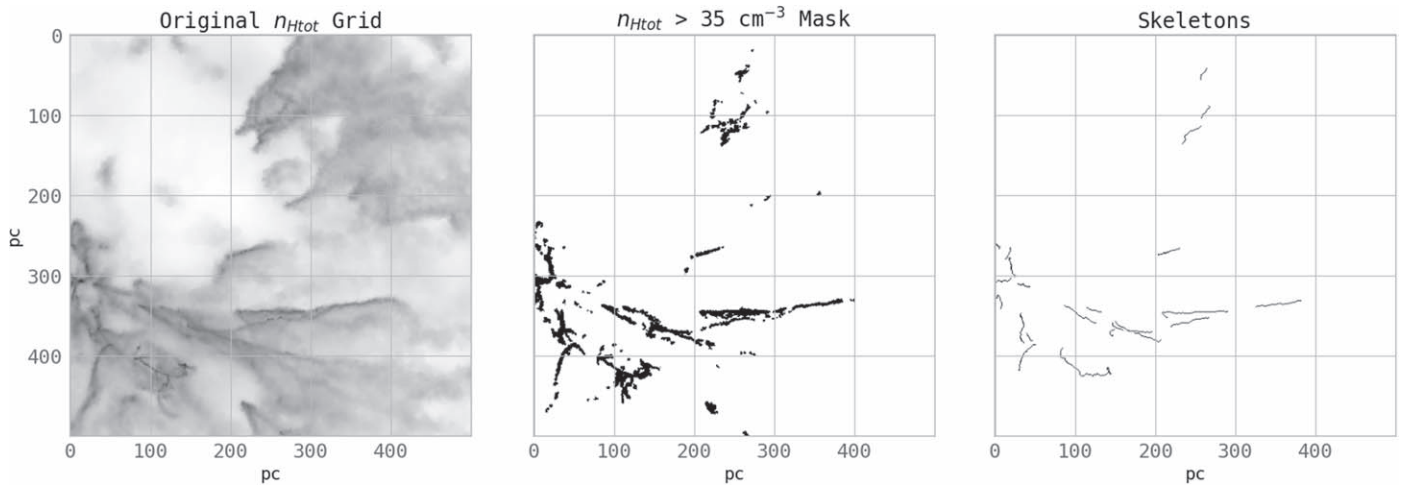


Figure 1. Example of cloud identification and skeletonization for a single simulation subgrid. The 3D grid is displayed with one axis collapsed. The first panel displays the total volume density of hydrogen nuclei in the grid. The second panel displays the same simulation grid after applying a density threshold mask of $n_{\text{Htot}} > 35 \text{ cm}^{-3}$, which is used to identify potential clouds. The third panel displays the resulting topological skeletons. Small clouds present in the second panel are removed from the data set and are not skeletonized. (Grid x1465y1400z1400.)

that survives thresholding and then flooding the regions surrounding those voxels. If a flooded feature contains more than 1000 nonempty voxels, then morphological closing via dilation and erosion is performed to fill small holes. The filling of small holes ensures the skeletonization can be performed efficiently. If a feature has fewer than 1000 voxels, that feature is removed from future cloud-seeking iterations. This ensures that small, spurious features are not analyzed. Significant features are then skeletonized using FilFinder. The skeleton length and total feature mass are derived, and if a skeleton’s total length is greater than 20 pc the skeleton is kept for future analysis. A 20 pc minimum length is chosen as the approximate lower bound due to the range of lengths observed for molecular clouds in the solar neighborhood (Zucker et al. 2021). See Figure 1 for a visual of the cloud identification and skeletonization process.

The simulated clouds can also be sorted based on whether or not they are forming stars. We flag star-forming clouds by iterating over each cloud and determining if there are any “star-forming sinks” (see sink particles introduced in Section 2) inside the cloud mask. If a cloud contains any star-forming sinks it is labeled as a star-forming cloud. This specification allows for a better cross-comparison between subsets of simulated clouds and the observed clouds in Zucker et al. (2021), which are all star-forming. Specifically, Zucker et al. (2021) analyzed a subset of clouds from the Star Formation Handbook (Reipurth 2008a, 2008b), which summarizes the 60 most important star-forming regions within 2 kpc of the Sun. We refer readers to Reipurth (2008a, 2008b) for more information on the star-forming properties of the observed clouds.

3.3. Simulated Cloud Analysis Pipeline

Once we extract a set of skeletons following the procedure outlined in Section 3.2, we then calculate normal vectors between skeleton points in order to define cloud-bisecting planes to derive radial profiles. This is accomplished by fitting a linear spline to the skeleton points (a function that connects individual points with piecewise linear functions), from which the first-order Cartesian derivatives for each point in the skeleton are derived. These derivatives are used to define a

normal vector pointing from the current point being analyzed to the next ordered point in the skeleton via $(-1, -\frac{dy}{dx}, -\frac{dz}{dx})$. This normal vector is used to define a circular 2D plane with a radius of 20 pc that cuts through the current point being analyzed and is orthogonal to the direction of the skeleton (see Figure 2).

An integral difference to the analysis done in Zucker et al. (2021) and this work is that the simulation grids have access to more information than the dust density that was converted to the total hydrogen gas density using a wavelength-dependent extinction curve. The slices defined herein are purely geometric in Cartesian space and do not encode any density or temperature information. We apply the slices to the simulation grid in order to extract and interpolate information on the density (n_{Htot} , n_{CO} , n_{H_2} , n_{H} , and n_{H^+}) and gas temperature (T_{gas}). We then derive radial profiles for each tracer. On a slice-by-slice basis the radial profile for each tracer is computed. We do this by first defining 40 radial distance bins (radius of 0–20 pc extending from the skeleton point with steps of 0.5 pc). For each radial distance bin we compute the median of the density and temperature values (see Figure 2). This process is repeated for each skeleton point and results in a set of radial profiles for each skeleton point in a cloud. Finally, each individual radial profile is averaged to obtain the cloud-averaged radial volume density profile of each chemical tracer as well as temperature profiles (Figure 3). In this work, only the skeleton points which have core values (taken from the first radial bin) greater than the density threshold mask ($n_{\text{Htot}} > 35 \text{ cm}^{-3}$) are included in the computation of the cloud average radial profiles to ensure we are not including spurious low-density patchy regions which may dilute the averaged profile.

The total gas density (n_{Htot}) radial profiles can be used directly as a basis for comparison with observations, since all 3D observational data from Zucker et al. (2021) are converted from the native units of the 3D dust map to n_{Htot} . If a simulated cloud’s total gas density radial profile is comparable to observations, a mapping can be formulated from n_{Htot} to the chemical and temperature profiles. This mapping can be then applied to observed radial profiles to test whether or not a chemical or temperature phase transition is occurring within a cloud.

Slices that bisect the cloud serve a dual purpose. In addition to providing a tool by which radial profiles can be extracted, they also provide a visualization of the morphology of a cloud. Comparative

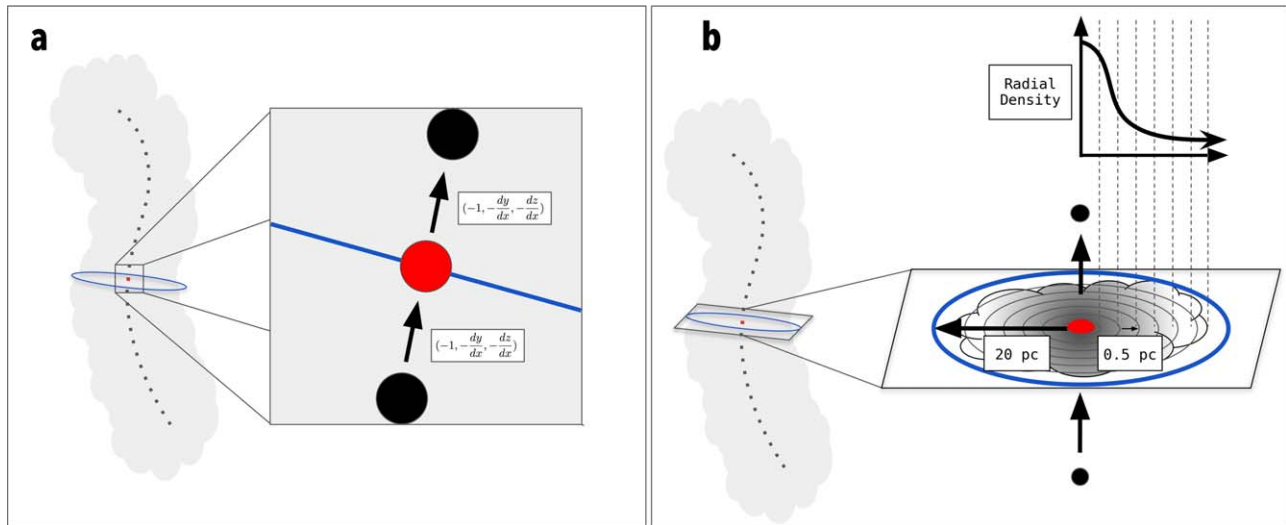


Figure 2. Graphical demonstration of the analysis pipeline. (a) Zoom-in view displaying the normal vector pointing between skeleton points. This normal vector is used to define a plane that bisects the cloud. (b) Zoom-in display of the 20 pc “slice” around a single skeleton point, which is then split into 0.5 pc radial bins and used to derive an average radial density profile for one skeleton point.

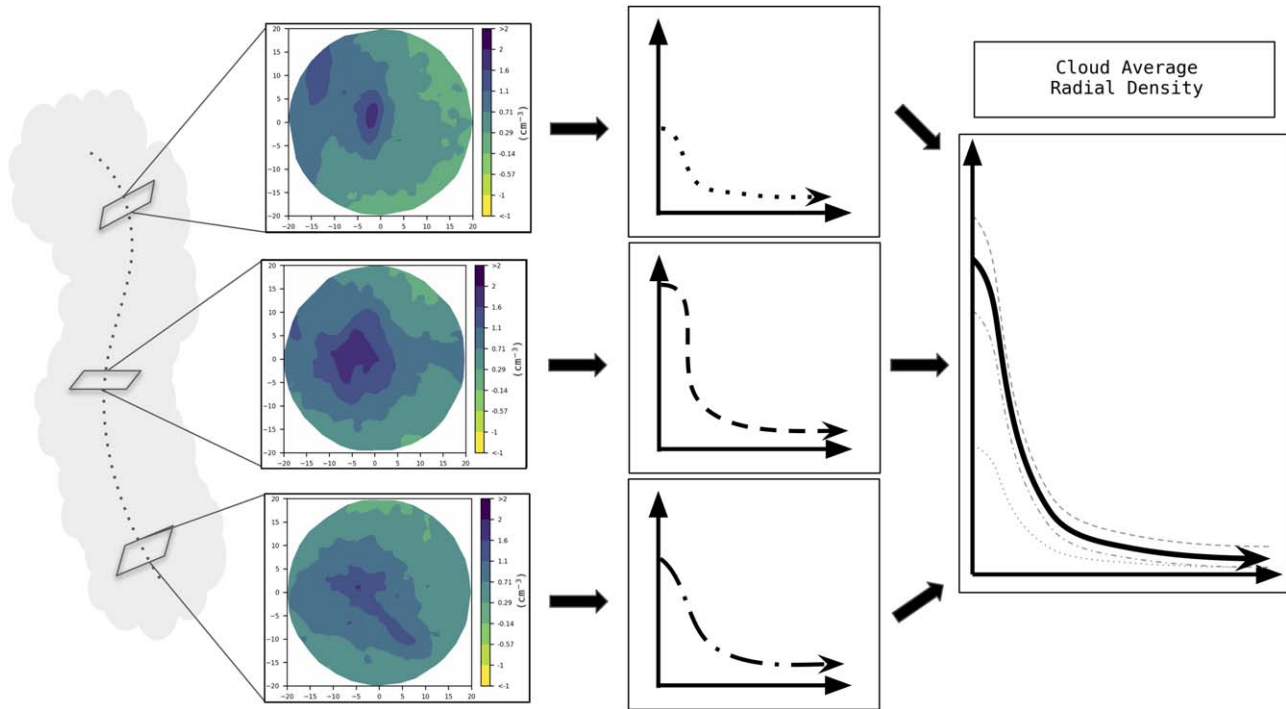


Figure 3. Deriving the cloud average radial volume density profile for an entire cloud. Density slices are computed for each skeleton point. The radial volume density profile for each skeleton point is derived and then averaged to construct a cloud-averaged radial volume density profile. Slices shown are from an observed molecular cloud (Perseus; Zucker et al. 2021) and individual radial volume density profiles are exaggerated.

analysis between simulated cloud and observed cloud 3D morphology can be advantageous for diagnosing discrepancies between simulations and observations. The methodology above computes radial profiles by taking the median across radial bins, a process which relies on the assumption of radial symmetry. If a cloud is highly asymmetric, taking the median of density values in a radial bin will not accurately capture cloud structure. Additionally, if the simulated clouds’ average radial density profiles for the total gas density are not matching observations, there is no clear method by which to diagnose potential differences between simulated clouds and observations. To provide a tool to diagnose visually why simulated clouds may not be consistent with

observations, the pipeline also outputs slices in each tracer and converts them into animations for visual analysis.

3.4. Fitting Data

Once we extract the cloud-averaged radial density profiles for the sample for all tracers ($n_{\text{H}_{\text{tot}}}$, n_{CO} , n_{H_2} , n_{H} , n_{H^+} , and T_{gas}),⁸ we perform single and multicomponent Gaussian fits to the density

⁸ The pipeline is also designed to run on the dust temperature data, T_{dust} . However, we do not consider T_{dust} for further analysis, since it represents the effective cooling from the dust.

tracers. We also perform sigmoid, as well as spline fits, to the gas temperature profiles.

We fit the gas temperature radial profile of each cloud to investigate the radial distance at which atomic gas is mostly CNM, mostly UNM, and mostly warm neutral medium (WNM). From Wolfire et al. (2003), atomic gas is in the CNM phase at 20–500 K and the WNM phase at 3000–8000 K. We therefore use the temperature profile fit to place markers at $T_{\text{gas}} \sim 250$ K and $T_{\text{gas}} \sim 3000$ K as the temperatures where atomic gas is “mostly” CNM and “mostly” WNM. We tested both sigmoid and spline fits to the temperature profiles. Due to cloud asymmetries, we found that the average temperature profile had large scatter and that sigmoid fits were incorrectly placing temperature markers. We opt to utilize the spline fit to determine more confidently which radial distance corresponds to $T_{\text{gas}} = 250$ K and $T_{\text{gas}} = 3000$ K in the temperature radial profile.

We fit the one- and two-component Gaussians to all density tracers. For the single-component Gaussian fit, our model is defined as:

$$n_X(r) = a \exp\left(\frac{-r^2}{2\sigma^2}\right), \quad (1)$$

where n_X is the volume density of tracer X , r is the radial distance (units of pc), a is the amplitude (peak density, units of cm^{-3}) and σ is the standard deviation (units of pc). The one-component fits were performed with a Levenberg–Marquardt least-squares fitter. For the two-component Gaussian fit, our model is defined as:

$$n_X(r) = a_1 \exp\left(\frac{-r^2}{2\sigma_1^2}\right) + a_2 \exp\left(\frac{-r^2}{2\sigma_2^2}\right), \quad (2)$$

where a_1 and σ_1 and a_2 and σ_2 are the amplitudes and standard deviations of the inner and outer Gaussians, respectively. The two-component fits were performed utilizing a sequential least-squares optimization algorithm. In Zucker et al. (2021) the first 2 pc are excluded from the fit since optical stellar photometric and astrometric measurements (the key ingredients in 3D dust mapping) become either scarce and/or unreliable in the core of clouds due to high levels of dust extinction. From the simulations we have access to core information without having to consider dust extinction. However, in order to ensure the best basis of comparison with observations we exclude the first 2 pc from our fits as well. The results stay consistent regardless if we include or exclude the inner 2 pc. We emphasize that the cloud average radial n_{H^+} profile is always a poor fit since ionized hydrogen has a radial profile with a nonzero center, which we do not take into account with our fits and is beyond the scope of this work.

We found that the choice to use a maximum number density of molecular gas and atomic gas to differentiate between molecular and atomic clouds was incorrectly flagging clouds. The atomic gas number density (n_{H_I}) in the cores of Cloud Factory–simulated clouds is on par with the molecular gas number density (n_{H_2} ; see the core density values in the Appendix, Figure A1). We decide to utilize the star-forming subset of simulated clouds as a better basis of comparison with observed clouds because all molecular clouds analyzed in Zucker et al. (2021) are known to be star forming.

The pipeline we describe here is publicly available and open source on GitHub. We include two versions of the pipeline: a Cloud Factory–specific version that loops over the Cloud Factory simulation grids, and a version that takes in a single 3D grid (observational or simulated) and performs the above analysis on whatever tracer the grid uses. Example notebooks and data are included so that simulators and observers can run a similar analysis on their own data.

4. Results

4.1. Ensemble Results

We applied the pipeline described in Section 3 to nine simulation subgrids with an appreciable amount of molecular gas and analyzed 125 clouds in total with 36 being flagged as star forming based on the presence of sink particles within their cloud masks. Out of the 36 star-forming clouds, only 24 had two-component Gaussian fits that successfully converged (see Section 4.2 for details; 11 clouds failed to converge with $\sigma_{1,2} \sim 0$ pc while one cloud was an outlier with $a_1 \geq 100 \text{ cm}^{-3}$). The ensemble results for the two-component Gaussian fits to the $n_{\text{H}_{\text{tot}}}$ distributions of the 24 simulated star-forming clouds are shown in Figure 4, alongside the observed solar neighborhood clouds from Zucker et al. (2021). The Gaussian peaks (a_1 and a_2) for simulated clouds (~ 50 and $\sim 8 \text{ cm}^{-3}$) agree mostly with observations (~ 30 and $\sim 9 \text{ cm}^{-3}$) by construction: recall that the same density threshold that was used to define observational skeletons was used to find clouds in the simulations. Simulated clouds deviate from observations for the remainder of the two-component Gaussian fit parameters. The most striking difference is the inner and outer Gaussian widths (σ_1 and σ_2 , respectively) for simulated star-forming clouds and observations. The widths for both the inner and outer Gaussians of the simulated star-forming clouds have lower values than observations. The Gaussian widths (σ_1 and σ_2) respectively extend out to ~ 3 and ~ 11 pc in the observations but only extend out to ~ 2 and ~ 5 pc in the simulations.

Ensemble $n_{\text{H}_{\text{tot}}}$ two-component and one-component Gaussian fits, as well as cloud length, mass, and n_{CO} , n_{H_2} , and n_{H_I} one-component Gaussian fits are highlighted in Table 1. Modeling uncertainties in the simulated cloud radial profile fits is difficult; there are no intrinsic uncertainties stemming from the simulations and any “uncertainty” in the reported fits would be due to the variations in the profiles on a slice-by-slice basis. We instead opt to report estimates of the population-level spread in the distribution of each reported value in Table 2.

The ensemble two-component Gaussian fit results are the first indication that the simulated clouds are not aligning with observational expectations. In particular, the width of the first Gaussian σ_1 indicates that the inner peaks of the simulated cloud radial profiles are thinner than the observed radial profile peaks, and the width of the second Gaussian σ_2 indicates that the tails of the simulated cloud radial profiles are falling off much faster than the observed radial profile tails. In order to diagnose the discrepancy between simulations and observations, we analyze the radial profiles and morphologies of specific clouds in detail.

4.2. Individual Cloud Examples

As a basis for comparison with the observational data, we apply our pipeline to the local Perseus molecular cloud, whose

Table 1
Properties of the Simulated Star-forming Clouds

| Grid # | Cloud # | Length | Mass | # Sinks | n_{Htot} | | n_{CO} | | n_{H_2} | | n_{H_I} | | | | | |
|-----------------|---------|--------|---------|---------|-------------------|------------|-----------------|------------|------------------|----------|------------------|----------|-------|----------|------|------|
| | | | | | a_1 | σ_1 | a_2 | σ_2 | a | σ | a | σ | a | σ | | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) |
| x1465y1390z1400 | 0 | 545 | 322,558 | 77 | 54.0 | 2.1 | 6.6 | 6.5 | 50.0 | 2.7 | 1.9e-06 | 1.0 | 21.0 | 2.4 | 10.0 | 1.0 |
| x1465y1390z1400 | 3 | 87 | 18,735 | 9 | 54.0 | 1.0 | 13.0 | 3.6 | 16.0 | 3.3 | 2.9e-09 | 1.7 | 4.6 | 3.2 | 7.2 | 1.7 |
| x1465y1390z1400 | 4 | 225 | 37,483 | 15 | 69.0 | 1.1 | 18.0 | 3.9 | 27.0 | 3.2 | 1.3e-07 | 1.5 | 12.0 | 3.0 | 2.4 | 1.5 |
| x1465y1390z1400 | 12 | 191 | 25,118 | 1 | 42.0 | 1.9 | 3.2 | 5.9 | 38.0 | 2.2 | 6.1e-09 | 1.2 | 3.8 | 1.8 | 31.0 | 1.2 |
| x1470y1400z1395 | 4 | 76 | 5812 | 1 | 23.0 | 2.6 | 9.5 | 2.6 | 33.0 | 2.6 | 3.6e-09 | 1.6 | 10.0 | 2.2 | 19.0 | 1.6 |
| x1465y1405z1395 | 8 | 169 | 63,898 | 3 | 58.0 | 2.0 | 3.1 | 7.1 | 54.0 | 2.2 | 2.4e-08 | 1.2 | 3.8 | 1.7 | 47.0 | 1.2 |
| x1475y1390z1395 | 0 | 46 | 6882 | 5 | 72.0 | 1.3 | 24.0 | 3.6 | 45.0 | 2.7 | 5.7e-06 | 0.6 | 1.8 | 1.5 | 41.0 | 0.6 |
| x1475y1390z1395 | 2 | 101 | 19,255 | 2 | 79.0 | 1.5 | 19.0 | 4.9 | 44.0 | 3.1 | 6.6e-07 | 1.0 | 12.0 | 2.3 | 24.0 | 1.0 |
| x1475y1390z1395 | 4 | 313 | 75,488 | 3 | 32.0 | 2.5 | 13.0 | 5.9 | 36.0 | 3.7 | 4.4e-07 | 1.1 | 7.2 | 3.3 | 17.0 | 1.1 |
| x1475y1390z1395 | 8 | 37 | 5176 | 1 | 30.0 | 2.0 | 1.7 | 7.1 | 28.0 | 2.3 | 5.1e-09 | 1.7 | 12.0 | 2.0 | 7.0 | 1.7 |
| x1475y1390z1395 | 18 | 18 | 48,825 | 19 | 57.0 | 1.7 | 9.0 | 4.7 | 44.0 | 2.4 | 9.8e-04 | 0.6 | 21.0 | 2.2 | 3.3 | 0.6 |
| x1475y1390z1395 | 22 | 122 | 19,959 | 8 | 54.0 | 1.4 | 12.0 | 3.8 | 49.0 | 1.9 | 2.2e-04 | 0.5 | 540.0 | 0.7 | 19.0 | 0.5 |
| x1465y1400z1400 | 3 | 33 | 15,656 | 5 | 86.0 | 2.1 | 9.8 | 5.9 | 80.0 | 2.6 | 5.8e-03 | 1.7 | 0.0 | 2.2 | 80.0 | 1.7 |
| x1465y1400z1400 | 13 | 43 | 3039 | 1 | 48.0 | 1.2 | 14.0 | 2.9 | 29.0 | 2.2 | 6.5e-08 | 0.6 | 3.3 | 1.1 | 23.0 | 0.6 |
| x1465y1400z1400 | 18 | 69 | 10,210 | 3 | 40.0 | 1.9 | 3.3 | 6.3 | 35.0 | 2.3 | 4.6e-09 | 1.4 | 13.0 | 1.8 | 16.0 | 1.4 |
| x1465y1400z1400 | 23 | 125 | 23,166 | 19 | 36.0 | 1.6 | 6.4 | 4.7 | 26.0 | 2.5 | 6.0e-08 | 1.1 | 13.0 | 2.4 | 0.7 | 1.1 |
| x1470y1400z1400 | 1 | 36 | 2294 | 1 | 49.0 | 1.7 | 6.1 | 1.7 | 55.0 | 1.7 | 1.8e-07 | 0.8 | 34.0 | 1.3 | 10.0 | 0.8 |
| x1465y1405z1400 | 0 | 207 | 47,893 | 8 | 78.0 | 1.4 | 9.3 | 4.4 | 47.0 | 2.2 | 3.0e-07 | 0.8 | 9.5 | 1.9 | 26.0 | 0.8 |
| x1465y1405z1400 | 4 | 128 | 33,625 | 2 | 69.0 | 1.9 | 6.3 | 8.7 | 56.0 | 2.5 | 2.1e-07 | 1.8 | 12.0 | 2.0 | 31.0 | 1.8 |
| x1465y1405z1400 | 8 | 21 | 1306 | 1 | 45.0 | 1.7 | 0.9 | 6.5 | 43.0 | 1.8 | 2.1e-08 | 1.0 | 17.0 | 1.8 | 4.5 | 1.0 |
| x1465y1405z1400 | 14 | 48 | 37,131 | 7 | 27.0 | 3.2 | 9.3 | 3.2 | 36.0 | 3.2 | 1.1e-06 | 1.2 | 6.3 | 3.2 | 21.0 | 1.2 |
| x1465y1405z1400 | 18 | 34 | 4622 | 1 | 15.0 | 1.2 | 6.1 | 5.0 | 7.4 | 4.5 | 3.8e-10 | 2.0 | 1.1 | 3.8 | 4.6 | 2.0 |
| x1465y1405z1400 | 28 | 41 | 4239 | 3 | 36.0 | 1.7 | 5.7 | 5.7 | 26.0 | 2.6 | 3.2e-09 | 1.4 | 8.1 | 2.3 | 8.9 | 1.4 |
| x1475y1390z1400 | 5 | 96 | 5630 | 1 | 30.0 | 2.1 | 5.3 | 4.2 | 30.0 | 2.5 | 9.5e-09 | 1.1 | 6.2 | 2.4 | 19.0 | 1.1 |

Note. Skeletonizing and analysis pipeline results for star-forming clouds in the Cloud Factory feedback-dominated suite of simulations. (1) Grid number identifier. (2) Cloud number identifier (each skeletonized cloud in one simulation grid is given an identifying number). (3) Length and (4) mass of each cloud. (5) Number of star-forming sinks in each cloud. Two-component Gaussian fit (amplitudes a_1 and a_2 and standard deviations σ_1 and σ_2) results to the column (6–9) n_{Htot} radial profiles. One-component Gaussian fit (amplitude a and standard deviation σ) results to the column (10–11) n_{Htot} , (12–13) n_{CO} , (14–15) n_{H_2} , and (16–17) n_{H_I} radial profiles.

Table 2
Population Statistics

| Type | Length (pc) | Mass (M_{\odot}) | n_{Htot} | | | | | | n_{CO} | | n_{H_2} | | n_{H_I} | |
|----------------------------|-------------------------|---------------------------|-------------------------------|---------------------|-------------------------------|----------------------|-----------------------------|---------------------|--------------------------------|---------------------|-----------------------------|---------------------|-----------------------------|---------------------|
| | | | a_1 (cm^{-3}) | σ_1 (pc) | a_2 (cm^{-3}) | σ_2 (pc) | a (cm^{-3}) | σ (pc) | a (cm^{-3}) | σ (pc) | a (cm^{-3}) | σ (pc) | a (cm^{-3}) | σ (pc) |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) |
| Simulated Star-forming | $81.5^{+114.6}_{-46.1}$ | $1.9e4^{+2.9e4}_{-1.4e4}$ | $48.5^{+21.5}_{-18.5}$ | $1.7^{+0.4}_{-0.4}$ | $7.8^{+5.5}_{-4.5}$ | $4.8^{+1.7}_{-1.3}$ | $37.0^{+14.3}_{-10.3}$ | $2.5^{+0.7}_{-0.3}$ | $9.8e - 8^{+3e-6}_{-9.3e-8}$ | $1.2^{+0.6}_{-0.4}$ | $9.8^{+8.5}_{-6.1}$ | $2.2^{+0.9}_{-0.6}$ | $18.0^{+13.0}_{-13.4}$ | $1.2^{+0.6}_{-0.4}$ |
| Simulated Non-star-forming | $36.0^{+45.0}_{-16.0}$ | $2.3e3^{+6.2e3}_{-1.3e3}$ | $38.0^{+18.0}_{-13.0}$ | $1.8^{+0.6}_{-0.5}$ | $9.2^{+6.6}_{-5.6}$ | $4.3^{+2.8}_{-1.6}$ | $34.0^{+14.0}_{-11.8}$ | $2.7^{+0.5}_{-0.6}$ | $5.5e - 9^{+8.1e-8}_{-5.1e-9}$ | $1.1^{+0.6}_{-0.4}$ | $3.3^{+7.5}_{-2.5}$ | $1.9^{+0.6}_{-0.5}$ | $27.0^{+9.5}_{-12.7}$ | $1.1^{+0.6}_{-0.4}$ |
| Observations | 62.0^{+14}_{-20} | $9.4e3^{+6.4e3}_{-6.4e3}$ | $31.3^{+10.9}_{-2.0}$ | $2.9^{+1.2}_{-0.3}$ | $9.2^{+2.4}_{-0.4}$ | $10.9^{+2.6}_{-1.7}$ | $30.8^{+10.2}_{-1.7}$ | $5.7^{+0.4}_{-1.3}$ | ... | ... | ... | ... | ... | ... |

Note. Summary table of the ensemble median cloud properties and their dispersions for the simulated star-forming and non-star-forming populations compared with observations. Each cell displays the median value (50th percentile) with the lower (50th percentile – 16th percentile) and upper (84th percentile – 50th percentile) bounds. (1) Simulation type. (2) Length and (3) mass of the clouds. Two-component Gaussian fit (amplitudes a_1 and a_2 and standard deviations σ_1 and σ_2) results to the column (4–7) n_{Htot} radial profiles. One-component Gaussian fit (amplitude a and standard deviation σ) results to the column (8–9) n_{Htot} , (10–11) n_{CO} , (12–13) n_{H_2} , and (14–15) n_{H_I} radial profiles.



Figure 4. Corner plot comparing the two-component Gaussian fits of observed clouds and simulated star-forming clouds. The diagonal shows a smoothed version of histograms comparing the two sets of clouds. The cloud peak densities (a_1 and a_2) are largely consistent across the two data sets. However, the cloud widths (σ_1 and σ_2) show more significant discrepancies.

radial density profile based on the 3D dust data from Leike et al. (2020) was originally analyzed in Zucker et al. (2021). Figure 5 displays our pipeline’s results for Perseus. The two radial profile plots display the two-component Gaussian and single-component Gaussian fit to Perseus’ radial profile. The gray points in the figure represent the slice-by-slice radial profiles that are averaged to compute the cloud-averaged radial volume density profile, represented by blue points since there is no temperature information present in the observational data. We also display multiple cross sections of the n_{Htot} density of Perseus taken along the skeleton. Our pipeline reproduces a two-component Gaussian fit that traces the data better than the single-component fit. Discrepancies between our fitted values and the values found in Zucker et al. (2021) come from slightly different modeling choices and that our pipeline does not take into account errors intrinsic to the observations (see Section 5).

Figures 6, 7, and 8 display three exemplary simulated star-forming clouds. The top panel displays the cloud’s location in the simulation grid, the skeleton of the cloud, and the number and location of star-forming sinks within the cloud. The bottom panel displays the one- and two-component Gaussian fits to the n_{Htot} radial profiles, where the temperature is displayed as a color bar and dotted lines demarcate the radial distances

corresponding to potential atomic gas thermal transition. Recall that the n_{Htot} radial profile is the only information we can directly compare with the observational data, as there is no observational information on the relative contribution from atomic versus molecular hydrogen gas. For the simulations we additionally include the radial profiles for molecular ($2 \times n_{\text{H}_2}$) and atomic n_{H} gas densities alongside n_{Htot} . We also include two representative slices taken halfway through the cloud’s skeletal length. These slices display the total gas density n_{Htot} and the gas temperature morphology.

The three exemplary fits display a wide variety of morphologies. Figure 6 shows a sheet-like cloud, Figure 7 shows a radially asymmetric cloud, and Figure 8 shows a cloud that is approximately radially symmetric. However these three fits display a common theme found in all clouds in the simulation: in general, the two-component Gaussian fit does not trace the data better than a single-component fit. Additionally, consistent with the small outer Gaussian width (σ_2) common across the entire ensemble of simulated clouds found in Figure 4, the radial profiles for the example simulated clouds fall off to $\sim 0 \text{ cm}^{-3}$ more quickly than Perseus’ profile (Figure 5). This trend is even more evident in the representative slices, which all display a lack of an extended gas envelope.

Perseus from Zucker et al. 2021

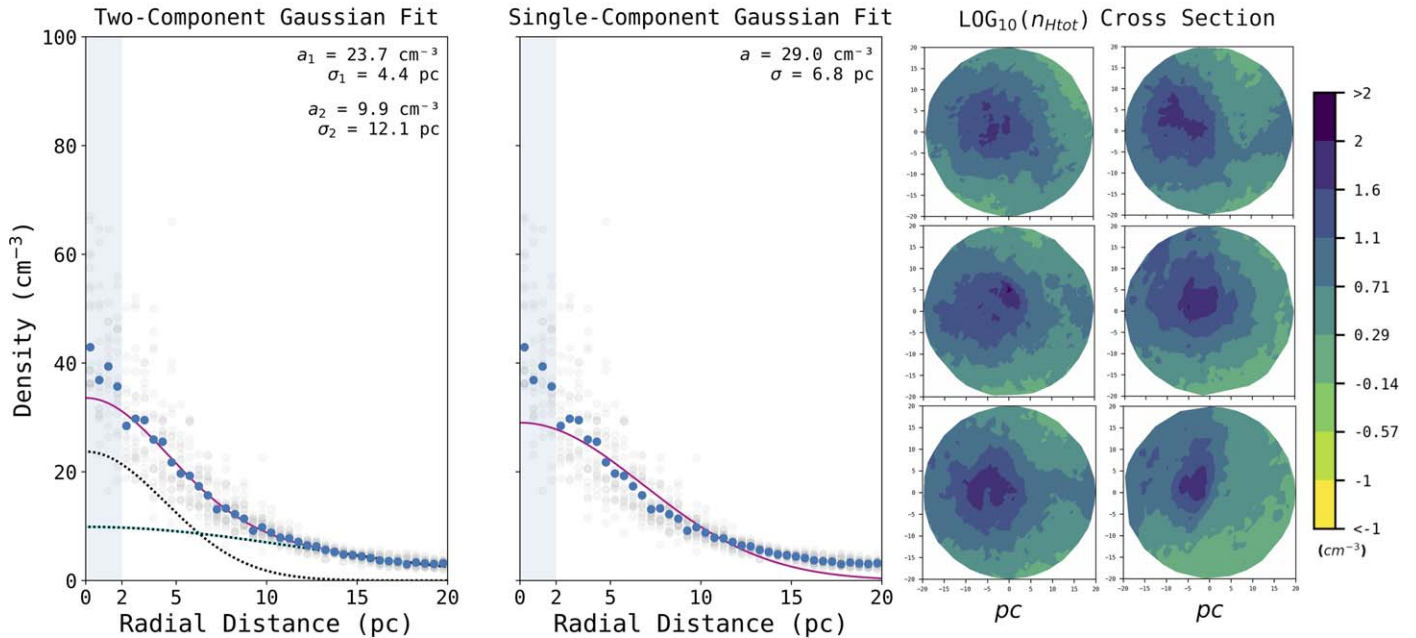


Figure 5. Cloud-averaged radial volume density (n_{Htot}) Gaussian fits for Perseus for the two-component (left panel) and single-component (middle panel) fits, shown alongside the density slices taken at six different positions along the skeleton (right panel). In the left and middle panels, the blue dots display the averaged radial profile, while the gray points indicate the variation in the radial profiles from slice to slice. The red line shows the best-fit function, which for the two-component Gaussian consists of inner (black dotted line) and outer (cyan dotted line) components. The inner 2 pc, where the 3D dust observations are unreliable, have been grayed out. The best-fit values for the two-component Gaussian fit (the inner and outer Gaussian widths σ_1 and σ_2 , respectively, and their corresponding amplitudes a_1 and a_2) and the single-component Gaussian fit (the width σ and amplitude a) are summarized at the top. As can be seen in the right panel, the slices display radial symmetry as well as an extended gas envelope.

We find that the lack of an extended gas envelope is a common feature across the entire sample of clouds.

While simulated non-star-forming clouds cannot be directly used to disentangle the precise thermal and chemical phase structure of observed star-forming clouds analyzed in Zucker et al. (2021), they can elucidate properties that are common within the entire simulation grid. The ensemble results including the entire sample size are shown in Table 2 and in Figure A2 in the Appendix. The majority of non-star-forming clouds displays similar properties to the star-forming ones for all Gaussian fit parameters. In particular, the Gaussian widths (σ_1 and σ_2), extend out to ~ 2 and ~ 5 pc, respectively, for the majority of the simulated clouds. A small subset of non-star-forming clouds do display larger σ_2 widths (≥ 10 pc). These clouds, while not star forming, tend to be found in very dense regions of the simulation grid and provide an interesting subset of clouds that we defer to future investigation.

5. Discussion

From the slice visuals in the three exemplary fits, we determine that the poor two-component Gaussian fits to the n_{Htot} radial profiles are due to radial asymmetry and the lack of an extended envelope. All three figures (Figures 6, 7, and 8) display the lack of an extended envelope and Figures 6 and 7 display radial asymmetry. The lack of an extended envelope presents itself as a rapid falloff to $\sim 0 \text{ cm}^{-3}$ at ~ 10 pc in the cloud average radial volume density profiles. Radial asymmetry presents itself as a larger scatter of gray points in Figures 6 and 7, with the gray points being representative of the individual slice-by-slice radial profiles that are averaged to compute the cloud-averaged radial volume density profile.

We turn to a more extreme example of asymmetry in a simulated star-forming cloud in Figure 9. This cloud displays radial asymmetry and lacks an envelope to such an extreme extent that the two-component Gaussian fit fails to converge. Asymmetry in this case causes the radial profile to take on a more exponential shape than a Gaussian profile. We can compare the density slice of this cloud to the density slices for Perseus. Perseus' density slices display approximate radial symmetry and a gas envelope that extends past 20 pc. The $\log(T_{\text{gas}})$ slices map onto the $\log(n_{\text{Htot}})$ slices quite well and we identify a temperature front that can exceed $\sim 10^6$ K mapping onto regions of the cloud with $\sim 0.001 \text{ cm}^{-3}$ density. In the Cloud Factory simulations there is a constant UV background that acts as a source of background heating. The resultant temperature field is driven by density variations where lower-density gas is hotter due to being less shielded from the ISRF. This example is in agreement with the asymmetric cloud in Figure 7. This cloud also displays a temperature front that maps onto the density structure, albeit a much colder one due to the skeleton tracing a part of the cloud that is being shielded. Similar temperature fronts map to regions of lower density in both Figures 6 and 8.

The lack of an extended gas envelope could be due to two factors: (1) too strong supernova feedback in comparison to the solar neighborhood or (2) a lack of magnetic fields, which would help keep clouds together (see, e.g., the discussion in Ganguly et al. 2023). Supernova feedback injects both thermal energy and momentum into the ISM and could be the cause of the low-density envelopes, and therefore the low σ_2 values. To probe for a potential cause we repeated the experiment on one grid of the potential-dominated suite of the Cloud Factory

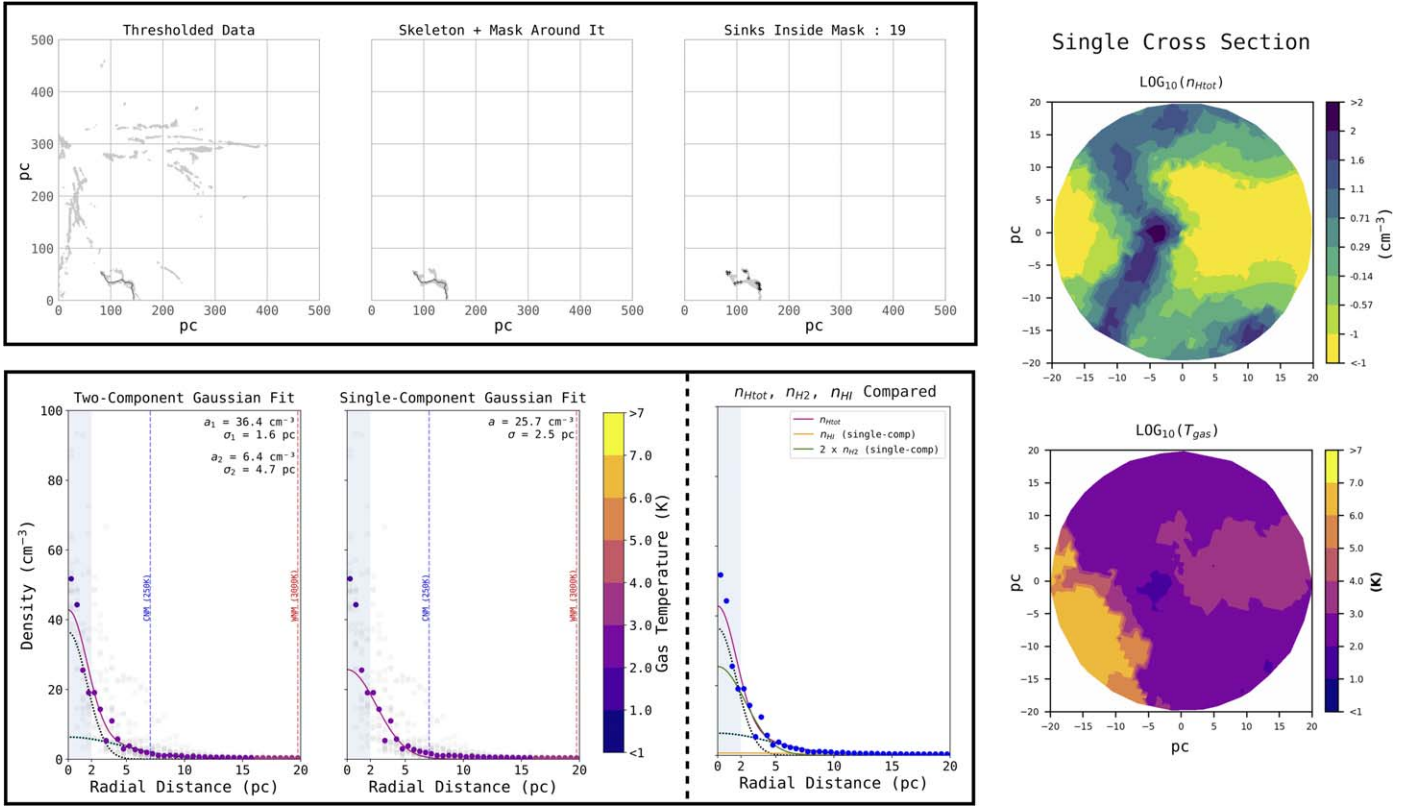


Figure 6. Sheet-like simulated cloud (grid x1465y1400z1400, cloud 23) with 19 star-forming sinks. In the top panel, we show the projected density field thresholded at a level of $n_{\text{Htot}} \geq 35 \text{ cm}^{-3}$ (top left panel) alongside the cloud’s mask and skeleton (top middle panel) and the locations of star-forming sinks (crossed points) within the cloud (top right panel). The bottom panel displays the cloud-averaged radial volume density (n_{Htot}) Gaussian fits to the simulated cloud with two-component (bottom, left panel) and single-component (bottom, middle panel) fits. In the bottom left and bottom middle panels, the colored points display the averaged radial profile with the average gas temperature at that radial distance (color bar), while the gray points indicate the variation in the radial profiles from slice to slice. The red line shows the best-fit function, which for the two-component Gaussian consists of inner (black dotted line) and outer (cyan dotted line) components. The inner 2 pc, where the 3D dust observations (see Figure 5) are unreliable, have been grayed out. The vertical dotted lines indicate temperature transitions from CNM to UNM ($\sim 250 \text{ K}$, blue dotted line) and UNM to WNM ($\sim 3000 \text{ K}$, red dotted line). The bottom right panel shows the n_{Htot} two-component Gaussian fits (with the inner and outer components) compared to the single-component Gaussian fits for n_{H_2} (orange) and n_{H_I} (green). The far right panels show representative density and temperature slices taken midway across the skeleton. As can be seen in the right panel, the slice morphology is much different than that found in Figure 5.

simulations, where supernova feedback is less prominent and the large-scale gravitational potential largely dictates the gas dynamics. The resultant cloud profiles were predominately sheet-like and displayed the same lack of extended gas envelopes. We display 2D representations of feedback-dominated and potential-dominated grids, as well as six different cloud morphologies from both grids, in Figure A3 in the Appendix. This result gives credence to magnetic fields being responsible for the lack of extended envelopes; however we leave a more detailed comparison between feedback and potential-dominated grids to future studies.

Our finding is in agreement with the work performed by Ganguly et al. (2023). Ganguly et al. (2023) utilized SILCC-Zoom simulations (Seifried et al. 2017), which include magnetic fields, self-gravity, supernova feedback, and nonequilibrium chemistry to study the effect of magnetic fields on molecular clouds. They found that the molecular clouds in their subset of simulations with magnetic fields contained more mass in their diffuse envelopes than clouds without magnetic fields. They additionally found that magnetic fields as a whole are more important for less dense structures. This result provides evidence that the lack of magnetic fields in the Cloud Factory suite may be driving the lack of extended envelopes in the simulated clouds.

To summarize: clouds in the feedback-dominated suite of the Cloud Factory simulations display a lack of extended envelope, a finding also shared by a subset of clouds in the potential-dominated suite. Star-forming clouds in the feedback-dominated case can be radially asymmetric filaments (Figures 7 and 9), symmetric filaments (Figure 8), or sheet-like (Figure 6) with a majority of the clouds being radially asymmetric. Star-forming clouds in the potential-dominated case are solely sheet-like. We propose that the lack of magnetic fields in combination with the constant UV background radiation is driving the lack of low-density envelopes in both simulation suites. That is, the lack of envelopes in both suites is due to clouds not being able to hold themselves together against the constant UV radiation field, thereby destroying the extended low-density envelopes. Additionally we argue that a combination of the galactic potential and supernova feedback is driving the cloud morphologies; the radially asymmetric profiles in the feedback-dominated suite is due to supernova feedback dominating, while sheet-like structures is due to potential forces dominating.

Our results are independent of modeling choices. We revisited the choice of restricting our analysis to clouds with lengths greater than 20 pc (Section 3) by rerunning the pipeline with the opposite criterion (restricting the analysis to clouds less than 20 pc) for a single feedback-dominated grid. We find

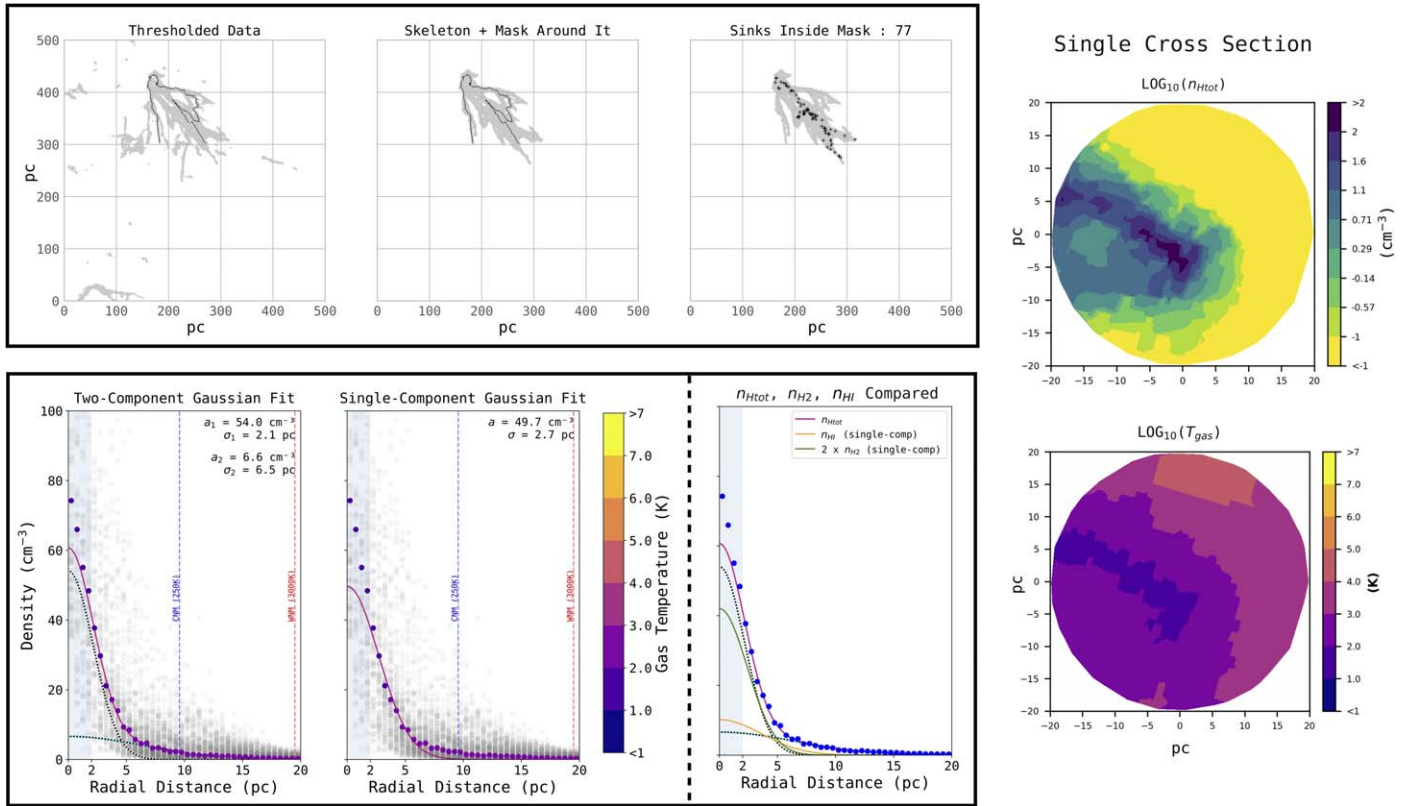


Figure 7. Same as Figure 6, but for an asymmetric cloud with 77 star-forming sinks. (Grid x1465y1390z1400, cloud 0.)

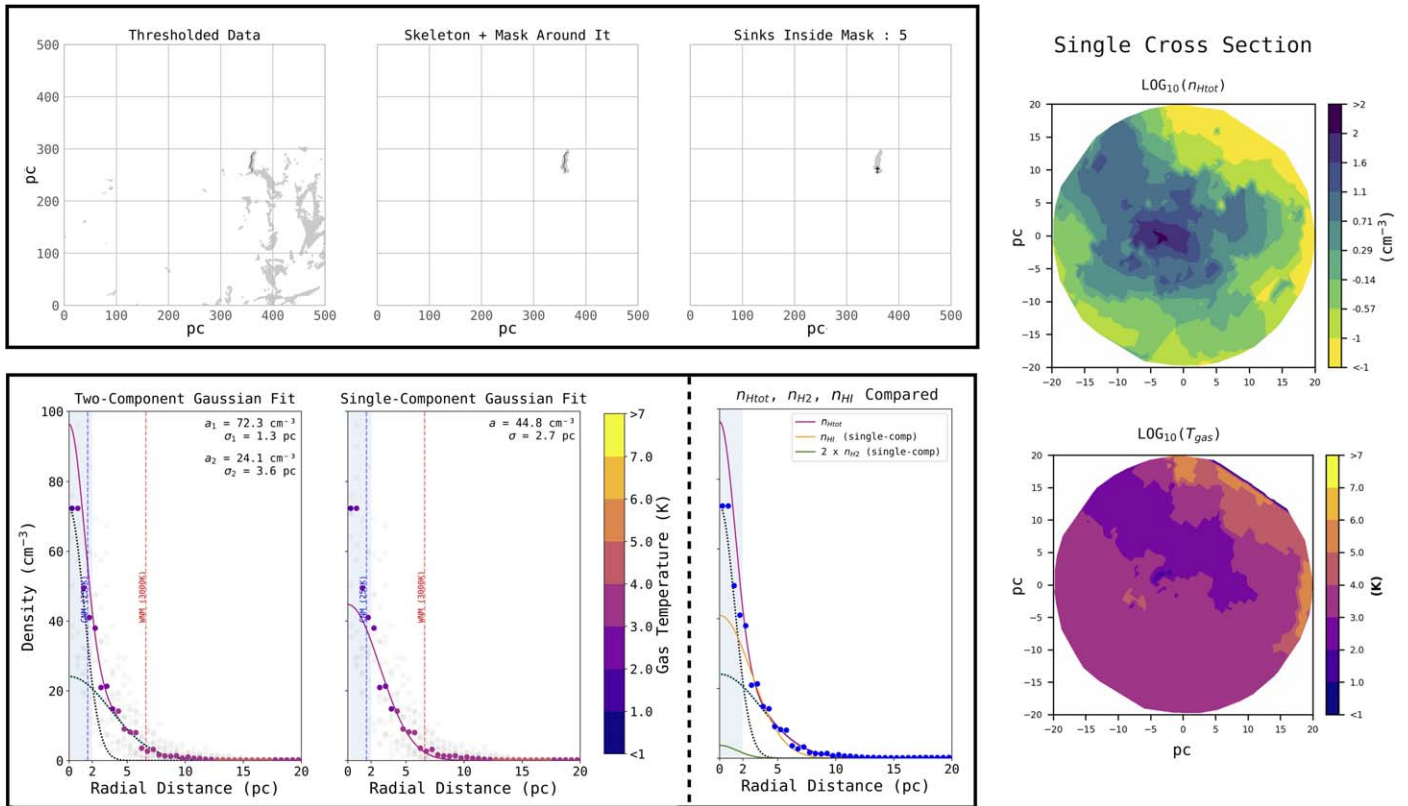


Figure 8. Same as Figure 6, but for a symmetric cloud with five star-forming sinks. (Grid x1475y1390z1395, cloud 0.)

that these small simulated clouds have similar morphologies to the longer clouds but—due to their radial profiles only having 20 slices—the shorter clouds displayed significantly more

scatter in their averaged profile. We additionally analyzed the effect of including a “shift” of the cloud center from the topological center to the densest region within the cloud’s

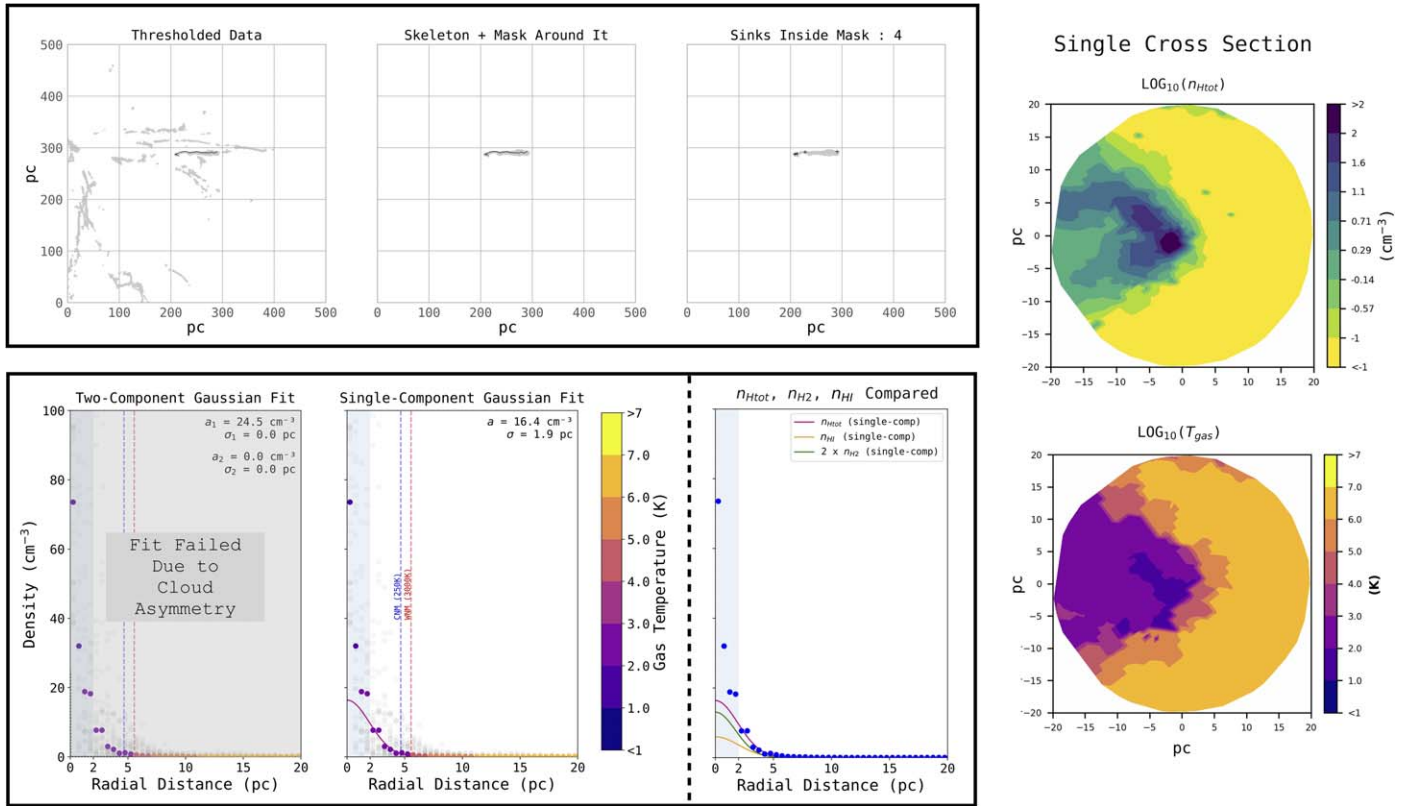


Figure 9. Same as Figure 6, but for an asymmetric cloud with four star-forming sinks. The two-component Gaussian fit failed due to an exponential-like radial profile. This failure likely stems from the assumption of radial symmetry in the derivation of the averaged radial profile. (Grid x14651400z1400, cloud 8).

mask, ensuring that the density peaks at a distance of 0 pc. This “shift” was explored in both Zucker & Chen (2018) and Zucker et al. (2021). Including this shift does not impact our main result, as the effect of slice-by-slice differences and cloud morphology impacts the resultant radial profile more profoundly. We note that shifting remains an option in the pipeline, and can be pertinent when specifically analyzing peak densities and width values in radial profiles of observations and simulations.

We finally revisit our assumption of radial symmetry when extracting radial profiles from each 20 pc slice. Due to the diversity of cloud morphologies found in the simulations, assuming radial symmetry is not appropriate for asymmetric and sheet-like clouds and new methods must be utilized to examine the density structure of these clouds. Departing from radial symmetry is outside of the scope of this work, and is why the pipeline provides slices as a tool to analyze whether radial symmetry is a good assumption or not.

In comparison to this work, Zucker et al. (2021) performed a Bayesian model comparison to show that a two-component Gaussian fit was preferred over other models, including a single-component Gaussian fit. Given that the simulations do not match observations, a similar Bayesian model comparison for the simulated cloud radial profiles is beyond the scope of this work. However, we note that the pipeline we present here should have broad potential applications, and more sophisticated Bayesian modeling of the extracted radial profiles can be applied to future simulations that include additional physics.

As shown in Figures 6, 7, 8, 9, and Appendix Figure A1, the total gas density, chemical composition, and temperature radial profile slices map onto each other quite well. In the future, if

simulated clouds better reproduce the extended envelopes seen in nearby observed clouds, radial profile slices of both temperature and other density tracers can be better inferred from the total gas density profiles. We would then be able to extrapolate additional information of the observed clouds from their total gas density, and test the radial distances at which local molecular filamentary clouds may exhibit thermal and/or chemical phase transitions.

We turn to a final example in Figure 10. This cloud has a symmetric radial morphology and contains no star-forming sinks. This cloud is still noteworthy within the broader sample since even though it is not star forming, it contains more molecular gas in the inner few parsecs than atomic gas. This cloud could be in the precursor stages of star formation. Tracking the evolution of morphology and chemical composition of clouds like these over time can elucidate the life cycle of a molecular cloud as it transitions from non-star forming to star forming, which we defer to future work.

6. Conclusion

We present a technique to extract and analyze clouds in resolved simulations of the ISM for the purpose of comparing to 3D observational data. We first extend the cloud identification and skeletonization methodology presented in Zucker et al. (2021) to search simulation grids for cloud-like structures. The radial profiles of clouds along their topological skeletons are then computed for multiple tracers. Density radial profiles are fit using one-component and two-component Gaussians and temperature radial profiles are fit using sigmoid and spline

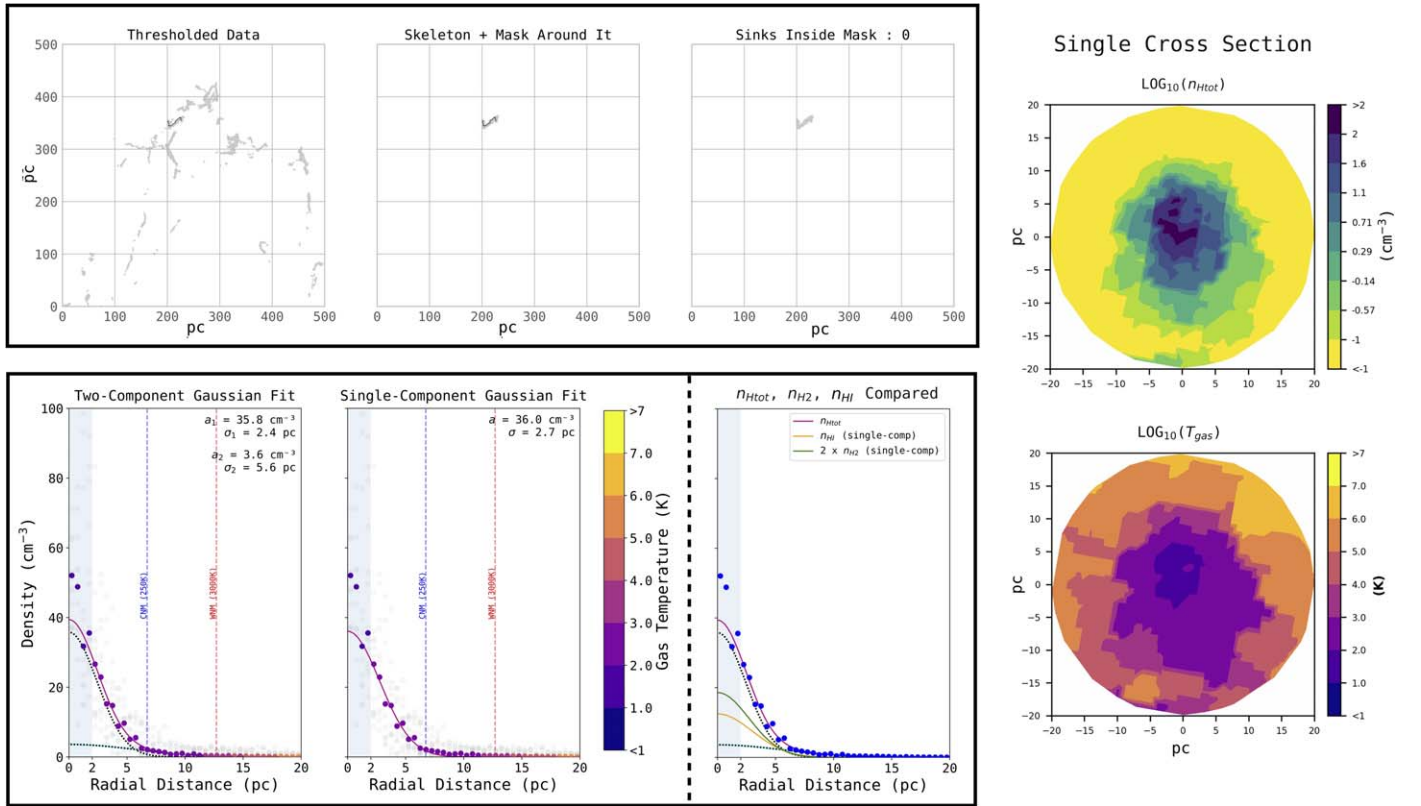


Figure 10. Same as Figure 6, but for a symmetric cloud with no star-forming sinks. Due to the high density of molecular gas, this cloud could be in the precursor stages of star formation. (Grid x14701400z1395, cloud 3.)

profiles. 2D slices along the skeleton are also computed to serve as a visualization of cloud morphology.

We apply the pipeline to the Cloud Factory feedback-dominated suite of simulations to extract radial profiles of two dozen star-forming clouds. We find that the simulated clouds are primarily radially asymmetric and lack low-density envelopes. In line with recent work (e.g., Ganguly et al. 2023), we propose that the lack of low-density envelopes is due to the lack of magnetic fields. We additionally propose that the radially asymmetric morphologies are due to supernova feedback. Initial tests indicate that clouds in the Cloud Factory potential-dominated suite of simulations are primarily sheet-like and also lack low-density envelopes, giving credence to the common lack of magnetic fields in both suites driving the lack of low-density envelopes and supernova feedback is driving the radially asymmetric morphologies. Future work will investigate the differences between the feedback- and potential-dominated suites in more detail, and compare the Cloud Factory clouds to clouds in other simulations.

New observational techniques that probe the 3D structure of local star-forming clouds provide an exciting new avenue by which simulations can be used to detangle the physical conditions occurring in observed clouds. We show that there is a current discrepancy between the 3D structure of clouds in hydrodynamic simulations and observations. Future iterations of the Cloud Factory suite will include magnetic fields (as in, e.g., Ganguly et al. 2023) and track cloud formation and evolution over time. On the observational front, there are new 3D dust maps being developed (e.g., Edenhofer et al. 2023) that extend to larger distances. These new 3D dust maps will allow for opportunities to

characterize cloud morphologies over a more diverse range of galactic environments. The cloud identification and analysis pipeline we present here provides a tool for future exploration of simulated and observed 3D cloud structure and is publicly available on Zenodo doi:10.5281/zenodo.10157333 and <https://github.com/elijah-mullens/Skeletonizing-and-Analyzing-Pipeline-for-3D-Interstellar-Cloud-Ensembles> GitHub (Skeletonizing and Analyzing Pipeline for 3D Interstellar Cloud Ensembles).

Acknowledgments

We acknowledge support from Space Telescope Science Institute’s 2022 Space Astronomy Summer Program. We thank STScI’s continued support of this project by funding two in-person poster presentations at the 241st meeting of the American Astronomical Society and Protostars and Planets VII. E.M. acknowledges that this material is based upon work supported by the National Science Foundation Graduate Research Fellowship under grant No. 2139899. We thank the anonymous referee for a very constructive and thorough report of our manuscript that significantly improved the quality of this work.

Appendix

Figure A1 shows example slices from multiple tracers. We found in Section 3.4 that the choice to use maximum number density of molecular gas and atomic gas to differentiate clouds was incorrectly flagging clouds due to atomic gas being on par with molecular gas density in the cores of clouds. We therefore utilized star-forming clouds in our analysis. This figure also shows that the total gas density maps onto the chemical

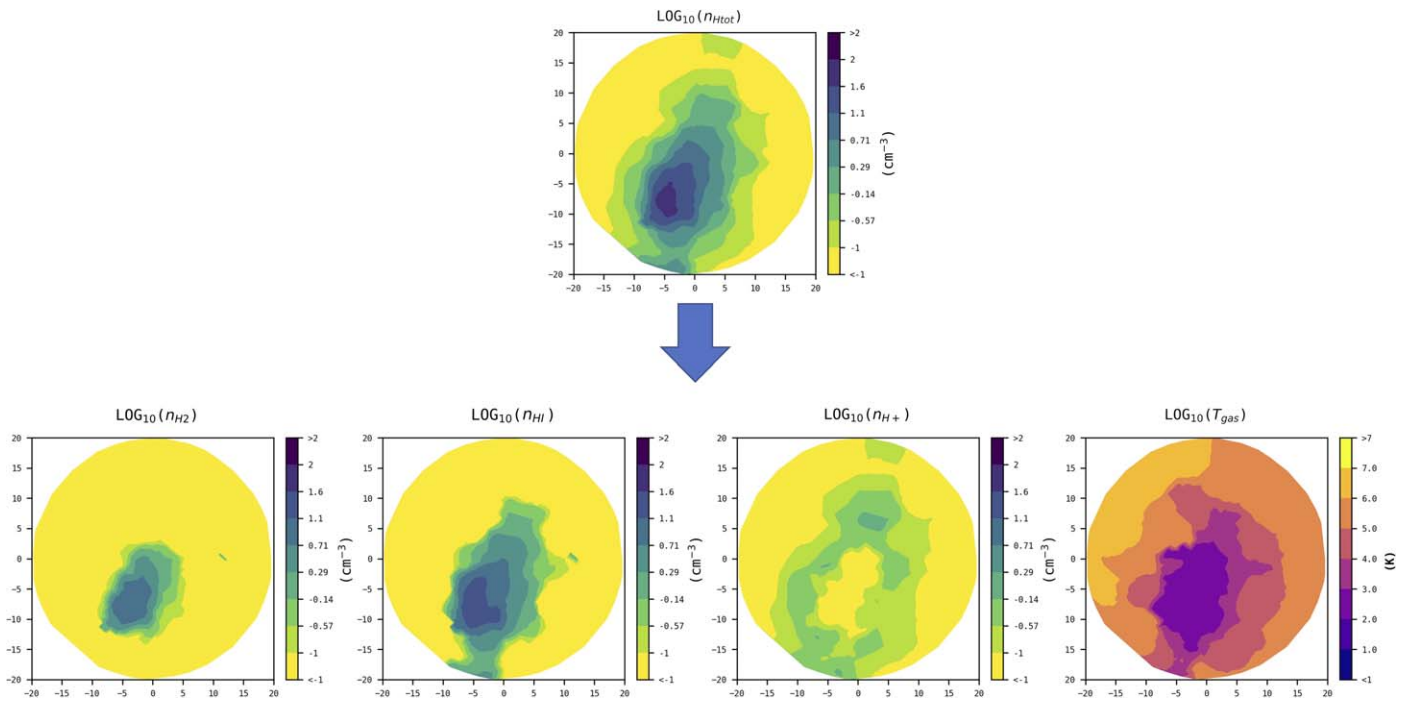


Figure A1. Example slices taken from multiple tracers ($n_{H_{tot}}$, n_{H_2} , n_{H_I} , n_{H^+} , and T_{gas}) midway along a skeleton. Once simulations can more faithfully reproduce the extended structure of gaseous envelopes observed in the solar neighborhood, we should be able to map these tracers on to the total hydrogen volume density to determine the radial distances at which chemical and/or phase transitions occur.

$n_{H_{tot}}$ Two-Component Gaussian Fits - Zucker 2021
All Simulated Clouds and Observations

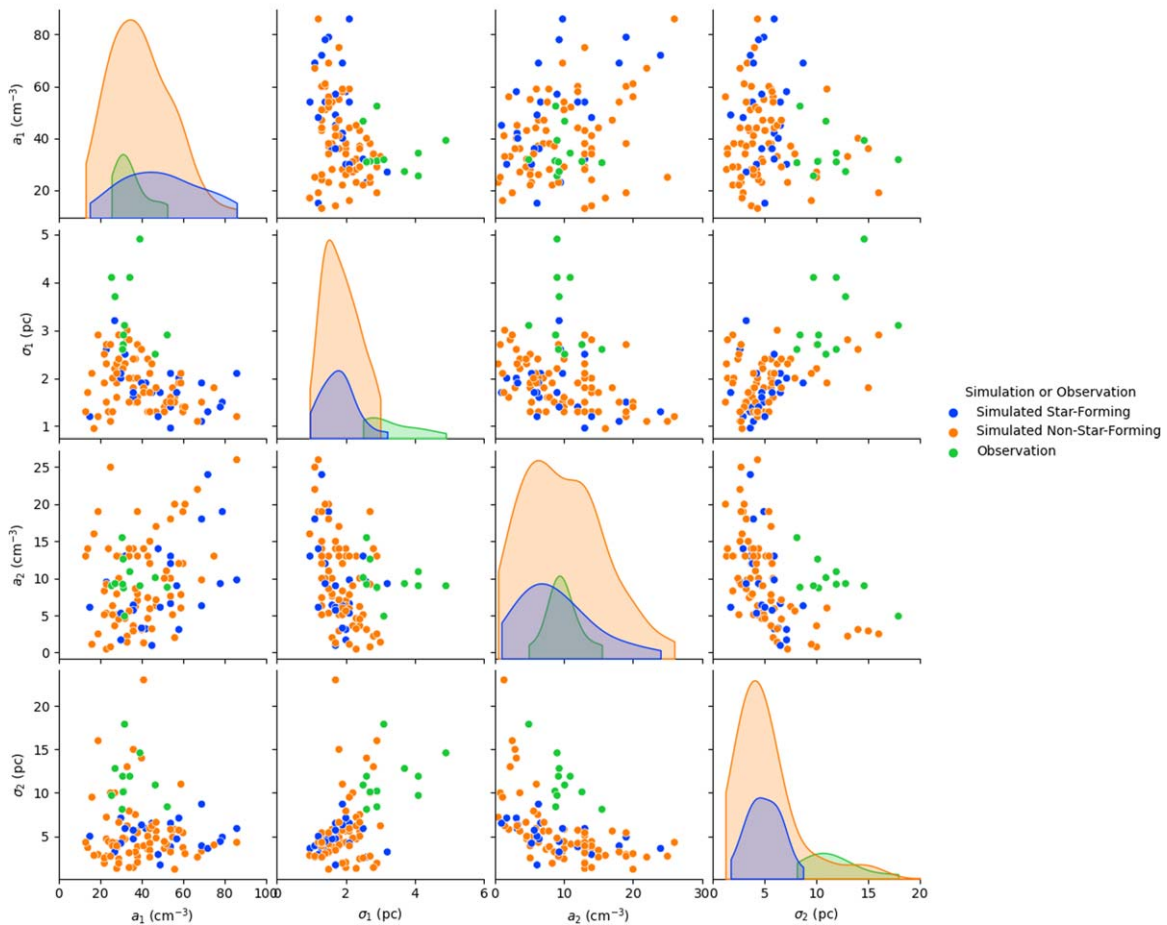


Figure A2. Same as in Figure 4, but with the inclusion of clouds with no star-forming sinks.

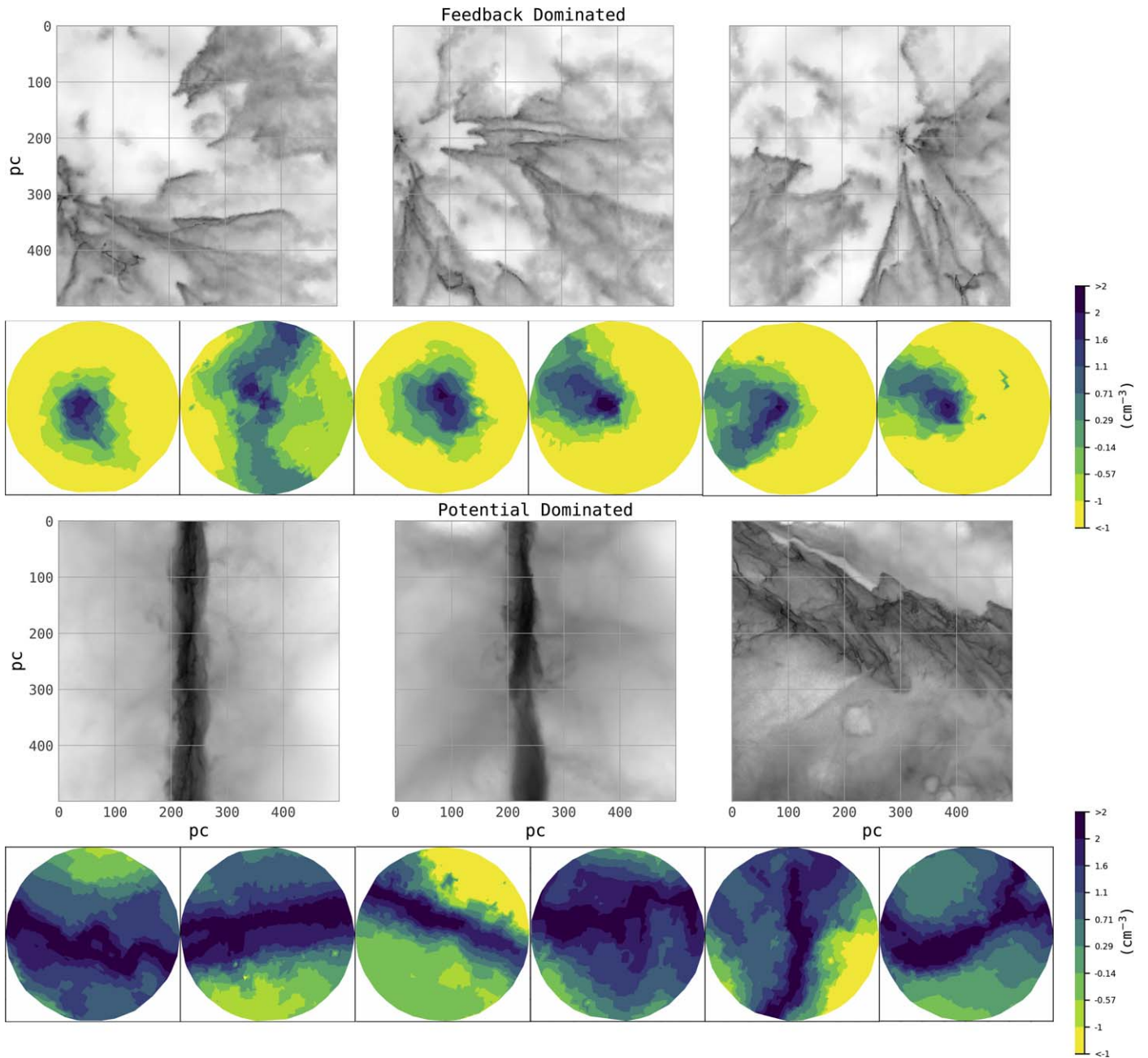


Figure A3. Figure displaying the difference between the feedback- (grid x1465y1400z1400) and potential-dominated (grid x1480y1390zp) grids, both 500×500 pc in size. Panels display projected 2D representations of the 3D density grids by taking a sum along one axis. The potential-dominated grid contains a disk of high-density gas compared to the feedback-dominated grid. Due to the higher density, in order to run the pipeline, a mask was applied at a thresholded density of 500 cm^{-3} in lieu of the 35 cm^{-3} used in the feedback-dominated case. The slices (20×20 pc in size) display the morphology of clouds found in each grid. The clouds in the potential-dominated case are more sheet-like and denser than clouds in the feedback-dominated case.

composition and temperature radial profiles, allowing for simulations to produce mappings from observed total gas density to other tracers once simulations map observations better.

Figure A2 shows the ensemble results of the entire sample size, specifically with the inclusion of non-star forming clouds and served as a point of comparison with Figure 4. These clouds display similar properties to star-forming clouds.

Figure A3 displays the difference between feedback- and potential-dominated grids, and show that magnetic fields might be the reason for the lack of extended envelopes in the simulations.

ORCID iDs

Elijah Mullens <https://orcid.org/0000-0003-0814-7923>
 Catherine Zucker <https://orcid.org/0000-0002-2250-730X>
 Claire E. Murray <https://orcid.org/0000-0002-7743-8129>
 Rowan Smith <https://orcid.org/0000-0002-0820-1814>

References

Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, *A&A*, **628**, A94
 André, P., Men'shchikov, A., Bontemps, S., et al. 2010, *A&A*, **518**, L102
 Armillotta, L., Krumholz, M. R., & Di Teodoro, E. M. 2020, *MNRAS*, **493**, 5273
 Bate, M. R., Bonnell, I. A., & Price, N. M. 1995, *MNRAS*, **277**, 362

- Blitz, L. 1993, in *Protostars and Planets III*, ed. E. H. Levy & J. I. Lunine (Tucson, AZ: Univ. Arizona Press), [125](#)
- Cahlon, S., Zucker, C., Goodman, A., Lada, C., & Alves, J. 2024, [ApJ](#), [961](#), [153](#)
- Chevance, M., Krumholz, M. R., McLeod, A. F., et al. 2022, in *ASP Conf. Ser.* 534, *Protostars and Planets VII*, ed. S. Inutsuka et al. (San Francisco, CA: ASP)
- Clark, P. C., Glover, S. C. O., & Klessen, R. S. 2012, [MNRAS](#), [420](#), [745](#)
- Colombo, D., Rosolowsky, E., Duarte-Cabral, A., et al. 2019, [MNRAS](#), [483](#), [4291](#)
- Cox, D. P., & Gómez, G. C. 2002, [ApJS](#), [142](#), [261](#)
- Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, [ApJ](#), [547](#), [792](#)
- Dharmawardena, T. E., Bailer-Jones, C. A. L., Founesneau, M., et al. 2023, [MNRAS](#), [519](#), [228](#)
- Dobbs, C. L., Krumholz, M. R., Ballesteros-Paredes, J., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), [3](#)
- Draine, B. T. 1978, [ApJS](#), [36](#), [595](#)
- Draine, B. T. 2009, in *ASP Conf. Ser.* 414, *Cosmic Dust—Near and Far*, ed. T. Henning, E. Grün, & J. Steinacker (San Francisco, CA: ASP), [453](#)
- Duarte-Cabral, A., Colombo, D., Urquhart, J. S., et al. 2021, [MNRAS](#), [500](#), [3027](#)
- Edenhofer, G., Zucker, C., Frank, P., et al. 2023, [arXiv:2308.01295](#)
- Froebrich, D., & Rowles, J. 2010, [MNRAS](#), [406](#), [1350](#)
- Ganguly, S., Walch, S., Seifried, D., Clarke, S. D., & Weis, M. 2023, [MNRAS](#), [525](#), [721](#)
- Gatto, A., Walch, S., Low, M. M. M., et al. 2015, [MNRAS](#), [449](#), [1057](#)
- Glover, S. C. O., & Mac Low, M. 2007a, [ApJS](#), [169](#), [239](#)
- Glover, S. C. O., & Mac Low, M.-M. 2007b, [ApJ](#), [659](#), [1317](#)
- Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, [ApJ](#), [887](#), [93](#)
- Habing, H. J. 1968, *BAN*, [19](#), [421](#)
- Hu, C.-Y., Sternberg, A., & van Dishoeck, E. F. 2021, [ApJ](#), [920](#), [44](#)
- Jackson, J. M., Rathborne, J. M., Shah, R. Y., et al. 2006, [ApJS](#), [163](#), [145](#)
- Jeffreson, S. M. R., Keller, B. W., Winter, A. J., et al. 2021, [MNRAS](#), [505](#), [1678](#)
- Kim, C.-G., Kim, J.-G., Gong, M., & Ostriker, E. C. 2023, [ApJ](#), [946](#), [3](#)
- Koch, E. W., & Rosolowsky, E. W. 2015, [MNRAS](#), [452](#), [3435](#)
- Leike, R. H., Glatzle, M., & Enßlin, T. A. 2020, [A&A](#), [639](#), [A138](#)
- Lombardi, M., Lada, C. J., & Alves, J. 2010, [A&A](#), [512](#), [A67](#)
- McMillan, P. J. 2017, [MNRAS](#), [465](#), [76](#)
- Men'shchikov, A. 2013, [A&A](#), [560](#), [A63](#)
- Nelson, R. P., & Langer, W. D. 1997, [ApJ](#), [482](#), [796](#)
- Pakmor, R., Springel, V., Bauer, A., et al. 2016, [MNRAS](#), [455](#), [1134](#)
- Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2011, [A&A](#), [536](#), [A25](#)
- Reipurth, B. 2008a, *Handbook of Star Forming Regions, Volume I: The Northern Sky*, Vol. 4 (San Francisco, CA: ASP)
- Reipurth, B. 2008b, *Handbook of Star Forming Regions, Volume II: The Southern Sky*, Vol. 5 (San Francisco, CA: ASP)
- Rezaei, Kh. S., & Kainulainen, J. 2022, [ApJL](#), [930](#), [L22](#)
- Ridge, N. A., Di Francesco, J., Kirk, H., et al. 2006, [AJ](#), [131](#), [2921](#)
- Seifried, D., Haid, S., Walch, S., Borchert, E. M. A., & Bisbas, T. G. 2020, [MNRAS](#), [492](#), [1465](#)
- Seifried, D., Walch, S., Girichidis, P., et al. 2017, [MNRAS](#), [472](#), [4797](#)
- Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, [MNRAS](#), [441](#), [1628](#)
- Smith, R. J., Treß, R. G., Sormani, M. C., et al. 2020, [MNRAS](#), [492](#), [1594](#)
- Sormani, M. C., Treß, R. G., Klessen, R. S., & Glover, S. C. O. 2017, [MNRAS](#), [466](#), [407](#)
- Springel, V. 2010, [MNRAS](#), [401](#), [791](#)
- Tress, R. G., Smith, R. J., Sormani, M. C., et al. 2020, [MNRAS](#), [492](#), [2973](#)
- Vergely, J. L., Lallement, R., & Cox, N. L. J. 2022, [A&A](#), [664](#), [A174](#)
- Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, [ApJ](#), [587](#), [278](#)
- Zucker, C., & Chen, H. H.-H. 2018, [ApJ](#), [864](#), [152](#)
- Zucker, C., Goodman, A., Alves, J., et al. 2021, [ApJ](#), [919](#), [35](#)