



# The VLA/ALMA Nascent Disk and Multiplicity (VANDAM) Survey of Orion Protostars.

## VI. Insights from Radiative Transfer Modeling

Patrick D. Sheehan<sup>1,2,5</sup> , John J. Tobin<sup>2</sup> , Leslie W. Looney<sup>3</sup> , and S. Thomas Megeath<sup>4</sup>

<sup>1</sup> Center for Interdisciplinary Exploration and Research in Astronomy, Northwestern University, 1800 Sherman Rd., Evanston, IL 60202, USA

<sup>2</sup> National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA 22901, USA

<sup>3</sup> Department of Astronomy, University of Illinois, Urbana, IL 61801, USA

<sup>4</sup> Ritter Astrophysical Research Center, Department of Physics and Astronomy, University of Toledo, Toledo, OH 43606, USA

Received 2021 October 7; revised 2022 February 18; accepted 2022 February 18; published 2022 April 14

### Abstract

We present Markov Chain Monte Carlo radiative transfer modeling of a joint ALMA 345 GHz and spectral energy distribution data set for a sample of 97 protostellar disks from the VLA and ALMA Nascent Disk and Multiplicity Survey of Orion Protostars. From this modeling, we derive disk and envelope properties for each protostar, allowing us to examine the bulk properties of a population of young protostars. We find that disks are small, with a median dust radius of  $29.4^{+4.1}_{-2.7}$  au and a median dust mass of  $5.8^{+4.6}_{-2.7} M_{\oplus}$ . We find no statistically significant difference between most properties of Class 0, Class I, and flat-spectrum sources with the exception of envelope dust mass and inclination. The distinction between inclination is an indication that the Class 0/I/flat-spectrum system may be difficult to tie uniquely to the evolutionary state of protostars. When comparing with Class II disk dust masses in Taurus from similar radiative transfer modeling, we further find that the trend of disk dust mass decreasing from Class 0 to Class II disks is no longer present, though it remains unclear whether such a comparison is fair owing to differences in star-forming region and modeling techniques. Moreover, the disks we model are broadly gravitationally stable. Finally, we compare disk masses and radii with simulations of disk formation and find that magnetohydrodynamical effects may be important for reproducing the observed properties of disks.

*Unified Astronomy Thesaurus concepts:* Protostars (1302); Protoplanetary disks (1300); Star formation (1569); Planet formation (1241)

*Supporting material:* figure set, machine-readable table

### 1. Introduction

Stars form from clouds of gas and dust that collapse under the force of their own gravity. Because of the initial angular momentum of the collapsing material, much of the material in the cloud forms into a disk rather than collapsing straight onto the star forming at the center. Material is then accreted through the disk onto the star, which in turn regulates much of the final buildup of stellar mass. Moreover, it is in these protostellar disks that planets are expected to form, and so understanding their properties throughout their evolution is crucial for understanding how planets form.

A classification scheme has been developed for young stellar objects, initially based on observational properties such as the near-infrared spectral index ( $\alpha$ ; e.g., Myers et al. 1987), the bolometric temperature ( $T_{\text{bol}}$ ; e.g., Chen et al. 1995), or the submillimeter luminosity of the source (e.g., Andre et al. 1993); however, this scheme has also been mapped to propose an evolutionary picture of star formation. In this picture, forming stars progress from young sources with “protostellar” disks, still embedded in substantial envelopes of infalling material that obscure much of the shorter wavelength light (Class 0), to systems with more mature disks but still embedded in less massive envelopes (Class I or flat spectrum), to envelope-free

pre-main-sequence stars with “protoplanetary” or “planet-forming” disks (Class II), to stars with only a small amount of, if any, remnant disk material (Class III) (e.g., Robitaille et al. 2006; Crapsi et al. 2008). The ages of sources in the Class 0 and I stages have been estimated at roughly 0.2 and 0.5 Myr, respectively, using counting statistics (e.g. Evans et al. 2009; Dunham et al. 2015), though estimates of the half-lives suggest potentially shorter typical timescales (e.g., Kristensen & Dunham 2018). Class II sources have typically been estimated to have ages of a few million years by comparing sources with evolutionary track models (e.g., Hillenbrand & Carpenter 2000).

Measurements of the properties of the youngest disks, the protostellar or Class 0/I disks, are keys to informing much of our understanding of the processes that drive star and planet formation. For many years it was unclear whether a disk could even form during the initial collapse of cloud material, as “magnetic braking” found in ideal magnetohydrodynamics (MHD) simulations arrested the formation of a disk (e.g., Allen et al. 2003; Mellon & Li 2008; Li et al. 2014). In recent years, a number of Keplerian-rotating Class 0/I disks have been identified (e.g., Tobin et al. 2012; Murillo et al. 2013; Codella et al. 2014; Harsono et al. 2014; Yen et al. 2017); even more compact, often elongated, continuum disk candidates have been detected around Class 0/I sources (e.g., Segura-Cox et al. 2016; Sheehan & Eisner 2017; Maury et al. 2018; Segura-Cox et al. 2018; Tobin et al. 2020a); and simulations have also found that the inclusion of nonideal MHD effects, turbulence, and/or misalignment between the axis of rotation and the magnetic field can overcome the magnetic braking catastrophe

<sup>5</sup> NSF Astronomy & Astrophysics Fellow.



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

(e.g., Dapp et al. 2012; Joos et al. 2012; Li et al. 2013; Hennebelle et al. 2016; Masson et al. 2016; Wurster et al. 2019). So while the question whether disks can form early in the star formation process may no longer remain, the relative importance of these different pieces of physics remains unclear. The properties of actual protostellar disks are therefore critical for placing constraints on these simulations.

Moreover, isolating disk properties, particularly of young, embedded protostars, is also crucial for understanding the early stages of planet formation. There is a significant amount of evidence that the amount of material available in protoplanetary disks is insufficient to form the masses of planetary systems that we observe (e.g., Najita & Kenyon 2014; Ansdell et al. 2016; Manara et al. 2018; Tychoniec et al. 2020), suggesting that by the protoplanetary disk phase much of the solid material in the disk may be already locked up in larger bodies that are not detectable with ALMA. Lending support to this idea are the wide array of substructures that have been found in nearly every protoplanetary disk observed with high enough spatial resolution (e.g., Brogan et al. 2015; Andrews et al. 2016; Isella et al. 2016; Andrews et al. 2018; Huang et al. 2018; Long et al. 2018; van der Marel et al. 2019). Moreover, purported protoplanets, or direct signatures of the presence of protoplanets, have been found in high-contrast near-infrared imaging and submillimeter continuum and spectral line imaging of a few of these substructured disks (e.g., Sallum et al. 2015; Keppler et al. 2018; Pinte et al. 2018; Teague et al. 2018; Isella et al. 2019; Pinte et al. 2020).

Despite their importance, the properties of protostellar disks have long been difficult to study owing to the dense envelopes of material in which they are embedded. Though millimeter observations are traditionally thought to be optically thin to dust, moderate-resolution millimeter images can still have a significant amount of envelope emission entangled with that of the disks (e.g., Dunham et al. 2014), making early millimeter studies of protostars difficult to interpret (e.g., Jørgensen et al. 2009; Andersen et al. 2019). The VLA and ALMA Nacent Disk and Multiplicity (VANDAM) Survey in Perseus (Tobin et al. 2016b; Tychoniec et al. 2018) and subsequent VANDAM: Orion Survey (Tobin et al. 2020a) provided the first high-resolution imaging surveys (spatial resolution of  $\sim 30$ – $40$  au) of entire populations of protostellar disks and put together a comprehensive picture of protostellar disk dust masses and radii, but even at these spatial resolutions envelopes can contribute to the images in difficult-to-disentangle ways.

One way to combat these issues is to use radiative transfer modeling to account for how the density, temperature, optical depth, viewing angle, and other effects ultimately come together to produce the emergent intensity distribution of young systems. Sheehan & Eisner (2017) applied rigorous radiative transfer forward modeling to CARMA + spectral energy distribution (SED) observations for a sample of 10 Class 0/I protostars in Taurus, and from this modeling they found that embedded disks are, on average, more massive than protoplanetary disks by a factor of a few, though with only 10 sources the significance could be improved. Moreover, conclusions about other disk properties (e.g., radius) could not be drawn. Still, that work provided a blueprint for understanding the properties of these young, embedded disks through careful modeling.

In this work, we build on that blueprint by applying the same radiative transfer modeling framework to protostellar disks in

the Orion Molecular Cloud Complex, observed as a part of the VANDAM: Orion (Tobin et al. 2020a) and HOPS (Manoj et al. 2013; Stutz et al. 2013; Furlan et al. 2016; Fischer et al. 2020) surveys. These surveys collectively observed more than 300 protostellar disks with ALMA, Spitzer, and Herschel to provide a rich data set including high spatial resolution continuum imaging and broadband SEDs. With 97 protostellar disks modeled, this constitutes the largest sample of sources with such careful modeling to date, by over an order of magnitude. As such, it presents an opportunity to begin to put together a picture of the structures of protostellar disks at early times ( $\lesssim 0.5$ – $1$  Myr; e.g., Evans et al. 2009; Dunham et al. 2015).

In Section 2 we discuss our observations, sample selection, and data reduction processes. Then, in Section 3 we give an overview of our modeling procedure, and we present the results of this modeling in Section 4. Finally, we compare the disk properties derived from our modeling to observationally derived disk properties for Class II disks, theoretical simulations of disk formation, and system properties derived from other observational methods (e.g., bolometric temperature, millimeter flux), and we discuss the implications of these in Section 5 and wrap up with our conclusions in Section 6.

## 2. Observations and Sample

The data we use for our modeling are drawn from the VANDAM Survey of Orion Protostars (Tobin et al. 2020a), which surveyed 328 protostars in the Orion Molecular Cloud complex. These observations include ALMA 345 GHz continuum and spectral line observations at  $0''.1$  spatial resolution, along with Very Large Array (VLA) 33 GHz continuum observations at  $0''.06$ , for all protostars surveyed. The data reduction for these observations, including self-calibration when the signal-to-noise ratio was high enough, is described in detail in Tobin et al. (2020a).

Due to the computational cost of running our modeling, we were unable to model all 328 protostars with the available computing time, so we necessarily had to make cuts to the full sample to make the modeling more tractable. From this sample, we excluded  $\sim 100$  protostellar multiple systems, as modeling of multiple systems is significantly more challenging (e.g., Sheehan & Eisner 2014). We also excluded  $\sim 30$  nondetections from the pool of protostars. Of the remaining  $\sim 200$  single protostars, we randomly selected 90 protostars for our modeling analysis, regardless of their signal-to-noise ratios. By selecting protostars completely at random, we hope to remove any potential biases that may be induced by not using the full sample. We do, however, also try to correct for our exclusion of the nondetections whenever appropriate, as we will describe below.

For our final sample of 97 protostars, including 25 Class 0 sources, 44 Class I sources, and 28 flat-spectrum sources, we collect the ALMA 345 GHz continuum observations for our modeling analysis. We exclude the VLA 33 GHz continuum observations, as there was not uniform coverage of all VANDAM: Orion targets (Tobin et al. 2020a), but also because modeling multiple millimeter observations would stretch our already thin computational resources and because our modeling code is not fully tested in that mode. In addition to our ALMA data set, we collect archival photometry and spectroscopy to include in our modeling analysis, primarily from the Herschel Orion Protostar Survey (HOPS; Furlan et al. 2016). This includes Spitzer IRAC and MIPS photometry and

IRS spectroscopy, Two Micron All Sky Survey photometry, and Herschel PACS and SPIRE photometry, when available. To account for systematics across different instruments, we assume a 10% flux calibration uncertainty for all flux measurements in our modeling.

We apply our modeling directly to the ALMA 345 GHz two-dimensional visibilities; however, we do show images of sources and their best-fit models for easier by-eye comparison. The images we show typically were made using Briggs weighting with a robust parameter of 0.5, to balance resolution and sensitivity, though for some faint sources we used a robust parameter of 2 to improve the sensitivity of the images. To ensure that in our modeling we are deriving reasonable parameter estimates and uncertainties, we checked the uncertainties on the observed visibility data by comparing  $\sigma_{\text{tot}} = 1/\sqrt{\sum_i (1/\sigma_i^2)}$  with the rms of a naturally weighted image and found that we should reduce the weights by a factor of 0.25 for the two to match. Though this is not a perfect comparison, as we are reducing the weights, this should provide more conservative estimates of the uncertainties on the data and therefore on the best-fit parameters.

### 3. Radiative Transfer Modeling

To model our sample, we use the `pdspy` code (Sheehan 2018), which follows the framework outlined in Sheehan & Eisner (2017) to fit full two-dimensional, axisymmetric, radiative transfer models simultaneously to a multiwavelength data set using Markov Chain Monte Carlo (MCMC) fitting. Our radiative transfer model takes a parameterized density structure for a protostellar system as an input to radiative transfer codes, which then generate synthetic observations of the system to compare with the observations. We describe the details of the model used below, but we also note that extensive documentation of the `pdspy` code is available online.<sup>6</sup>

Our model includes a protostar, disk, and envelope with an outflow cavity. For simplicity, and because protostellar properties are unknown for the majority, if not entirety, of our sample, we assume that the central protostar has a temperature of 4000 K, reasonable for a generic young, low-mass protostar. We do, however, leave the luminosity of that protostar,  $L_*$ , as a free parameter in our fit. As we assume a blackbody spectrum for simplicity,  $L_*$  is primarily controlled by varying the protostellar radius, and so this may not exactly emulate the true spectrum of a protostar, which may have significant accretion luminosity as well. As most of this emission is reprocessed by the envelope, this should not substantially affect the results of the modeling.

The model also includes a protostellar disk following the prescription for a viscously accreting disk, with the surface density described by

$$\Sigma = \Sigma_0 \left( \frac{R}{R_c} \right)^{-\gamma} \exp \left[ - \left( \frac{R}{R_c} \right)^{2-\gamma} \right] \quad (1)$$

(Lynden-Bell & Pringle 1974), where  $R$  is the radius within the disk as typically defined in a cylindrical coordinate system.  $R_c$  is the radius in the disk at which the surface density is exponentially tapered, and  $\gamma$  controls the surface density power-law shape, along with how the disk is tapered.  $\Sigma_0$  is the

surface density at  $R_c$ , but we instead parameterize the model in terms of the total disk mass,  $M_{\text{disk}}$ , from which the surface density normalization can be calculated as

$$\Sigma_0 = \frac{(2 - \gamma) M_{\text{disk}}}{2\pi R_c^2}. \quad (2)$$

The vertical structure of the disk is given by the typical structure assumed for a flared accretion disk, with the density defined as

$$\rho = \frac{\Sigma(R)}{\sqrt{2\pi} h(R)} \exp \left[ - \frac{1}{2} \left( \frac{z}{h(R)} \right)^2 \right]. \quad (3)$$

The function  $h(r)$  defines the scale height of the disk as a function of radius, as

$$h(r) = h_0 \left( \frac{R}{1 \text{ au}} \right)^\psi, \quad (4)$$

with  $h_0$  specifying the scale height of the disk at 1 au and  $\psi$  defining the flaring of the disk. Finally, we also truncate the disk at an inner radius of  $R_{\text{in}}$ , inside of which the density drops to zero.

We also include an envelope in our model based off of the rotating collapsing model from Ulrich (1976). The density of the envelope is given by

$$\rho = \frac{\dot{M}}{4\pi} (GM_* r^3)^{-\frac{1}{2}} \left( 1 + \frac{\mu}{\mu_0} \right)^{-\frac{1}{2}} \left( \frac{\mu}{\mu_0} + 2\mu_0^2 \frac{R_c}{r} \right)^{-1}, \quad (5)$$

where  $r$  is defined as is typical in a spherical coordinate system, and  $\mu = \cos \theta$ .  $R_c$  is the critical radius, inside of which the density profile begins to flatten into a disk-like structure. In our model, we truncate the envelope at an outer radius  $R_{\text{env}}$  beyond which the density drops to zero and define  $R_c = R_{\text{disk}}$ . We also truncate the density at the same inner radius as the disk,  $R_{\text{in}}$ . Finally, rather than parameterize the envelope in terms of the accretion rate,  $\dot{M}$ , we integrate over the entire density structure to calculate the total mass,  $M_{\text{env}}$ , and normalize the density properly from this value.

Finally, we include in our envelope model an outflow cavity in which the density of the envelope is reduced by some factor,  $f_{\text{cav}}$ . The structure of this cavity is defined by

$$z > 1 \text{ au} + R^\xi, \quad (6)$$

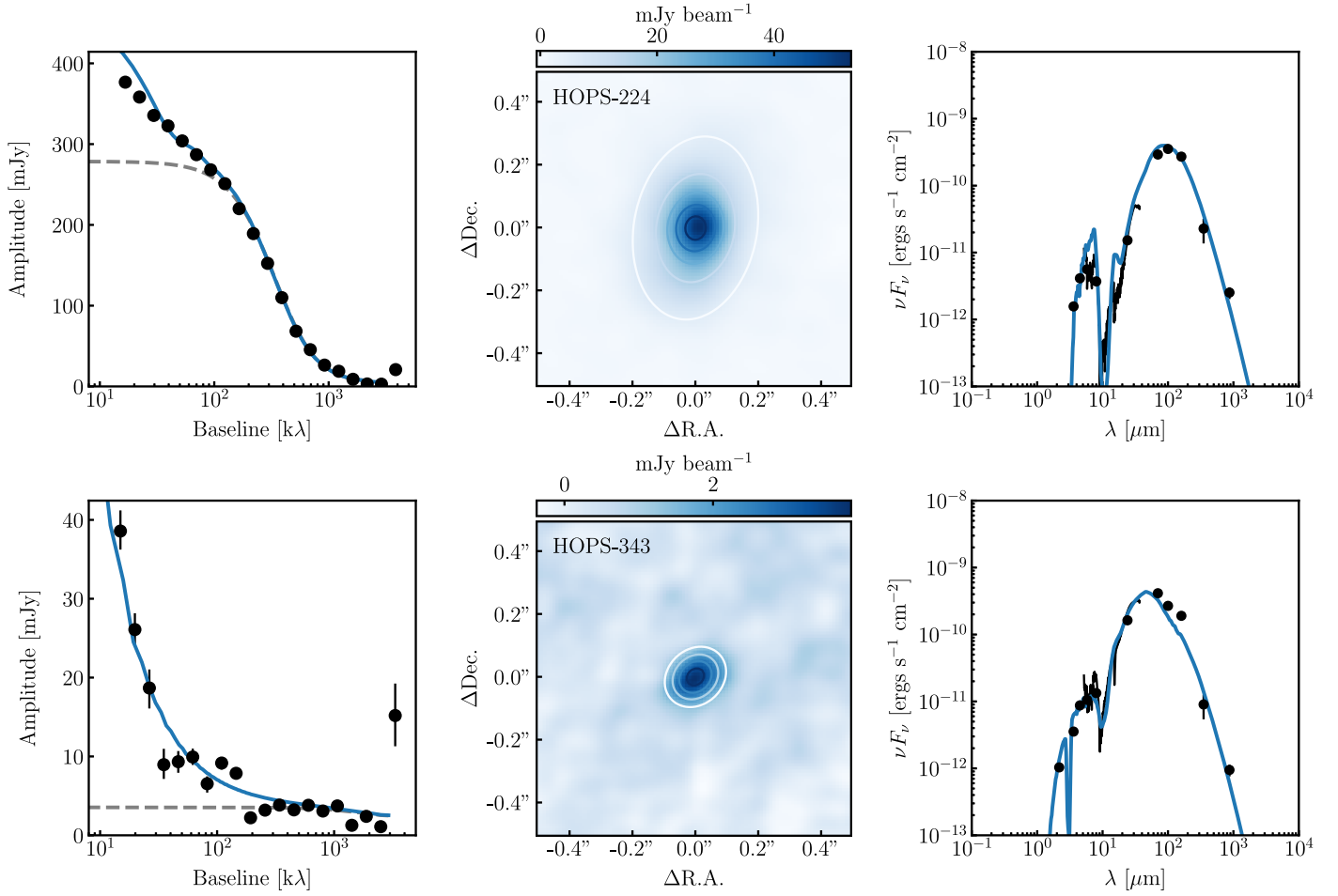
where  $\xi$  defines the shape of the cavity and opening angle. The opening angle, defined as the full angle across the cavity at  $R_{\text{env}}$  from the position of the protostar, is therefore

$$\theta_{\text{open}} = 2 \tan^{-1} \left( \frac{r}{1 \text{ au} + R_{\text{env}}^\xi} \right). \quad (7)$$

We show a simple schematic of this geometry in the Appendix.

SEDs can also be sensitive to extinction from foreground cloud emission that is not a part of the envelope. In previous works we have included foreground extinction as a parameter in our fit when well motivated (e.g., Sheehan & Eisner 2017). Here, for simplicity, because of computational constraints and because foreground extinction can be very degenerate with envelope properties, we leave this out of our model. The spatial information provided by the millimeter visibilities should help

<sup>6</sup> Documentation available at [pdspy.readthedocs.io](https://pdspy.readthedocs.io).



**Figure 1.** Examples of radiative transfer models compared with our observational data set. In the left column, we show the one-dimensional, azimuthally averaged 345 GHz visibility profiles, with the best-fit model shown in blue and the disk contribution to that model as gray dashed lines. For ease of viewing, we show the azimuthally averaged visibilities, but the fit is done to the full, two-dimensional data set. In the middle column we show the 345 GHz image with the model shown in contours. The contours are at levels of 10%, 30%, 50%, 70%, and 90% of the peak flux value, with any contours that fall below  $3\sigma$  excluded, and are meant primarily to demonstrate that the model profile matches the data in two dimensions. Finally, on the right we show the SED with the best-fit model in blue.

(The complete figure set (97 images) is available.)

to mitigate this effect, but it should be noted that for some sources this could impact the envelope properties.

We provide our model with dust opacities following a similar recipe to Woitke et al. (2016), with dust grains initially composed of 70% astronomical silicates (Draine 2003) and 30% carbonaceous material (Zubko et al. 1996), though we then add water ice (Hudgins et al. 1993) at 50% the level of silicates+carbon, for final abundances of 47% silicates, 20% carbonaceous material, and 33% water ice. We follow the distribution of hollow spheres prescription (Min et al. 2005). Finally, we assume that the grains have a power-law size distribution,  $n \propto a^{-p}$  with a minimum size of  $0.05 \mu\text{m}$ . In the envelope, where dust grain growth is likely less advanced, we assume that the maximum dust grain size is  $a_{\text{max}} = 1 \mu\text{m}$  and  $p = 3.5$ . In the disk, however, we allow both  $a_{\text{max}}$  and  $p$  to vary.

This density structure is provided to the RADMC-3D Monte Carlo radiative transfer code (Dullemond 2012), which is first used to calculate the thermal structure of the disk + envelope system. Then, using the density and thermal structure, we use RADMC-3D to generate synthetic SEDs and 345 GHz millimeter images. We then Fourier-transform the synthetic millimeter images to compare directly with our two-dimensional ALMA visibility data. The inclusion of synthetic

observations adds two additional parameters to our model, the position angle P.A. and inclination  $i$ , for a total of 15 free parameters for each fit:

$$\hat{\theta} = \{L_*, M_{\text{disk}}, R_{\text{disk}}, R_{\text{in}}, h_0, \gamma, \psi, M_{\text{env}}, R_{\text{env}}, f_{\text{cav}}, \xi, a_{\text{max}}, p, i, \text{P.A.}\}.$$

To compare these synthetic models with our observational data set for each source, we use the MCMC fitting code *emcee* (Foreman-Mackey et al. 2013), which employs an affine-invariant Monte Carlo sampler to sample parameter space and converge on regions that best fit the data. In our runs we use 200 walkers spread out over a wide range of parameter space. To simultaneously fit our two independent data sets (SED and ALMA 2D visibilities), we use a goodness-of-fit metric that is a linear combination of the  $\chi^2$  values for each data set separately,

$$X^2 = w_{\text{SED}} \chi_{\text{SED}}^2 + w_{\text{vis}} \chi_{\text{vis}}^2. \quad (8)$$

In general, we use  $w_{\text{SED}} = w_{\text{vis}} = 1$ , though in some cases we increase  $w_{\text{SED}}$  to help improve the fit to the SED in the model, as the SED typically has many fewer data points to fit than the visibilities. Moreover, because the SED is more sensitive to the envelope, particularly at short wavelengths, while the visibilities are particularly sensitive to the disk and the two-



**Table 1**  
Best-fit Radiative Transfer Model Parameters

Sources	$L_*$ ( $L_\odot$ )	$M_{\text{disk,dust}}$ ( $M_\oplus$ )	$R_{\text{disk,dust}}$ (au)	$F_{\nu,230\text{GHz}}$ (mJy)	$M_{\text{env,dust}}$ ( $M_\oplus$ )	$R_{\text{env,dust}}$ (au)	$a_{\text{max}}$ ( $\mu\text{m}$ )	$i$ (deg)	$\kappa_{\nu,345\text{GHz}}$ ( $\text{cm}^2 \text{g}^{-1}$ )	$\beta$
HOPS-2	$0.43^{+1.32}_{-0.14}$	$2.8^{+21.8}_{-1.5}$	$11.3^{+8.4}_{-6.7}$	2.876	$0.30^{+23.78}_{-0.18}$	$164^{+9867}_{-56}$	$2.5^{+1060.7}_{-1.3}$	$54^{+11}_{-37}$	3.832	1.604
HOPS-3	$2.30^{+0.81}_{-1.20}$	$15.0^{+11.5}_{-4.7}$	$98^{+22}_{-19}$	15.529	$0.132^{+0.240}_{-0.068}$	$221^{+106}_{-21}$	$397^{+692}_{-235}$	$77.8^{+1.0}_{-1.6}$	11.039	1.286
HOPS-13	$0.73^{+0.25}_{-0.19}$	$0.54^{+0.30}_{-0.38}$	$33^{+22}_{-21}$	0.605	$0.61^{+112.69}_{-0.35}$	$167^{+10661}_{-64}$	$4.1^{+1680.6}_{-2.9}$	$30^{+26}_{-21}$	3.829	1.602
HOPS-16	$0.44^{+0.12}_{-0.10}$	$15.4^{+11.7}_{-8.2}$	$12.4^{+2.3}_{-3.3}$	8.140	$0.35^{+45.56}_{-0.16}$	$146^{+5402}_{-43}$	$95^{+3054}_{-91}$	$26^{+12}_{-19}$	4.398	1.818
HOPS-18	$1.93^{+0.33}_{-0.24}$	$12.6^{+1.6}_{-5.5}$	$83.1^{+5.4}_{-10.7}$	9.097	$53^{+293}_{-32}$	$2289^{+5836}_{-1118}$	$144^{+670}_{-94}$	$77.5^{+1.5}_{-1.7}$	4.981	2.010
HOPS-29	$1.87^{+1.07}_{-0.28}$	$5.4^{+4.0}_{-2.6}$	$15.4^{+3.9}_{-2.2}$	8.534	$165^{+632}_{-150}$	$2891^{+20971}_{-1025}$	$81^{+86}_{-54}$	$50.4^{+12.7}_{-5.9}$	4.089	1.696
HOPS-36	$0.884^{+0.132}_{-0.090}$	$18^{+67}_{-13}$	$22.6^{+8.6}_{-9.5}$	16.079	$0.29^{+0.77}_{-0.12}$	$114^{+246}_{-13}$	$3989^{+45468}_{-3716}$	$55.4^{+2.7}_{-2.7}$	5.957	0.846
HOPS-41	$1.27^{+0.40}_{-0.36}$	$6.8^{+6.4}_{-2.2}$	$8.1^{+1.4}_{-2.0}$	9.679	$95^{+908}_{-66}$	$3846^{+14624}_{-2027}$	$1.26^{+68.65}_{-0.23}$	$27^{+12}_{-17}$	3.816	1.597
HOPS-42	$0.95^{+0.18}_{-0.23}$	$26.6^{+5.3}_{-7.9}$	$110^{+18}_{-22}$	15.409	$0.80^{+91.78}_{-0.14}$	$161^{+23152}_{-10}$	$809^{+82310}_{-623}$	$77.53^{+0.97}_{-1.29}$	5.193	1.511
HOPS-43	$2.17^{+0.61}_{-0.32}$	$0.50^{+0.62}_{-0.26}$	$88^{+816}_{-33}$	0.683	$1545^{+591}_{-672}$	$28211^{+3222}_{-12816}$	$49^{+10918}_{-48}$	$73.3^{+5.5}_{-8.8}$	3.965	1.652

**Note.** The full table contains many more rows than shown here.

(This table is available in its entirety in machine-readable form.)

dimensional structure of the system, it is possible to achieve good fits to both data sets, despite this difference. Because of our adjustments to the weights, however, the uncertainties from our modeling should not be treated as true statistical uncertainties, but they do provide a reasonable estimate of the range of parameter values that can provide a good fit to the data.

As radiative transfer models are computationally expensive to generate, we use a range of supercomputing resources to do our model fitting. The majority of the modeling was done on Bridges at the Pittsburgh Supercomputing Center and Comet at the San Diego Supercomputer Center with a 5 million core-hour allocation through the National Science Foundation (NSF) XSEDE program. Additional models were run on Schooner at the Oklahoma Supercomputing Center for Education & Research (OSCER), and testing was also done on Stampede2 at the Texas Advanced Computing Center (TACC). Fits with *emcee* can be parallelized because, for each step, the 200 walkers can have their likelihood functions evaluated independently. As such, the radiative transfer models for multiple walkers can be computed in parallel. This parallelization can be further tuned by running the RADMC-3D radiative transfer models using multiple cores, creating a hybrid MPI-OpenMPI parallelization scheme. After some testing, we found that the most efficient way to run models was using  $\sim 5$  supercomputer nodes, each with 20–28 cores per node. We would then run the models for 5 walkers simultaneously on each node (for a total of  $\sim 25$  walkers running models simultaneously), with each walker using 4–6 cores to run RADMC-3D (with  $\sim 100$ –150 cores in total).

#### 4. Results

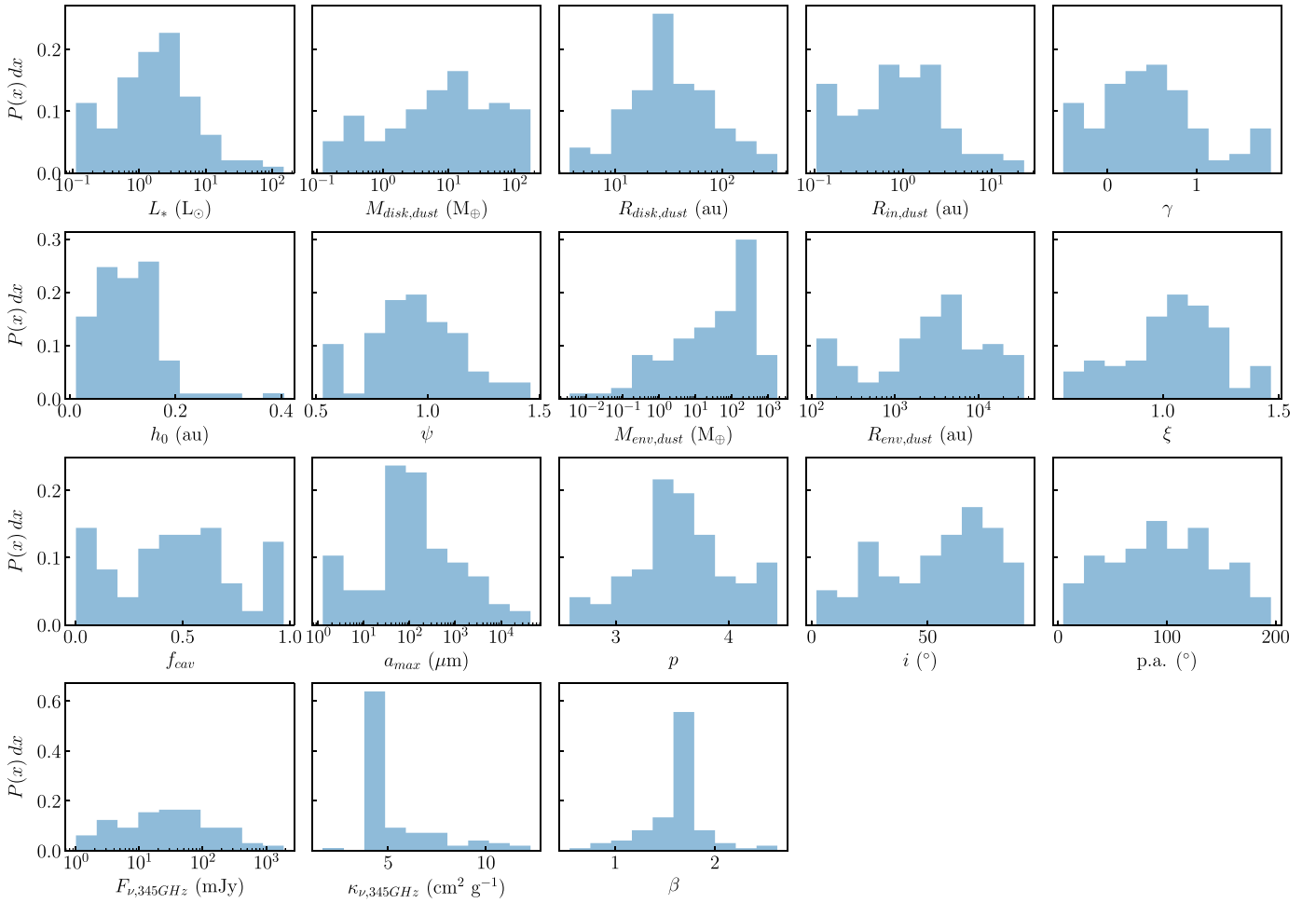
We show example model fits, using the maximum a posteriori model for each source, in Figure 1. We note that in Figure 1 we show the one-dimensional, azimuthally averaged visibility profile for ease of interpretation, but the models are fit to the full, two-dimensional visibility data set. We also list a subset of the most relevant best-fit parameter values in Table 1, with the full table available online and in an accompanying machine-readable table with all best-fit values and derived quantities discussed throughout the remainder of this work. The best-fit parameters are calculated using the maximum

likelihoods of marginalized one-dimensional posteriors for each parameter, after the burn-in steps are discarded. To do so, we use a kernel density estimation to estimate the probability density function (pdf) from the marginalized samples on a fine one-dimensional grid and then report the grid value where the pdf is maximized. The uncertainties on those values are determined by the range around those values containing 95% of the post-burn-in walkers.

We have also provided the full data resulting from our model fits in an online repository at doi:[10.5281/zenodo.5842333](https://doi.org/10.5281/zenodo.5842333), so that the community might make use of our results in their own work. This includes the full posterior distributions for each parameter of each source in our model, along with those additional values derived from our fit parameters as discussed throughout the remainder of the text. We also include, for each source, the *pdspy* configuration file and data files modeled, along with a script that demonstrates how to use *pdspy* to work with the models. With these tools, the resulting models should be easily accessible for anyone to generate and use for further studies.

We note that throughout the remainder of this discussion, all masses, both disk and envelope, account for only the dust and ignore the gas in the system. We follow this convention throughout our analysis, unless otherwise noted, as our radiative transfer modeling is most directly sensitive to the dust content of the system. We therefore refer to disk and envelope properties derived from our modeling as  $M_{\text{disk,dust}}$ ,  $R_{\text{disk,dust}}$ ,  $M_{\text{env,dust}}$ , and  $R_{\text{env,dust}}$  in the relevant tables, figures, and discussion and list masses in units of  $M_\oplus$  to remind readers of this. For readers interested in the gas mass, we do, however, include estimates of the total mass assuming a gas-to-dust ratio of 100:1 in the online version of Table 1 and the machine-readable table accompanying this work, and we refer to these parameters, for example, as  $M_{\text{disk}}$  to indicate the total mass.

In Figure 2, we show the distributions of parameter values from our modeling over our full sample, and therefore a picture of the demographics of protostellar disk properties. We also show the distribution of three additional properties derived from the modeling that are relevant for comparing with other studies of disk demographics: the 345 GHz flux of the disk in our models ( $F_{\nu,345\text{GHz}}$ ), the 345 GHz opacity from our models ( $\kappa_{\nu,345\text{GHz}}$ ), and the spectral index between 230 GHz and 345 GHz ( $\beta$ ). We also list the median value for each property



**Figure 2.** Histograms of the distribution of best-fit parameter values found from our modeling for our sample of 90 protostellar disks. The last row shows quantities that are not formally parameters of our model but are derived from the best-fit models.

measured by our modeling over our sample of protostars in Table 2. Median values were calculated using the Kaplan–Meier fitter in the *lifelines* package (Davidson-Pilon et al. 2019). To estimate the uncertainty on these median values, we generated 1000 realizations of the distribution of protostellar system (star + disk + envelope) properties by randomly sampling a set of parameters from the posterior for each protostar for each realization. We calculated the median parameter values for each of these 1000 realizations, and the uncertainty reported in Table 2 is given by adding the 68% inclusion range of medians from the realizations around the median value calculated from the Kaplan–Meier fitter and adding this in quadrature with the counting uncertainty reported by the Kaplan–Meier fitter.

Of particular note, we find that the protostellar disks are small, with a median dust radius of  $29.4^{+4.1}_{-3.2}$  au. Large disks with  $R_{\text{disk}} > 100$  au do exist, though they only account for  $11.3^{+4.6}_{-3.4}\%$  of our sample. This is in reasonable agreement with Maury et al. (2019) and Tobin et al. (2020a), both of which found that large protostellar disks are less common. The median embedded disk dust mass across all different classifications is  $5.8^{+4.6}_{-2.7} M_{\oplus}$ , and the median envelope dust mass is  $75.9^{+13.7}_{-51.3} M_{\oplus}$ .

The total envelope masses, assuming a gas-to-dust ratio of 100, range from  $\sim 10^{-4}$  to  $1 M_{\odot}$  with a handful of sources with even lower masses. This range is in reasonably good agreement

with envelope masses estimated by Furlan et al. (2016), who matched a grid of radiative transfer models with the SEDs for the same set of sources considered here. They are somewhat lower than what has been found from single-dish observations of other star-forming regions, which tend to fall in the  $0.1\text{--}10 M_{\odot}$  range (e.g., Enoch et al. 2008; Sadavoy et al. 2014; Pezzuto et al. 2021). Those surveys, however, tend to find preferentially young, likely Class 0 sources, which tend to have more massive envelopes. Even considering just the Class 0 sources in our sample, however, we are still lacking the envelopes with  $>1 M_{\odot}$  of material. Though this could be due to true differences in the envelope mass distributions between regions, it is also very likely that the difference could be due to the scales considered. Though we do fit single-dish fluxes from the HOPS Survey when available, our ALMA observations are primarily tracing emission on scales up to a few thousand au and may be missing out on larger-scale cloud emission.

To explore how protostellar disk and envelope properties change with evolutionary stage, we show the cumulative distribution of parameter values split into the three classes of sources examined here, Class 0, Class I, and flat spectrum, in Figure 3. We note that for most parameters we have no a priori reason to believe that our observations are biased toward missing a particular range of values, and so we assume that the sample is complete. The exception to this is the disk dust mass, as we selected only sources that were detected with ALMA,

**Table 2**  
Comparison of Class 0/I/Flat-spectrum Properties

Parameters	Medians				Two-sample Tests <sup>a</sup>		
	All ( <i>N</i> = 97)	Class 0 ( <i>N</i> = 25)	Class I ( <i>N</i> = 44)	Flat ( <i>N</i> = 28)	0 versus I	0 versus Flat	I versus Flat
$L_*$ ( $L_\odot$ )	1.90 <sup>+0.27</sup> <sub>-0.30</sub>	2.90 <sup>+0.99</sup> <sub>-0.46</sub>	1.84 <sup>+0.11</sup> <sub>-0.71</sub>	1.59 <sup>+0.55</sup> <sub>-0.66</sub>	6.4	0.0	0.0
$M_{\text{disk,dust}}$ ( $M_\oplus$ )	5.8 <sup>+4.6</sup> <sub>-2.7</sub>	7.1 <sup>+14.3</sup> <sub>-2.0</sub>	4.9 <sup>+1.0</sup> <sub>-2.7</sub>	14.0 <sup>+1.3</sup> <sub>-7.0</sub>	69.0	0.0	0.0
$R_{\text{disk,dust}}$ (au)	29.4 <sup>+4.1</sup> <sub>-3.2</sub>	35.6 <sup>+17.1</sup> <sub>-10.0</sub>	26.9 <sup>+4.5</sup> <sub>-3.3</sub>	29.5 <sup>+6.1</sup> <sub>-4.4</sub>	0.0	0.0	0.0
$R_{\text{in,dust}}$ (au)	0.8 <sup>+0.2</sup> <sub>-0.1</sub>	1.7 <sup>+0.7</sup> <sub>-0.3</sub>	1.0 <sup>+0.5</sup> <sub>-0.3</sub>	0.2 <sup>+0.1</sup> <sub>-0.0</sub>	18.5	<b>100.0</b>	<b>100.0</b>
$\gamma$	0.4 <sup>+0.1</sup> <sub>-0.1</sub>	0.4 <sup>+0.1</sup> <sub>-0.0</sub>	0.4 <sup>+0.2</sup> <sub>-0.2</sub>	0.5 <sup>+0.1</sup> <sub>-0.2</sub>	0.0	0.5	0.6
$h_0$ (au)	0.11 <sup>+0.01</sup> <sub>-0.01</sub>	0.09 <sup>+0.03</sup> <sub>-0.02</sub>	0.10 <sup>+0.02</sup> <sub>-0.01</sub>	0.12 <sup>+0.02</sup> <sub>-0.01</sub>	2.5	7.5	1.9
$\psi$	0.93 <sup>+0.05</sup> <sub>-0.01</sub>	0.98 <sup>+0.06</sup> <sub>-0.05</sub>	0.99 <sup>+0.06</sup> <sub>-0.08</sub>	0.89 <sup>+0.02</sup> <sub>-0.06</sub>	1.4	<b>94.4</b>	93.5
$M_{\text{env,dust}}$ ( $M_\oplus$ )	75.9 <sup>+13.7</sup> <sub>-51.3</sub>	232.3 <sup>+21.3</sup> <sub>-102.5</sub>	24.5 <sup>+33.6</sup> <sub>-13.8</sub>	13.7 <sup>+25.7</sup> <sub>-8.0</sub>	15.5	<b>98.2</b>	6.2
$R_{\text{env,dust}}$ (au)	3429.4 <sup>+602.3</sup> <sub>-1116.5</sub>	2130.4 <sup>+1492.6</sup> <sub>-294.3</sub>	3513.6 <sup>+720.0</sup> <sub>-1251.6</sub>	4031.6 <sup>+1733.5</sup> <sub>-2015.0</sub>	4.0	6.4	0.2
$\xi$	1.056 <sup>+0.027</sup> <sub>-0.031</sub>	1.077 <sup>+0.037</sup> <sub>-0.035</sub>	1.056 <sup>+0.082</sup> <sub>-0.041</sub>	1.020 <sup>+0.051</sup> <sub>-0.130</sub>	2.4	1.6	12.1
$f_{\text{cav}}$	0.48 <sup>+0.04</sup> <sub>-0.06</sub>	0.48 <sup>+0.03</sup> <sub>-0.11</sub>	0.51 <sup>+0.06</sup> <sub>-0.11</sub>	0.41 <sup>+0.13</sup> <sub>-0.01</sub>	2.9	3.7	1.3
$a_{\text{max}}$ ( $\mu\text{m}$ )	101 <sup>+34</sup> <sub>-17</sub>	91 <sup>+45</sup> <sub>-12</sub>	82 <sup>+50</sup> <sub>-24</sub>	149 <sup>+60</sup> <sub>-71</sub>	0.3	4.4	2.4
$p$	3.55 <sup>+0.07</sup> <sub>-0.05</sub>	3.54 <sup>+0.14</sup> <sub>-0.10</sub>	3.53 <sup>+0.14</sup> <sub>-0.07</sub>	3.57 <sup>+0.25</sup> <sub>-0.15</sub>	0.8	1.0	2.0
$i$ (deg)	57.2 <sup>+3.2</sup> <sub>-2.0</sub>	57.2 <sup>+7.8</sup> <sub>-2.9</sub>	69.4 <sup>+2.6</sup> <sub>-4.3</sub>	29.9 <sup>+3.1</sup> <sub>-4.9</sub>	0.1	<b>99.7</b>	<b>100.0</b>
P.A. (deg)	96.6 <sup>+8.6</sup> <sub>-7.0</sub>	91.3 <sup>+32.6</sup> <sub>-5.2</sub>	98.1 <sup>+13.3</sup> <sub>-6.7</sub>	86.9 <sup>+22.1</sup> <sub>-11.5</sub>	0.0	0.0	0.1
Derived Quantities							
$F_{\nu,345\text{GHz}}$ (mJy)	26.353 <sup>+11.239</sup> <sub>-5.069</sub>	59.913 <sup>+28.087</sup> <sub>-30.849</sub>	19.194 <sup>+4.236</sup> <sub>-4.263</sub>	39.255 <sup>+5.903</sup> <sub>-8.303</sub>	100.0	0.0	0.0
$\kappa_{\nu,345\text{GHz}}$ ( $\text{cm}^2 \text{ g}^{-1}$ )	4.2 <sup>+0.2</sup> <sub>-0.1</sub>	4.0 <sup>+0.3</sup> <sub>-0.0</sub>	4.1 <sup>+0.8</sup> <sub>-0.1</sub>	4.4 <sup>+0.3</sup> <sub>-0.3</sub>	0.0	0.0	0.0
$\beta$	1.6 <sup>+0.0</sup> <sub>-0.0</sub>	1.6 <sup>+0.0</sup> <sub>-0.0</sub>	1.6 <sup>+0.0</sup> <sub>-0.0</sub>	1.6 <sup>+0.0</sup> <sub>-0.0</sub>	0.0	0.0	0.0

**Note.**

<sup>a</sup> These columns report the percentage of 1000 randomly sampled realizations of the distribution of each parameter that have  $p < 0.05$  from two-sample tests comparing different protostar classes. Each realization is generated by randomly sampling a parameter value from the posterior for each source, for all sources in the sample. We bold parameter/Class combinations for which  $>90\%$  of all realizations have  $p < 0.05$ , suggesting that the distributions are distinct.

and the millimeter brightness is related to disk dust mass. Though in reality it is more complicated than this, for simplicity we add additional sources with upper limits on their disk dust mass equivalent to the lowest dust mass disk that we modeled until the fraction of detected sources in our sample matches the overall survey (0.88 for all sources, 0.91 for Class 0 sources, 0.85 for Class I sources, and 0.88 for flat-spectrum sources; Tobin et al. 2020a), when generating cumulative distributions or running two-sample tests for disk dust masses.

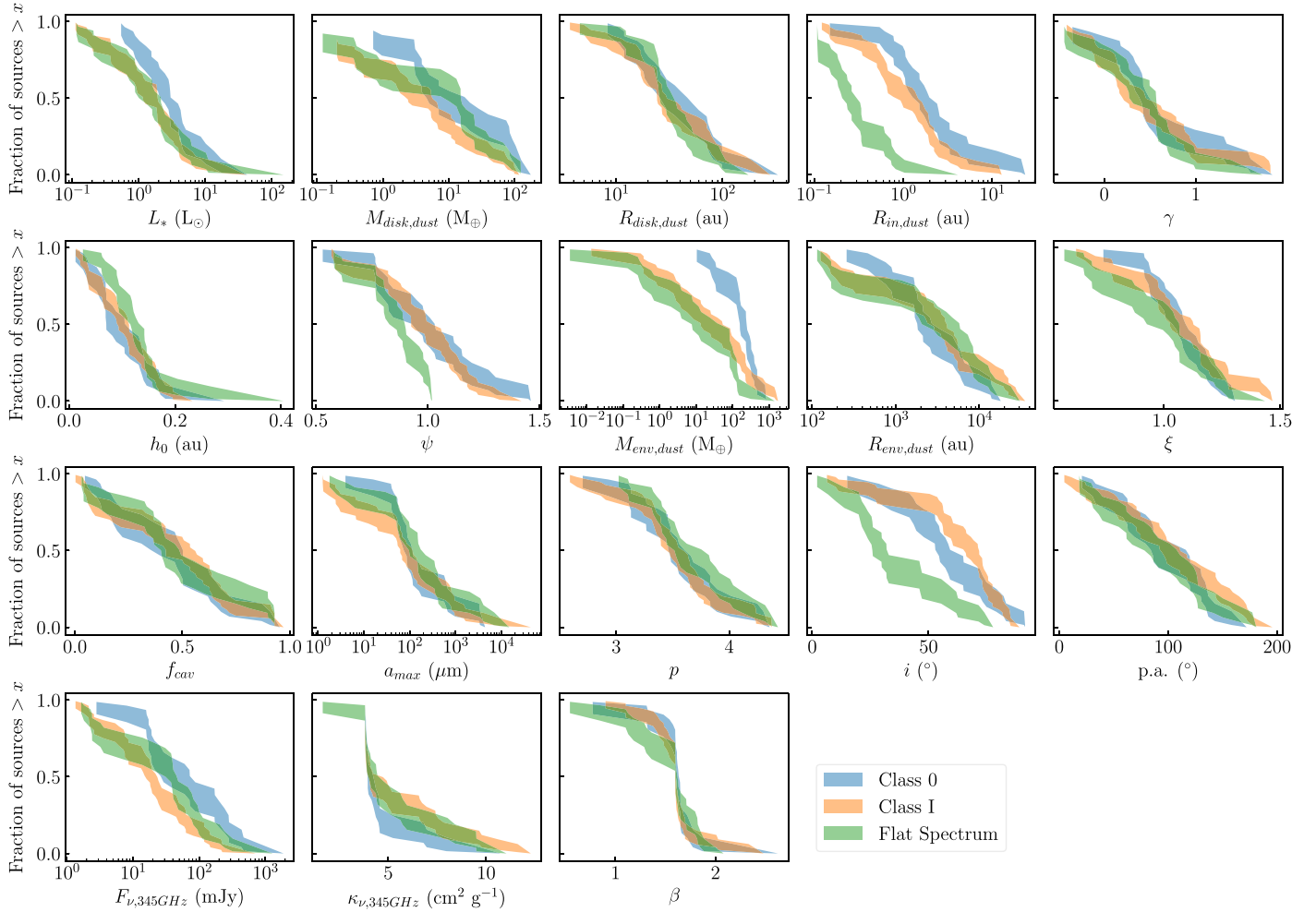
To test whether there are any significant differences in the distribution of parameters between the separate classes, we run two-sample tests using the log-rank test in the `lifelines` package to compare the distribution of parameter values of each of the three classes against each other. To account for uncertainties in the measured parameters, instead of using the “best-fit” parameters, we create 1000 realizations of the source parameters by randomly selecting parameters from the posterior for each source for each realization. We then calculate the  $p$ -value associated with the two-sample test for each realization and consider the fraction of times for which the test reported a statistically significant difference in the distributions ( $p < 0.05$ ). The results of this calculation are reported in the last three columns of Table 2, with comparisons that we consider significant shown in bold.

For most parameters, we find that the distributions for each class are consistent with being drawn from the same underlying distribution. One exception to this is that flat-spectrum sources have lower envelope dust masses than Class 0 sources (100% of realizations with  $p < 0.05$ ), which is what might be expected if the Class 0/I/flat-spectrum scheme represents an evolutionary sequence. There are some realizations where Class I

sources have lower envelope dust masses than Class 0 sources (15% of realizations with  $p < 0.05$ ) and flat-spectrum sources have lower envelope dust masses than Class I sources (6% of realizations with  $p < 0.05$ ), but these are not significant. Furthermore, we find some evidence that Class I sources have disk dust masses that are lower than Class 0 sources (69% of realizations with  $p < 0.05$ ). We cannot, however, confidently distinguish Class 0 disk dust masses from flat-spectrum disk dust masses or Class I disk dust masses from flat-spectrum disk dust masses (0% of realizations with  $p < 0.05$ ).

Our inability to distinguish between Class 0 and flat-spectrum disk dust masses is in contrast with Tobin et al. (2020a), who found that their masses were drawn from different underlying distributions. This appears to be due to a knee in the cumulative distribution of flat-spectrum disk dust masses, where the flat-spectrum distribution actually crosses the Class 0 distribution.

Interestingly, we find that flat-spectrum sources preferentially have lower inclinations than Class 0 or I sources ( $>99\%$  of realizations with  $p < 0.05$ ). If the classification of a source was purely determined by its evolutionary stage, we would not expect the inclination to be significantly different from class to class, as we find here. One potential reason for this difference could be that sources with lower inclination have lower extinction to the hotter central regions of the disk and the protostar itself, and so we see more near-infrared emission from these sources. As such, they are more likely to be classified as flat spectrum. Conversely, higher-inclination sources will have more near-infrared extinction, and so they are more likely to be classified as Class 0/I.



**Figure 3.** Cumulative distributions showing the fraction of protostellar disks with a measured parameter value above a given value. We separately show the cumulative distributions for sources identified as Class 0 (blue), Class I (orange), and flat spectrum (green) to highlight any potential differences between populations at different evolutionary stages.

Flat-spectrum sources also have  $\psi$  and  $R_{\text{in}}$  values that are smaller than those of Class 0 or I sources ( $>93\%$  of realizations with  $p < 0.05$ ). It is possible that this may point to underlying physical differences. For example, the difference in  $\psi$  may indicate that flat-spectrum disks are flatter, which may in turn hint that large dust grains have settled more in older disks. These differences may instead, however, be related to systematics in our relatively simple model. As flat-spectrum sources tend to have more near-infrared emission, they may require smaller inner dust radii to ensure that there is hot material to produce this emission, and flatter disks to ensure that material is not obscured by optical depth effects. Higher-resolution ALMA observations and/or more detailed models may be needed to better determine whether these are not simply systematics of our model.

## 5. Discussion

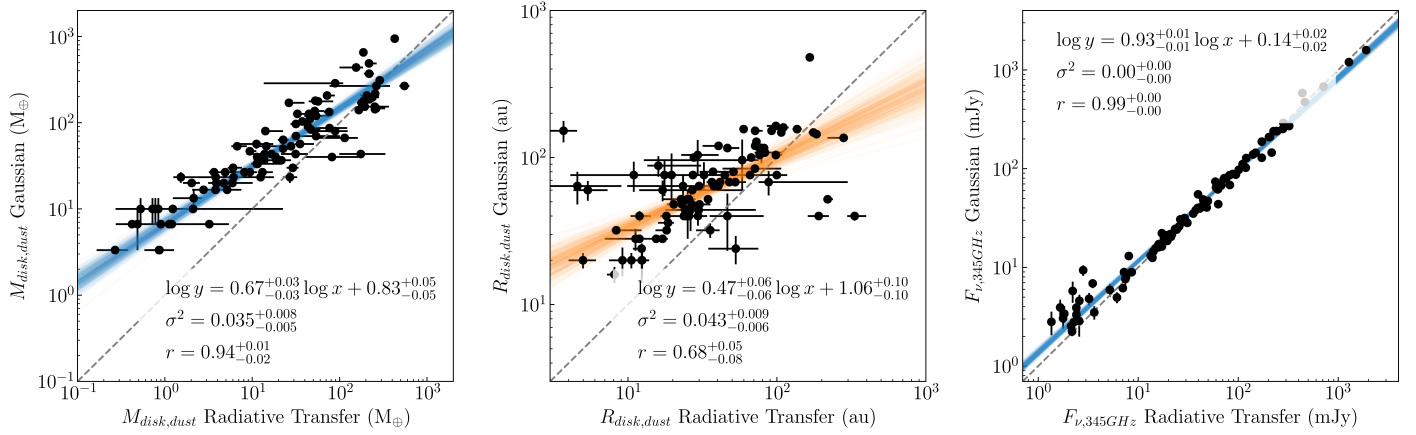
### 5.1. Comparison with Simple Estimators of Disk Structure

Disk surveys have traditionally relied on simple estimates to measure disk properties owing, in large part, to the complexity and computational costs involved in doing radiative transfer modeling of protostellar and protoplanetary disks. For example, disk dust masses are often calculated from the submillimeter flux by assuming optically thin dust emission (Hildebrand 1983)

with a uniform temperature either of 20 K (e.g., Ansdell et al. 2016; Pascucci et al. 2016) or that scales with protostellar luminosity (e.g., Andrews et al. 2013; Tobin et al. 2020a) and an estimate of the dust opacity at the relevant submillimeter wavelength (e.g., Beckwith et al. 1990). Disk sizes are typically calculated by fitting images with two-dimensional Gaussian functions (e.g., Tobin et al. 2020a), by fitting with Nuker profiles (e.g., Tripathi et al. 2017) or by using the curve-of-growth method to calculate the radius enclosing 95% of the total flux (e.g., Ansdell et al. 2016).

These simple estimates of disk dust masses and radii are useful for providing quick measurements of disk properties for large samples of disks without the need for the significant computational resources we use in this work. These estimates are, however, also limited by the fact that disk temperatures and densities are inherently three-dimensional, coupled with optical depth effects, the viewing angle of the system, and the scattering of light by dust grains (e.g. Liu 2019; Zhu et al. 2019). These difficulties are compounded when considering protostellar disks, which are embedded within an envelope of infalling dusty material that can also be bright in the submillimeter, making uniquely identifying the disk difficult. Some previous works have attempted to remove envelope emission by considering only the flux at a specific interferometer baseline that is thought to be large enough to resolve





**Figure 4.** Comparisons between the disk dust masses (left), radii (middle), and 345 GHz fluxes (right) measured by our radiative transfer modeling and as measured by fitting a two-dimensional Gaussian fit to the VANDAM: Orion images of our sources. We also show the results of a loglinear fit to each data set and list the best-fit parameter values, as well as show a dashed line where the values on each axis are equal for comparison.

out typical size scales of envelopes (e.g., Jørgensen et al. 2009; Andersen et al. 2019), though even those estimates can still be subject to significant envelope contamination (e.g., Dunham et al. 2014). Other works have extended this by fitting simple analytic disk intensity models directly to the visibility data, sometimes even including an envelope component to help separate disk from larger-scale emission, (e.g., Harsono et al. 2014; Maury et al. 2014; Segura-Cox et al. 2016, 2018; Maury et al. 2019), but are still limited by the relatively simple conversion from flux to dust mass.

Tobin et al. (2020a) presented estimates of protostellar disk dust masses and radii for sources in the VANDAM: Orion survey by assuming that the emission fit in the image plane with a two-dimensional Gaussian function could be attributed uniquely to emission from a protostellar disk. They used the typical conversion from submillimeter flux to dust mass,

$$M_d = \frac{d^2 F_{\nu}}{\kappa_{\nu} B_{\nu}(T)} \quad (9)$$

(e.g., Hildebrand 1983), assuming a 345 GHz opacity of  $1.84 \text{ cm}^2 \text{ g}^{-1}$  and a temperature that was scaled based on the protostellar luminosity,  $T = 43 \text{ K} (L/L_{\odot})^{0.25}$ . This temperature scaling was derived from a grid of radiative transfer models, in an effort to account for the effects of protostellar luminosity on the disks average temperature. They also reported disk sizes using the  $2\sigma$ , deconvolved, major-axis size from two-dimensional Gaussian fits to each source in the image plane.

As we have now modeled 97 of the sources from Tobin et al. (2020a) using our disk radiative transfer modeling infrastructure, we can directly test how these simple calculations of protostellar disk properties compare with models that account for more physics in deriving properties. In Figure 4, we show a comparison of the disk dust masses and radii, as measured by our radiative transfer modeling framework, compared with the dust masses and radii presented in Tobin et al. (2020a). For reference, we also show the best-fit loglinear model for a fit to each set of data using the `linmix` package,<sup>7</sup> which enables us to account for errors on multiple independent variables (Kelly 2007), along with the  $x = y$  line as a dashed line.

We find that the disk dust masses derived from our modeling differ substantially from the dust masses found with simple

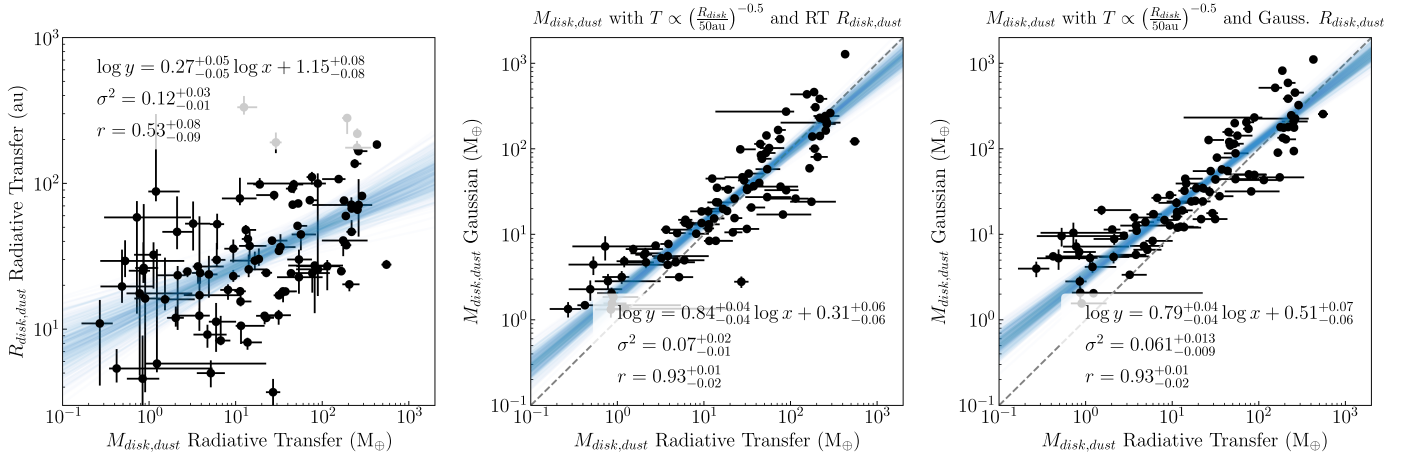
estimates from the flux. While there is a strong loglinear correlation between  $M_{\text{disk,Gaussian}}$  and  $M_{\text{disk,RT}}$ , the disk dust masses measured in Tobin et al. (2020a) overestimate the mass compared to radiative transfer modeling, with the severity of the discrepancy increasing for lower-mass disks.

To test whether this discrepancy is the result of improperly separating disk and envelope, we also compare the 345 GHz flux inferred from our best-fit disk model for each source with the 345 GHz flux measured in Tobin et al. (2020a) in the rightmost panel of Figure 4. From a loglinear fit, we find a tight correlation with a slope very close to 1, indicating that there is very good agreement between the fluxes measured in Tobin et al. (2020a) and the fluxes we recover from our modeling. Though the slope is nominally different from 1 with high significance, this appears to be due to a handful of faint sources for which the Gaussian fit overestimated the flux. If we only fit sources with  $F_{\nu,345\text{GHz,RT}} > 5 \text{ mJy}$ , we find that the slope is perfectly consistent with 1.

This tight one-to-one correspondence between the flux measured from the different methods suggests that the discrepancy between disk dust masses measured from the different methods is not due to difficulty separating disk from envelope. Though this may still be a problem for lower-resolution observations, with the resolution of  $\sim 30\text{--}40 \text{ au}$  that we have here, we can confidently identify the disk flux from millimeter images with a two-dimensional Gaussian fit. The good agreement between the fluxes measured in different ways therefore instead indicates that the difference is in the conversion from millimeter flux to dust mass. In particular, this suggests that there may be physics missing from the simple estimates presented in Equation (9), even with the modifications to the temperature from Tobin et al. (2020a).

That said, Tobin et al. (2020a) also found from the radiative transfer modeling on which their luminosity correction was based that the average disk temperature should scale like  $R^{-0.5}$ , though they ultimately used a fiducial radius of  $50 \text{ au}$  in their calculations for simplicity and to avoid reliance on possibly inaccurate radii from Gaussian fits. We do, however, find that there is a correlation between  $M_{\text{disk,dust}}$  and  $R_{\text{disk,dust}}$  in our sample, as we show in Figure 5, so it is possible that this correction is indeed important. If we recalculate dust masses from the ALMA fluxes while including the  $R^{-0.5}$  scaling in calculating the temperature, using  $R_{\text{disk,dust}}$  from our modeling, as shown in the middle panel of Figure 5, we find better

<sup>7</sup> <https://github.com/jmeyers314/linmix>



**Figure 5.** Left: comparison of the disk dust masses and radii measured using our radiative transfer modeling framework, demonstrating that the two are correlated, though with significant scatter. Middle: a comparison of the disk dust masses found from our modeling with disk dust masses derived from fitting the disk with a two-dimensional Gaussian in the image plane (e.g., Tobin et al. 2020a) and using Equation (9) with  $T = 43 \text{ K} (R_{\text{disk}}/50 \text{ au})^{-0.5} (L/1 L_{\odot})^{0.25}$  and  $R_{\text{disk}}$  measured from our modeling. Right: same as in the middle panel, but using  $R_{\text{disk}}$  measured in Tobin et al. (2020a). We also show the results of a loglinear fit to each data set and list the best-fit parameter values, as well as show a dashed line where the values on each axis are equal for comparison.

agreement with our radiative-transfer-derived disk dust masses. The best-fit line still has a slope shallower than 1, but this may again be due to the discrepancy between fluxes for the faintest sources, as these are presumably generally the lowest-mass disks. Above  $\sim 5 M_{\oplus}$  the slope appears to be closer to 1. There is still a significant amount of scatter in the relationship, though, likely because we are marginalizing over many disk parameters that contribute to radiative transfer effects. As such, while this correction helps to reproduce the radiative-transfer-modeling-measured distribution of disk dust masses on average, individual disks may still have significant differences from their radiative-transfer-measured mass. In the right panel of Figure 5 we show the same comparison but using  $R_{\text{disk,dust}}$  measured from Gaussian fitting, as this is more analogous to what would be done by an observer not using radiative transfer modeling. We find that this also reduces the discrepancy between the two mass measurements, though the slope is still shallower than 1 and is not corrected quite as well as in the case of using the radiative-transfer-modeling-measured radii. This is likely because of the discrepancy between the disk radii we measure and those found from Gaussian fitting, as can be seen in Figure 4 and discussed below.

Ultimately, we suggest that future works use the relation

$$T = 43 \text{ K} \left( \frac{L_{*}}{1 L_{\odot}} \right)^{0.25} \left( \frac{R_{\text{disk}}}{50 \text{ au}} \right)^{-0.5} \quad (10)$$

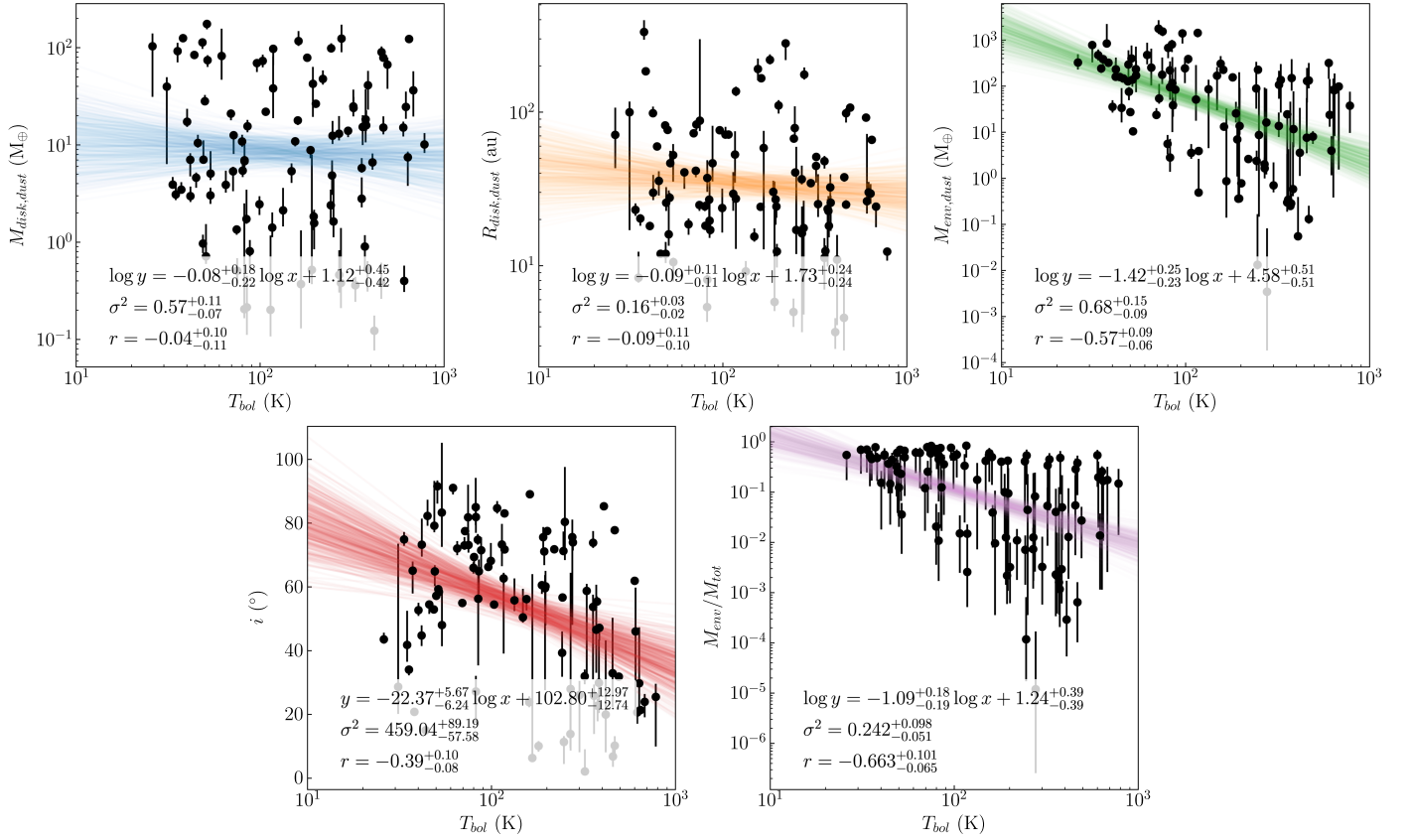
to calculate the average disk temperature, as it includes additional physical effects in simple estimates of disk dust masses. We do caution, however, that this relationship should only be applied to *embedded* disks, as the models it is based on included an envelope that can serve to keep the disk warmer than it would be otherwise. For protoplanetary disks, the temperature pre-factor may be different, though the scaling with luminosity and disk radius would likely be similar. We also note that these corrections are model dependent and could still produce incorrect disk dust masses if the assumptions underlying the model are also incorrect. That said, the particular corrections applied here, that brighter stars produce warmer disks and that smaller disks with more material close to

the star are on average warmer, should be reasonably model agnostic.

We also find that the disk dust radii measured by the two-dimensional Gaussian fits also frequently disagree with the disk dust radii measured from our modeling. The Gaussian fits tend to more severely overestimate disk radii for smaller disks, while they actually *underestimate* the sizes of a number of the largest disks, some by quite a lot (see Figure 4). The discrepancy between radiative-transfer-measured disk radii and Gaussian-fitting-measured disk radii is likely due to the low spatial resolution, of only  $\sim 40 \text{ au}$ . As such, smaller disk radii are increasingly difficult to accurately deconvolve from the beam. Additionally, disks with massive envelopes may appear larger in the image plane, though the strong correlation between fluxes measured from our modeling and from two-dimensional Gaussian fitting suggests that this may not be a major effect. On the other hand, for larger disks, two-dimensional Gaussian functions become increasingly poor representations of the resolved disk brightness profile and often tend to fit the more compact central peak rather than the extended, low surface brightness disk structure. By fitting directly in the  $u$ - $v$  plane, we avoid the difficulties of measuring deconvolved sizes, and the more flexible geometry of our disk model allows us to better match the full brightness profiles of the disks in our sample.

## 5.2. Protostellar Disk and Envelope Evolutionary Trends

In Section 4, we discussed how the distributions of disk properties from our best-fit models change as a function of the protostellar classification, but for the most part we found only marginal evidence that the distributions were distinct for different classes. In this section, we further examine trends as a function of protostellar “age” by comparing a selection of disk and envelope properties with three separate quantities that have been proposed as tracers of protostellar evolution. In particular, we consider the bolometric temperature ( $T_{\text{bol}}$ ), the ratio of envelope mass to total mass ( $M_{\text{env}}$ ), and the age as inferred from simple models of protostellar evolution (e.g., Andre et al. 1994; Saraceno et al. 1996; André et al. 2008; Molinari et al. 2008; Fischer et al. 2017). For each comparison, we run a loglinear fit to the data using the `linmix` package, which



**Figure 6.** Distributions of a sample of parameters measured from our radiative transfer modeling framework compared with the bolometric temperature measured for each source. We also show the best-fit loglinear, or semi-loglinear in the case of inclination, model for each comparison by plotting 100 samples drawn from the posterior of an MCMC fit, and we list the best-fit parameters, slope, intercept, and variance for the fit in each panel, along with the correlation coefficient.

accounts for errors on both variables being fit (Kelly 2007), in order to test whether there is a trend with protostellar age in each quantity.

### 5.2.1. Bolometric Temperature ( $T_{\text{bol}}$ )

Protostars have traditionally been classified into Class 0/I/II/III and flat-spectrum sources according to their bolometric temperature (e.g., Chen et al. 1995) or near-infrared spectral index (e.g., Myers et al. 1987). Though these classifications have been defined observationally, they are also thought to roughly follow the evolutionary state of the young star through the effect the envelope has on the SED of a protostar (e.g., Whitney et al. 2003; Crapsi et al. 2008). Sources with low bolometric temperatures and (steeply) rising SEDs in the infrared are best explained by significant dust obscuration, presumably by a massive envelope of material while they are still young. Conversely, sources with higher bolometric temperatures or SEDs that are shallower in the infrared or even decreasing likely do not suffer from as much extinction and therefore do not have as substantial envelope obscuration and are older. Properties such as the bolometric temperature or the classification as Class 0/I/flat spectrum have therefore sometimes been treated as estimators of the age of a protostellar system (e.g., Evans et al. 2009; Dunham et al. 2015; Kristensen & Dunham 2018; Andersen et al. 2019).

The comparison of disk properties with  $T_{\text{bol}}$  is shown in Figure 6 and largely reinforces the trends, or lack thereof, found in Section 4. Disk dust mass and radius show no statistically significant correlation with  $T_{\text{bol}}$ . While we do see

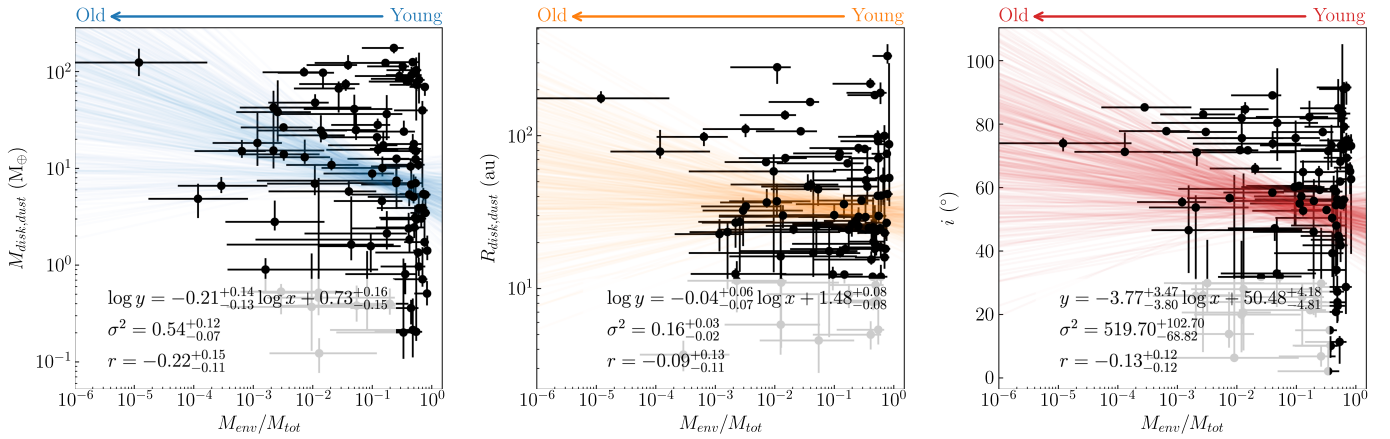
that envelope dust mass decreases with  $T_{\text{bol}}$ , there remain a number of sources with large envelope dust masses for larger values of  $T_{\text{bol}}$ , particularly when compared with their disk dust mass. And though the relative importance of the envelope as compared with the disk, as measured by the ratio of their masses,  $M_{\text{env,dust}}/M_{\text{disk,dust}}$ , does decrease with increasing  $T_{\text{bol}}$ , there remain sources with high  $T_{\text{bol}}$  that have very substantial envelopes, with  $M_{\text{env,dust}}/M_{\text{disk,dust}} > 10$  or low- $T_{\text{bol}}$  sources for which the importance of the envelope is perhaps less than expected ( $M_{\text{env,dust}}/M_{\text{disk,dust}} < 1$ ).

Finally, as was discussed in Section 4, we find a statistically significant dependence of source inclination on bolometric temperature. As a whole, these results would seem to suggest that classifications of young stars that rely on SED diagnostics such as  $T_{\text{bol}}$  may to some degree trace protostellar evolution but are also contaminated by difficult-to-disentangle radiative transfer effects. Though this has, to some degree, always been known to be a difficulty of the classification scheme (e.g., Calvet et al. 1994; Chiang & Goldreich 1999; Furlan et al. 2016), this work provides the first clear demonstration of this for a large sample of protostars.

### 5.2.2. Ratio of Envelope Mass to Total Mass ( $M_{\text{env}}/M_{\text{tot}}$ )

In response to the degeneracies involved in relating SED-derived properties back to underlying physical properties, Robitaille et al. (2006) instead suggested using the physical properties of a system as a more reliable method for estimating the evolutionary stage of protostars. One would naively expect that as the natal cloud collapses and material from the envelope





**Figure 7.** Distributions of a sample of parameters measured from our radiative transfer modeling framework compared with the ratio of envelope mass to total mass measured for each source. We also show the best-fit loglinear, or semi-loglinear in the case of inclination, model for each comparison by plotting 100 samples drawn from the posterior of an MCMC fit, and we list the best-fit parameters, slope, intercept, and variance for the fit in each panel, along with the correlation coefficient.

accretes onto the disk and star, the amount of matter present in each component (envelope, star, and disk) should evolve with time. Initially, the bulk of the system material should be present in the envelope, but over time the mass of the central protostar and the disk should increase while the mass of the envelope should decrease.

This ratio may not be a perfect representation of the age of a protostellar system;  $M_{\text{env}}$  (and therefore  $M_{\text{env}}/M_{\text{tot}}$ ) may not decrease linearly, and possibly not even monotonically, with time. Simulations of global cloud collapse demonstrate that the star and disk formation process can be significantly impacted by environment, including the late accretion of envelope material onto the disk and star (e.g., Bate 2018; Kuffmeier et al. 2018, 2020a) and how the outflow interacts with envelope material (e.g., Offner & Arce 2014). Nonetheless,  $M_{\text{env}}/M_{\text{tot}}$  should roughly track the age of the system as the material in the envelope depletes and the central protostar becomes less and less obscured. More importantly, it should more exactly measure the physical properties of the system that quantities such as  $T_{\text{bol}}$  are attempting to emulate, namely, the “embeddedness” of the system or how substantial of an envelope the system has.

$M_{\text{env}}/M_{\text{tot}}$  has not traditionally been used as a tracer of protostellar evolution, however, largely because of the difficulty of measuring disk and envelope masses compared with the ease of measuring bolometric temperatures and near-to mid-infrared spectral indices. However, with our detailed physical modeling we are able to measure disk and envelope masses for a sizable population of protostars.

In Figure 7, we show how protostellar disk and envelope properties scale with  $M_{\text{env}}/M_{\text{tot}}$ . We remind readers that because stellar masses are inherently gas masses, we use the total (gas + dust) masses assuming a gas-to-dust ratio of 100 to calculate this ratio. Though protostellar masses have been measured kinematically for a few sources in our sample (e.g., Tobin et al. 2020b), they are unknown for the vast majority of our sample owing to the difficulty of detecting photospheric properties of embedded sources through the heavy extinction from their envelopes (e.g., Greene et al. 2018). Instead, to properly and probabilistically account for the unknown masses of the majority of the sources in our sample, for each source we match the posterior distributions from our radiative transfer modeling with a sample of masses drawn from the full Chabrier

initial mass function (IMF; Chabrier 2003). As such, the error bars shown in Figure 7 should account for the large uncertainty in the stellar masses, using the IMF as a prior. We also assume a gas-to-dust ratio of 100 to convert the dust masses that we measure here to total masses.

We also show the results of loglinear fits to the data in Figure 7 to investigate whether there are any statistically significant trends in disk properties with protostellar evolution. As the VANDAM: Orion sample was selected on the basis of  $T_{\text{bol}}$  and not  $M_{\text{env}}/M_{\text{tot}}$ , it is likely that below some value of  $M_{\text{env}}/M_{\text{tot}}$  our sample is not complete. As such, we exclude sources with  $M_{\text{env}}/M_{\text{tot}} < 0.01$  from our fits, as these sources appear to be older sources with high inclinations that led to them being classified as protostars. As it is unclear where the appropriate boundary in  $M_{\text{env}}/M_{\text{tot}}$  should lie, we have explored more stringent cuts in  $M_{\text{env}}/M_{\text{tot}}$  and find that our fits do not change substantially.

In contrast with the comparison with  $T_{\text{bol}}$ , we find that there is no statistically significant correlation between inclination and  $M_{\text{env}}/M_{\text{tot}}$ , which is to be expected, as there is no reason to believe that sources at differing stages of evolution should have differing inclination distributions. Likewise, for most parameters of our model we find no statistically significant correlation with  $M_{\text{env}}/M_{\text{tot}}$ .

We do find hints, at the not quite  $2\sigma$  significant level, of a trend of increasing  $M_{\text{disk,dust}}$  as  $M_{\text{env}}/M_{\text{tot}}$  decreases ( $r = -0.22^{+0.15}_{-0.11}$ ), regardless of exact cutoff value for the  $M_{\text{env}}/M_{\text{tot}}$  we choose. Though the trend is somewhat weak, this may be an indication that disk dust mass *increases* with time during the early stages of star formation as material from the envelope continues to accrete. We note that this is opposite the weak trends found by Tobin et al. (2020a) or Andersen et al. (2019), who used  $T_{\text{bol}}$  as a proxy for evolution. Still, it is important to note that this trend is weak, if it exists at all, and exhibits a large amount of scatter. Modeling of a larger sample of protostellar disks may help to better determine whether the trend is real.

We again see no trend in disk dust radius with  $M_{\text{env}}/M_{\text{tot}}$ . This lack of a clear trend of increasing disk size with time might underscore the effects of dust radial drift relative to the gas (e.g., Weidenschilling 1977; Birnstiel et al. 2010), which may wash out the growth of the gas disk with time in observations tracing dust. Alternatively, or perhaps in addition,



weak or no growth of protostellar disks with time may hint at the importance of magnetic fields in regulating the angular momentum of infalling material and preventing large disks from forming (e.g., Wurster et al. 2019; Hennebelle et al. 2020). Observations of the sizes of gas disks, difficult as they may be, might be needed to help further illuminate which physics regulates the sizes and evolution of young disks.

It is important to note that while the IMF is an appropriate representation of the final masses of stars, it may not be an appropriate distribution to represent the masses of protostars in our sample that are young and still forming. It would be ideal, instead, to sample from a protostellar mass function (PMF) that accounts for the masses of these protostars as they form (e.g., McKee & Offner 2010), or better yet to use directly measured dynamical masses (e.g., Tobin et al. 2012, 2020b) for the sample. In the absence of either, however, we hope that the use of the IMF enables a reasonable estimate for the protostellar mass with broad error bars to account for and propagate the large uncertainty in the estimate. Nonetheless, the lack of measured masses and our use of the IMF in their place could have an impact on the correlations, or lack thereof, found here. Future efforts to provide more realistic estimates of the sources in this sample are crucial for refining this analysis to search for evolutionary trends.

### 5.2.3. A Simple Model of Protostellar Evolution

Finally, we also consider a simple model of protostellar evolution (e.g., Andre & Montmerle 1994; Saraceno et al. 1996; André et al. 2008; Fischer et al. 2017) to produce a set of evolutionary tracks in the  $L_{\text{bol}} - M_{\text{env}}$  plane and compare the tracks with those values as inferred from our modeling. The model uses a simple prescription for how the envelope mass, stellar mass, and mass accretion rate evolve with time and then ties to pre-main-sequence evolutionary tracks to estimate properties of the protostar such as radius or luminosity.

To generate the tracks, we follow the prescription of André et al. (2008). For a given final stellar mass,  $M_{\text{star}}$ , we assume that the star is formed from a core with mass  $M_{\text{env},0} = M_{\text{star}}/\epsilon$ , where  $\epsilon$  is the star formation efficiency of individual cores. At any given time, the instantaneous mass accretion rate is given by  $\dot{M}_{\text{acc}} = \epsilon M_{\text{env}}/\tau$ , where  $\tau$  is the characteristic timescale for protostellar evolution. We create an array of times from  $t = 0$  to  $t = 10$  Myr, spaced logarithmically. At any given time,  $t_i$ , the mass of the star, the mass of the envelope, and the accretion rate can be calculated from the masses and accretion rate at the previous time step from

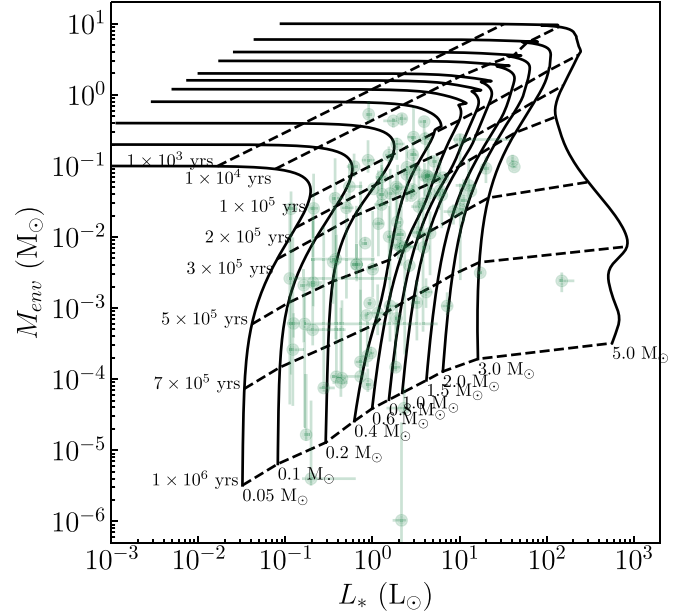
$$M_{\text{star},i} = M_{\text{star},i-1} + \dot{M}_{\text{acc},i-1}(t_i - t_{i-1}), \quad (11)$$

$$M_{\text{env},i} = M_{\text{env},i-1} - \dot{M}_{\text{acc},i-1}/\epsilon(t_i - t_{i-1}), \quad (12)$$

$$\dot{M}_{\text{acc},i} = \epsilon M_{\text{env},i}/\tau. \quad (13)$$

Following André et al. (2008), we assume that  $\epsilon = 0.5$  (e.g., Matzner & McKee 2000; McKee & Tan 2002) and  $\tau = 1 \times 10^5$  yr.

To determine the total luminosity at any given time, we calculate the contributions of the accretion luminosity and protostellar luminosity ( $L_{\text{star}} = L_{\text{acc}} + L_{\text{star}}$ ). In this way,  $L_{\text{star}}$  is analogous to the same parameter in our radiative transfer models, which accounts for the total luminosity that is reprocessed and reemitted across the entire spectrum, as discussed at the beginning of Section 3. The accretion



**Figure 8.** A demonstration of the  $L_{\text{star}} - M_{\text{env}}$  evolutionary tracks described in Section 5.2.3. The protostellar evolutionary tracks are shown in black, with the curves of a constant final mass shown as solid lines and isochrones shown as dashed lines. The sources in our modeled sample are shown as green points.

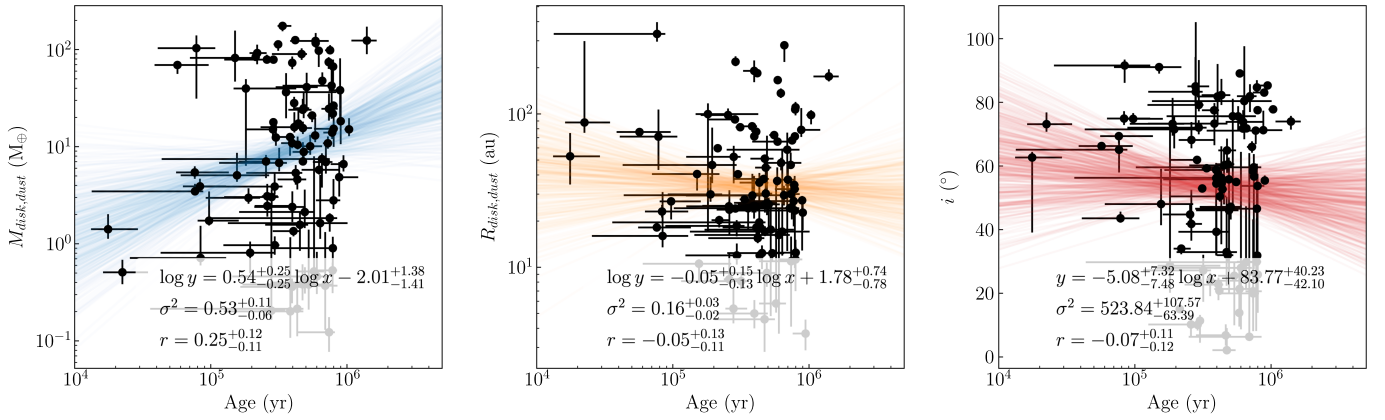
luminosity at any given time is given by

$$L_{\text{acc}} = \eta \frac{GM_{\text{star}}\dot{M}_{\text{acc}}}{R_{\text{star}}}. \quad (14)$$

In this equation,  $M_{\text{star}}$  and  $\dot{M}_{\text{acc}}$  are calculated in Equations (11)–(13),  $G$  is the gravitational constant, and we use  $\eta = 1$  in line with André et al. (2008). To calculate  $R_{\text{star}}$  and  $L_{\text{star}}$  at any given time, we use the evolutionary tracks from (Feiden 2016) from the current mass ( $M_{\text{star},i}$ ) and age ( $t_i$ ). When stars have mass  $< 0.1 M_{\odot}$ , below the lower limit of the evolutionary tracks and also of hydrogen burning, we use the scaling relationships found by Fischer et al. (2017,  $R_{\text{star}} \propto M_{\text{star}}^{0.34}$  and  $L_{\text{star}} \propto M_{\text{star}}^{1.34}$ ) to scale the evolutionary tracks from  $0.1 M_{\odot}$  to the relevant stellar mass. Though approximate, this should give a rough idea of the properties of stars that fall outside of the boundaries of the evolutionary tracks.

With this setup, we calculate evolutionary tracks in the  $L_{\text{bol}} - M_{\text{env}}$  plane for a range of protostar masses from  $0.01$  to  $5 M_{\odot}$ . We show these evolutionary tracks and how the sources in our sample compare in Figure 8. As these tracks involve stellar mass, we again use the total envelope mass, assuming a gas-to-dust ratio of 100, from our modeling for comparison with the tracks. Our tracks are qualitatively similar to previous works (e.g., Andre & Montmerle 1994; Saraceno et al. 1996; André et al. 2008; Fischer et al. 2017), though we note that their exact values may differ from work to work owing to the slightly different assumptions used by each study.

To estimate the age of the protostars in our sample with these evolutionary tracks, we use two-dimensional linear interpolation to map any given  $(L_{\text{star}}, M_{\text{env}})$  pair onto the tracks and thereby infer the age, as demonstrated in Figure 8. For each source, we repeat this process for every sample point in the posterior distribution to produce a distribution of ages for the source. The best-fit age of the source is then determined as the



**Figure 9.** Distributions of a sample of parameters measured from our radiative transfer modeling framework compared with the age measured for each source using the simple protostar evolutionary tracks shown in Figure 8. We also show the best-fit loglinear, or semi-loglinear in the case of inclination, model for each comparison by plotting 100 samples drawn from the posterior of an MCMC fit, and we list the best-fit parameters, slope, intercept, and variance for the fit in each panel, along with the correlation coefficient.

peak of the posterior distribution of ages, with the uncertainties from the 95% confidence interval around those ages.

We show how a selection of protostellar disk and envelope properties, as inferred from our modeling, compare with the age inferred from these evolutionary tracks in Figure 9. We also show the results of a loglinear fit to the data in each comparison. To ensure that the comparison is not affected by completeness of the sample in age, as the sample was selected based on  $T_{\text{bol}}$ , we only consider sources with inferred ages  $>10^{4.5}$  yr to exclude sources that are outliers in age, though we find that the choice of cutoff does not significantly affect our results.

The results presented in Figure 9 are very similar to what we found when using  $M_{\text{env}}/M_{\text{tot}}$  as an evolutionary tracer. We find no statistically significant correlation between inclination and age inferred from the evolutionary tracks, which is, again, sensible, as there is no reason to believe that sources of different ages should have preferentially different inclinations. We also again find that the disk dust mass actually increases with time, albeit with a very large spread, suggesting that the dust mass of the disk tends to grow as material continues to fall onto the disk from the envelope. We do note that the correlation is only slightly more than  $\sim 2\sigma$  significant ( $r = 0.25^{+0.12}_{-0.11}$ ), though, so it would be useful to add additional sources to this analysis to confirm the trend with higher significance. Finally, we again do not find a statistically significant trend in disk dust radius as a function of age, indicating little or no growth of the disk with time, though whether this lack of growth is due to radial drift, magnetic fields sapping angular momentum of infalling material, or something else remains open.

We do caution, however, that the ages inferred from these evolutionary tracks should be treated as rough estimates of the ages of these sources and not absolute values. The model used to generate the tracks is quite simple and ignores the potentially complicated accretion history of protostellar envelopes. Moreover, it does not account for the possibility of episodic accretion (e.g., Kenyon et al. 1990; Dunham et al. 2010). Outbursts in young sources are known to occur (e.g., Safron et al. 2015; Yoo et al. 2017; Lee et al. 2020), but it is, for the most part, unknown whether any of the sources in our sample could be in an outburst state. If any were, they would appear more luminous than these tracks would predict, and therefore

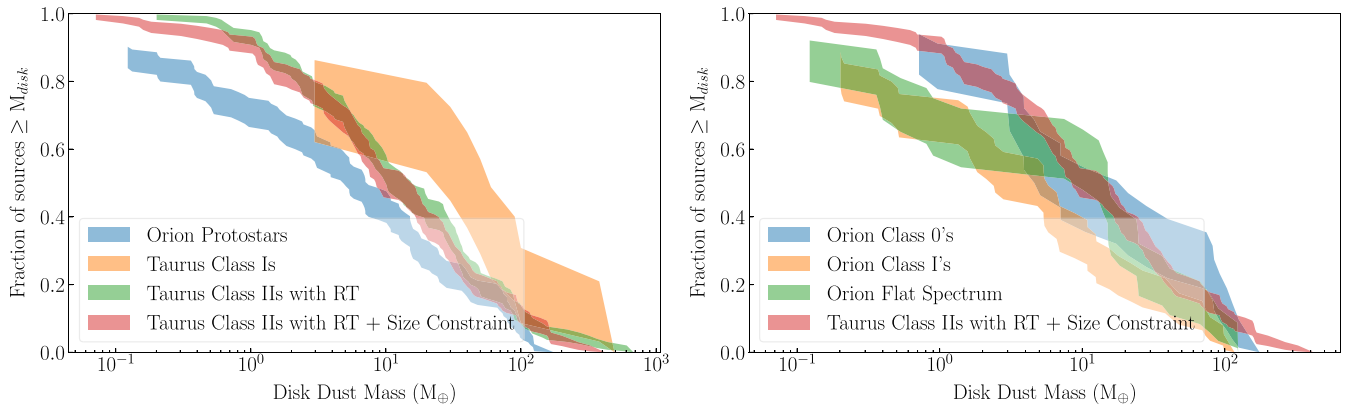
we would infer the wrong age. Nonetheless, they should give some idea of the rough age of the systems in a physically motivated way.

### 5.3. Comparison with Other Star-forming Regions

Continuum surveys of protoplanetary disks have become routine with ALMA over the past decade, with a number of different star-forming regions encompassing a range of ages and environments having been surveyed (Ansdell et al. 2016; Barenfeld et al. 2016; Pascucci et al. 2016; Ansdell et al. 2018; Eisner et al. 2018; Cieza et al. 2019). Such surveys have dramatically improved our understanding of the bulk properties of protoplanetary disks, altered our view of the planet-forming potential of these disks (e.g., Najita & Kenyon 2014; Manara et al. 2018), have demonstrated that protoplanetary disk dust masses decrease with age (e.g., Pascucci et al. 2016; Ansdell et al. 2018; Williams et al. 2019) with some interesting exceptions (e.g., Cazzoletti et al. 2019), and have also found evidence that environment may play some role in disk evolution (e.g., Eisner et al. 2018; Ansdell et al. 2020; Van Terwisga et al. 2020).

Surveys have also begun to push toward younger, Class 0/I protostellar disks, including the VANDAM: Orion Survey studied here, but also as a part of the ODISEA Survey (Cieza et al. 2019), the MASSES Survey (Stephens et al. 2018), and the VANDAM: Perseus Survey (e.g., Tobin et al. 2016a) and subsequent ALMA follow-up (Tychoniec et al. 2020). These surveys have painted a picture of protostellar disks as more massive than protoplanetary disks, or at least brighter (e.g., Sheehan & Eisner 2017; Andersen et al. 2019; Williams et al. 2019; Tobin et al. 2020a), suggesting that they may better represent the initial mass budget in disks for forming planets (Sheehan & Eisner 2017; Tychoniec et al. 2020).

Tobin et al. (2020a) compared the disk dust masses and radii measured for the VANDAM: Orion survey using two-dimensional Gaussian fits with the results of existing millimeter surveys of protoplanetary disks employing similar methods in great detail. As we demonstrated in Section 5.1, the fluxes measured by Tobin et al. (2020a) are generally in good agreement with the disk fluxes derived by our modeling here, and as the surveys described above generally use similar methodologies to Tobin et al. (2020a), we opt not to rehash those comparisons here, as we would not expect those results to



**Figure 10.** Left: cumulative distributions of protostellar disk dust masses as found from the samples of this work, Sheehan & Eisner (2017), and Ballering & Eisner (2019) using radiative transfer modeling. The distributions from Ballering & Eisner (2019) correspond to the disk dust masses derived with and without using the disk mass–disk radius relation found by Tripathi et al. (2017) as a prior in the modeling. Right: cumulative distribution of protostellar disk dust masses from this work, split by classification, compared with the Class II disk dust mass distribution for Taurus, using the disk mass–disk radius relation as a size constraint in the modeling.

change substantially other than from our smaller sample size. Instead, we compare the results of our modeling with the results of two other studies, Sheehan & Eisner (2017) and Ballering & Eisner (2019), that have done similar radiative transfer modeling to account for more physics in deriving disk properties. As such, comparing with these surveys should provide the fairest comparison of disk properties across different star-forming regions and states of evolution.

Sheehan & Eisner (2017) employed the same radiative transfer framework we use here to model a sample of 10 Class 0/I protostars in Taurus observed with the Combined Array for Research in Millimeter-wave Astronomy (CARMA) and separate disk and envelope contributions. They found that the disk-only fluxes were on average greater than the fluxes of Class II disks by a factor of a few, though with only 10 sources the result was only significant at the  $2.5\sigma$  significance level. Ballering & Eisner (2019) used a similar radiative transfer modeling framework to model 132 Class II disks, also in the Taurus star-forming region, finding that the dust masses that they derived from their modeling were  $1\text{--}5\times$  higher than the dust masses found from the simple flux-to-mass conversion used more generally. Though their modeling did not include spatial information by fitting resolved millimeter observations directly, they did repeat their fits both including the disk mass–disk size relation measured by Tripathi et al. (2017) as a prior and also leaving the disk radius unconstrained. We also note that Ribas et al. (2020) found a similar result from an independent modeling analysis of a subset of 23 of these same disks.

In Figure 10, we show a comparison between the distributions of disk dust masses derived here and those in Sheehan & Eisner (2017) and Ballering & Eisner (2019). For the results from Ballering & Eisner (2019), we show the distribution both with and without the size constraint. To rigorously compare the distributions, we use the `lifelines` package to run a log-rank two-sample test. We find that the Orion protostellar disk dust masses can be statistically distinguished from both the Taurus Class I sources ( $p$ -value = 0.009) and the Taurus Class II sources whether a size constraint is included ( $p$ -value = 0.0003) or not ( $p$ -value =  $1.7 \times 10^{-6}$ ).

If we split the protostellar disks modeled here into their Class 0/I/flat-spectrum designations, we find that Class 0 disks have a distribution of disk dust masses that is consistent with that of

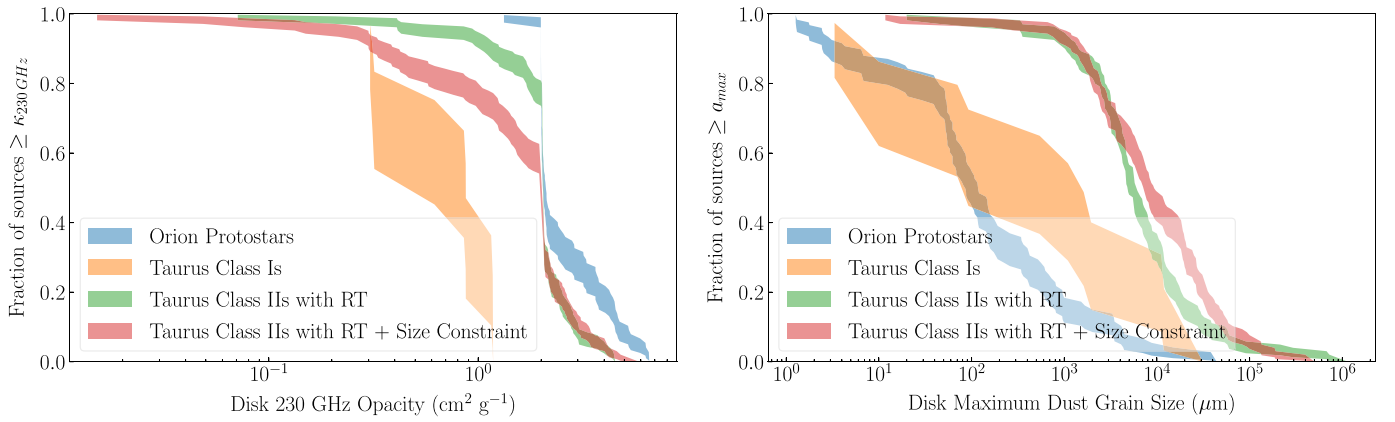
the Class II disks ( $p$ -value = 0.58), while both the Class I and flat-spectrum distributions are distinct ( $p$ -values of 0.0003 and 0.03, respectively). The Class I distribution, interestingly, mirrors the Class 0/Taurus Class II distributions at high dust masses but then has a knee that better follows the Class I disk dust masses.

Interestingly, and in contrast to previous works (e.g., Sheehan & Eisner 2017; Tychoniec et al. 2018; Andersen et al. 2019; Tobin et al. 2020a), we do not find a monotonic trend of disk dust masses decreasing from Class 0 to Class II stages. Instead, protostellar disks as a whole are less massive than Class II disks, and disk dust mass decreases from the Class 0 to Class I stages before seemingly rising again while progressing to the flat-spectrum and Class II stages. We do note that the latter effect may be mitigated by the issues with the Class 0/I/flat-spectrum system, as discussed in Section 5.2. This difference is the result of Class II disk dust masses from radiative transfer modeling being several times larger than dust masses from estimates from fluxes, in combination with the lower disk dust masses found in our radiative transfer modeling when compared with disk dust mass estimates from sub-millimeter fluxes.

It is important to note, however, that this comparison may not be entirely fair, as the embedded protostars in Taurus, a relatively low mass star-forming region, may be substantially different from those in Orion, which is substantially more massive. Indeed, if we compare the Taurus protostars with the Taurus Class II sources, we find that they cannot be statistically distinguished ( $p$ -values  $\sim 0.16$ ), though even this is in contrast with previous results owing to the larger Taurus disk dust masses when done with radiative transfer modeling. A more apt comparison for the disks in our sample may be Class II disks also in Orion. Though a comprehensive survey has not been published (S. E. van Terwisga et al. 2022, in preparation; but see also Grant et al. 2021, for a survey of Herschel-detected protoplanetary disks), particularly not with radiative transfer modeling, when comparing submillimeter fluxes the disks in the Orion Nebular Cluster (ONC) do seem to be lower mass than typical Class II disks (e.g., Eisner et al. 2018).

Another aspect of these comparisons that needs to be accounted for is the underlying protostellar/pre-main-sequence mass distribution. As disk dust masses are correlated with stellar masses (e.g., Andrews et al. 2013; Ansdell et al. 2016; Barenfeld et al. 2016; Pascucci et al. 2016), to properly





**Figure 11.** Left: cumulative distributions of protostellar disk 230 GHz dust opacity as found from the samples of this work, Sheehan & Eisner (2017), and Ballering & Eisner (2019) using radiative transfer modeling. The distributions from Ballering & Eisner (2019) correspond to the fits that did and did not use the disk mass–disk radius relation found by Tripathi et al. (2017) as a prior in the modeling. Right: same as on the left, but for maximum dust grain size.

compare disk dust masses, it is imperative to ensure that the disk masses are drawn from identical populations of stars. As stellar masses are known only for a handful of protostars in the VANDAM: Orion survey (e.g., Tobin et al. 2020b), we cannot meet this requirement and are instead forced to assume that it is met. But if the distributions are indeed drawn from differing underlying stellar mass distributions, then the comparisons would be invalid.

Finally, though each of these studies employed radiative transfer modeling to measure disk dust masses, there are differences that could potentially lead to systematic offsets between the properties measured. Though our work and Sheehan & Eisner (2017) employ the same code, we use slightly different dust opacities in our modeling and model interferometric data at a different wavelength (230 GHz vs. 345 GHz). On the other hand, we use similar dust opacities to Ballering & Eisner (2019), but their modeling did not directly model any imaging data to provide spatial constraints on the model. To test how much these differences in opacity may be playing a role, we show cumulative distributions for the 230 GHz dust opacity and maximum dust grain sizes in Figure 11. We do indeed find a fairly significant difference, of a factor of  $\sim 2\times$  between typical dust opacities from Sheehan & Eisner (2017) as compared with this work, which may help to explain the difference in dust masses. Though there are some small differences in the 230 GHz opacity between our work and Ballering & Eisner (2019), we overall have similar values, with the majority falling near the canonical opacity of  $\sim 2 \text{ cm}^2 \text{ g}^{-1}$ . That said, our sample has very different maximum dust grain sizes and therefore likely different broadband opacity curves that could influence the disk dust mass through the temperature structure and thereby hinder a fair comparison of the samples. On the other hand, these differences could reflect real differences in the properties of disks across differing evolutionary stages and environments. Ultimately, it would be ideal to model all of these samples, and more, in a uniform manner to ensure that disk properties are being compared in as fair a manner as possible.

Assuming, however, that our comparisons are fair, these results raise interesting questions about the potential for planet formation in disks. The onset of disk surveys with ALMA has raised questions about whether Class II disks have enough dust mass in them to form the population of planets that are being found by various exoplanet surveys (e.g., Najita &

Kenyon 2014; Manara et al. 2018). Early studies of Class 0/I/flat-spectrum disks suggested that they were, on average, more massive than Class II disks, suggesting that these young disks might be a better representation of the initial reservoir of material for forming planets (e.g., Sheehan & Eisner 2017; Tychoniec et al. 2018, 2020). In comparing the disk dust masses measured using detailed radiative transfer modeling, however, it is less clear that young disks are indeed more massive and more readily capable of forming planets.

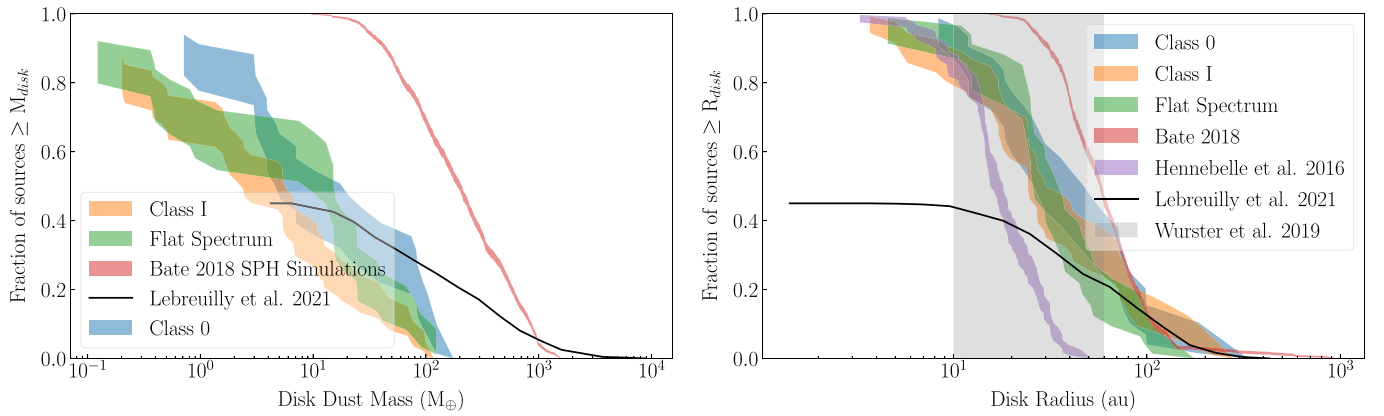
One potential resolution to this problem is that disk dust masses measured at 870  $\mu\text{m}$  or 1.3 mm are systematically being underestimated. Disk dust masses measured from longer-wavelength observations do indeed find higher disk masses for both embedded disks and protoplanetary disks (e.g., Tobin et al. 2020a; Tychoniec et al. 2020; Tazzari et al. 2021). Whether this discrepancy is the result of dust grain growth (e.g., Ricci et al. 2010; Tazzari et al. 2021), optical depth, or some combination of the two, as well as how that plays into the disk dust masses, however, likely requires detailed modeling to account for all of these effects simultaneously. Even with modeling, this picture may be complicated by the effects of dust grain scattering (e.g., Liu 2019; Zhu et al. 2019) across unseen substructures that can produce a low spectral index even in the face of small dust grains (e.g., Lin et al. 2020).

Overall, this work provides a new look at how disk properties evolve from protostellar to protoplanetary disks that contrasts results from earlier works. It is, however, likely that this picture will continue to evolve as improved models are confronted with better data.

#### 5.4. Constraints on Disk Formation Theory

When and how protostellar disks form during the star formation process has been a long-standing problem in our understanding of star and planet formation. For many years, it was unclear whether protostellar disks even could form during the early stages of cloud collapse and star formation, as ideal MHD simulations suggested that magnetic fields may prevent entirely the formation of disks (e.g., Allen et al. 2003; Galli et al. 2006; Hennebelle & Fromang 2008; Mellon & Li 2008; Li et al. 2014). In more recent years, the inclusion of nonideal MHD effects such as ambipolar diffusion, Hall effect, and ohmic resistivity (Dapp & Basu 2010; Li et al. 2011; Machida et al. 2011; Tsukamoto et al. 2015; Hennebelle et al. 2016; Masson et al. 2016; Tsukamoto et al. 2017b; Wurster &





**Figure 12.** Cumulative distributions of disk dust mass (left) and radius (right) of disks in our sample, as compared with the masses and radii of disks formed in simulations from Bate (2018), Hennebelle et al. (2016), Wurster et al. (2019), and Lebreuilly et al. (2021). The masses and radii from Bate (2018) come from SPH simulations of the global collapse of a molecular cloud to form a population of stars and disks, whereas the distribution of radii from Hennebelle et al. (2016) is the result of applying Equation (15) to the disks in our sample. Wurster et al. (2019) and Lebreuilly et al. (2021) both run simulations similar to Bate (2018) but include the effects of nonideal MHD in their models. We note that our observations probe dust masses and radii, while these simulations probe gas disk properties. To compare masses directly, we divide the simulation gas masses by a gas-to-dust ratio of 100, while we assume that gas and dust disks are of comparable size.

Bate 2019), as well as misalignments between magnetic field and angular momentum axis of the cloud (Hennebelle & Ciardi 2009; Joos et al. 2012; Seifried et al. 2012) or turbulence (e.g., Santos-Lima et al. 2012), can lead to the formation of disks early in the star formation process, and the question has now switched to the importance of these effects and the characteristics of the disks that are formed (Machida et al. 2011; Masson et al. 2016; Wurster & Lewis 2020a, 2020b; Hennebelle et al. 2020; Xu & Kunz 2021). For example, if the Hall effect is dominant, it may lead to a bimodal distribution in disk radii as disk formation is enhanced for configurations where angular momentum is anti-aligned with the magnetic field and suppressed when they are aligned (e.g., Tsukamoto et al. 2015, 2017a), though whether such bimodality will persist through subsequent evolution is unclear (e.g., Zhao et al. 2020).

In recent years, computational capacity has recently reached the point where global simulations of cloud collapse that follow gas from the molecular cloud scale down to the size scales of disks as they form are now becoming feasible (e.g., Kuffmeier et al. 2017; Bate 2018; Wurster et al. 2019), enabling the demographics of entire populations of protostellar disks to be simulated. Comparisons between the demographics derived from such simulations that include different suites of input physics and the observed demographics then provide a powerful constraint on the physics of disk formation.

In Figure 12, we show comparisons of the cumulative distributions of protostellar disk dust masses and radii with the cumulative distribution of protostellar disk properties from simulations by Bate (2018). Bate (2018) used smoothed particle hydrodynamics to consider pure hydrodynamics during the collapse of a  $500 M_{\odot}$  cloud to form a population of over 100 protostars and disks. Though both single stars and multiple systems are formed in their simulations, in Figure 12 we consider only the single stars to best match with our own sample. We also note that the disk radii in Bate (2018) are calculated using the radius inside of which 63.2% of the total disk mass is enclosed based on their observation that for the density profile in Equation (1), for  $\gamma < 2$ ,  $R_c$  always encloses a fraction  $(1 - 1/e)$  of the total disk mass, or 63.2%. As their criterion is based on the surface density model that we use in our analysis, our radii should be directly comparable. Finally,

as Bate (2018) considered gas simulations while we have measured dust properties, we note that in comparing disk masses we have divided the simulated disk masses by a gas-to-dust ratio of 100.

As can be seen in Figure 12, the disks produced in the Bate (2018) simulations tend to predict masses and sizes that are too large when compared with the disk dust masses and radii in our sample. This discrepancy may reflect the lack of magnetic fields in their simulations, which are known to remove angular momentum from the cloud, which should in turn lead to smaller disks. Indeed, using simple relations related to the timescales of magnetized collapse, Hennebelle et al. (2016) derived that young disks should have sizes of

$$R \approx 18 \text{ au} \left( \frac{\eta_{\text{AD}}}{0.1 \text{ s}} \right)^{\frac{2}{9}} \left( \frac{B_z}{0.1 \text{ G}} \right)^{-\frac{4}{9}} \left( \frac{M_* + M_{\text{disk}}}{0.1 M_{\odot}} \right)^{\frac{1}{3}}, \quad (15)$$

where  $\eta_{\text{AD}}$  is the ambipolar diffusion coefficient and  $B_z$  is the magnetic field in the inner region of the core. Hennebelle et al. (2020) later showed that this relation does a reasonably good job of predicting the disk sizes from a suite of MHD simulations within a factor of a few. Though exact values of  $\eta_{\text{AD}}$ ,  $B_z$ , and  $M_*$  are unknown, we can estimate the range of radii we would expect assuming their standard values for  $\eta_{\text{AD}}$  and  $B_z$ ,  $M_*$  sampled randomly from a Chabrier IMF, and using the disk mass calculated for each disk individually via our modeling. We show the distribution of disk radii calculated in Figure 12.

Interestingly, the disk sizes predicted by Hennebelle et al. (2016) using fiducial values for  $B_z$  and  $\eta_{\text{AD}}$  appear to be too *small* compared to the disks in our sample, possibly suggesting that magnetic braking might be too efficient at sapping angular momentum from the cloud. We note that there may be correlations between  $M_*$ ,  $M_{\text{disk}}$ , and  $R_{\text{disk}}$  that we are not accounting for by randomly assigning stellar masses and that might impact the exact distribution. Alternatively, magnetic fields were systematically weaker in our real disks than their fiducial values; this could also help to bring the two into agreement. For example, a reduction in the typical field strength by a factor of 3–4 would bring the distributions into better agreement. An increase in the strength of ambipolar

diffusion by a factor of 10–20 could also help remedy this disagreement. This might be achieved by altering the cosmic-ray ionization rate, to which ambipolar diffusion is sensitive (e.g., Marchand et al. 2016), which has numerically been shown to regulate the sizes of protostellar disks that are formed (e.g., Kuffmeier et al. 2020b). In both cases, the change from the fiducial values of  $B_z$  or  $\eta_{\text{AD}}$  in concert with assuming a lognormal distribution of their values with a standard deviation of  $\sim 1$  dex also helps to improve the agreement with the spread of observed disk radii.

To further examine the role of magnetic fields on disk formation in a global environment, Wurster et al. (2019) followed on the work of Bate (2018) by simulating the global collapse of a cloud to form stars and disks including MHD, while varying initial conditions such as the strength of the magnetic field and whether nonideal MHD effects were included in the calculations. For expediency, the simulations followed a cloud of only  $50 M_\odot$ , and so fewer disks were formed overall, making the comparison of demographics more challenging. The disks that did form around single protostars, though, had typical sizes of 10–60 au, in reasonable agreement with the typical sizes of disks in our sample.

Perhaps most realistically, Lebreuilly et al. (2021) recently ran simulations of the collapse of a  $1000 M_\odot$  cloud to form a population of 191 protostars and 42 disks when including nonideal MHD effects along with radiative transfer and stellar feedback. We show the cumulative distributions of the properties of those disks in Figure 12. We note that we follow Lebreuilly et al. (2021) in scaling the distribution according to the fraction of stars with disks (particularly their Figure 4, bottom middle panel), as they note that the smaller disks in their simulation were likely disrupted by numerical effects. We find that at the large radius end of the distribution their simulated disks are in good agreement with the distribution of disk dust radii that we measure. At smaller radii our distributions do diverge, but this may again be due to numerical effects in the simulations. The disk masses from Lebreuilly et al. (2021), on the other hand, tend to predict that the most massive disks are significantly more massive than what we find in our sample, but come to be in better agreement at the lower end of the mass distribution.

It is important to note, however, that while these comparisons may provide insights into disk formation physics, there are a number of limitations that must be kept in mind. Most notably, these figures compare dust disk radii derived from observations with gas disk radii as measured by simulations. As dust particles tend to decouple from the gas and experience aerodynamic drag that leads to radial drift of the dust (Weidenschilling 1977; Birnstiel et al. 2010), it is not necessarily the case that the sizes being compared should be in agreement (Lebreuilly et al. 2020). Observations of Class II protoplanetary disks have indeed demonstrated that gas disk radii are, on average,  $2.5\times$  larger when measured with gas observations than when measured with dust observations (e.g., Ansdell et al. 2018; Boyden & Eisner 2020; Sanchis et al. 2021). There is some question as to whether this is a true difference in size, as opposed to an optical depth effect (e.g., Trapman et al. 2019), and Class II disks are significantly older than the disks considered here, but nonetheless the difference between gas and dust radii may limit the utility of such comparisons. Moreover, while differing levels of physics are included in the simulations we consider, none include the

effects of stellar feedback, which has been shown to have a significant impact on cloud collapse and presumably therefore disk formation and evolution (e.g., Guszejnov et al. 2021).

Regardless of these caveats, such simulations provide a starting point for placing constraints on the physics that is important in setting the size scales of protostellar disks that will be improved on as both simulations and our observational data improve.

### 5.5. Gravitational Stability of Protostellar Disks

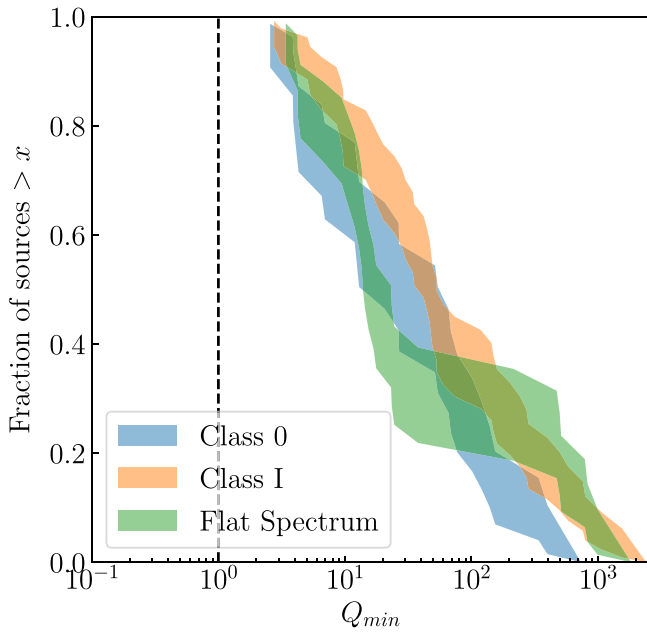
Self-gravity has long been thought to be an important driver of star and planet formation and the evolution of protoplanetary disks. In young, massive disks it has been suggested that gravitational instabilities might be a source of angular momentum transport (Vorobyov & Basu 2007, 2008; Machida et al. 2010; Zhu et al. 2010). Instabilities in disks might also be a source of companion stars in multiple systems (Adams et al. 1989; Kratter & Matzner 2006; Stamatellos & Whitworth 2009). Moreover, it has been suggested that gravitational instabilities might form planets in the outer regions of disks, where core accretion is less effective (e.g., Boss 1997; Durisen et al. 2007; Boss 2011), though the general consensus seems to be that planetary-mass companions formed via gravitational instabilities are difficult to keep from growing into brown dwarfs or stars (Kratter et al. 2010; Zhu et al. 2012; Forgan & Rice 2013). Nonetheless, gravitational instabilities may drive spiral arms that could be conducive to dust trapping that would otherwise enhance planet formation (e.g., Dipierro et al. 2015).

To test the gravitational stability of the disks in our sample and determine whether self-gravity is important, we use our radiative transfer modeling framework to calculate the Toomre  $Q$  profile for each disk in our sample. We do this via the equation

$$Q(R) = \frac{c_s(R) \Omega(R)}{\pi G \Sigma(R)}, \quad (16)$$

where  $c_s$  is the sound speed,  $\Omega$  is the rotational frequency, and  $\Sigma$  is the surface density, all of which vary as a function of radius  $R$  in the disk. Rather than calculate the Toomre  $Q$  profile for each sample of parameters in our posterior, which would be quite computationally expensive ( $\sim 10,000$  radiative transfer models per source), we calculate the profile only for the best-fit model. We use midplane temperature, averaged vertically over the five cells closest to the midplane, to calculate the sound speed for a gas with a mean molecular mass of 2.37, and we assume a gas-to-dust ratio of 100. Rather than assume a mass for the central source, we randomly draw  $\sim 10,000$  masses from the Chabrier IMF and use those to estimate the range of  $Q$  values for each protostar in our sample.

In Figure 13, we show the minimum Toomre  $Q$  value found in each disk as cumulative distributions split by source classification, as well as compared with  $T_{\text{bol}}$  and  $M_{\text{env}}/M_{\text{tot}}$ . We note that for simplicity we calculate the cumulative distributions using only  $M_* = 0.1 M_\odot$ , the peak of the Chabrier IMF. Though accounting for the stellar masses may alter the shape of these distributions, without direct knowledge of the stellar masses it may be difficult to account for their effect in a self-consistent way. The uncertainties shown in the comparisons with  $T_{\text{bol}}$  and  $M_{\text{env}}/M_{\text{disk}}$  are derived from the range of values calculated by drawing stellar masses randomly from the IMF.



**Figure 13.** Comparison of the minimum value of Toomre’s  $Q$  for the disks in our sample as a function of protostellar classification.

We find that, collectively, protostellar disks do not seem to be broadly gravitationally unstable. About 20% of the disks in our sample have  $Q_{\min} < 10$ , with a handful approaching  $Q_{\min} \sim 1$  for low-mass protostars, but few seem to be definitively subject to self-gravity. Disks with more modestly low values, of  $Q \sim 2$ , may still be capable of forming spirals and transporting angular momentum (e.g., Tomida et al. 2017), but there still remain few disks that approach this level of instability.

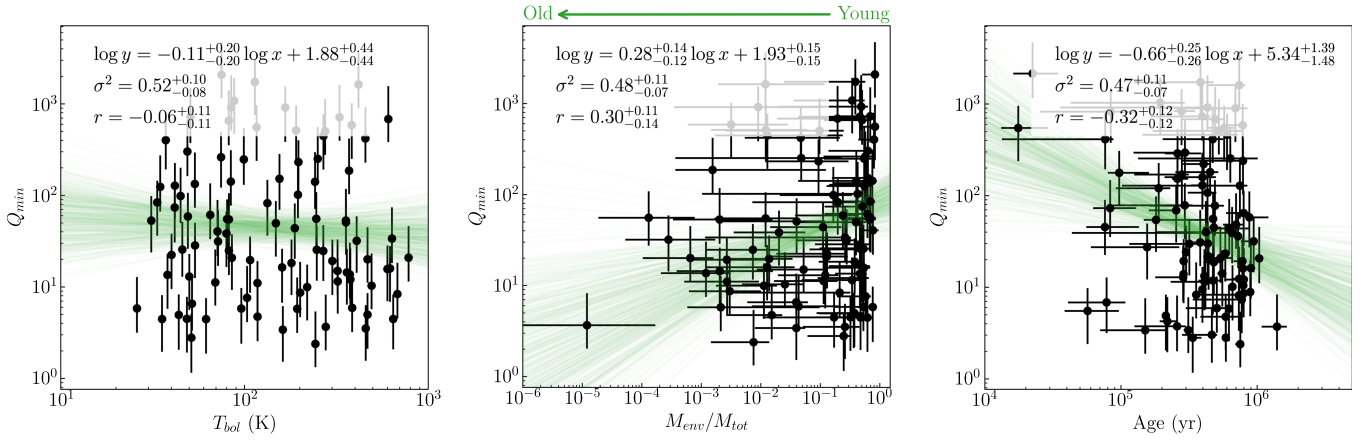
In Figure 14, we explore how  $Q_{\min}$  compares with the simple estimators of protostellar evolution discussed in Section 5.2. If we consider the Class 0/I/flat-spectrum identification as an indicator of evolution, or equivalently the bolometric temperature, we do not see much evidence that the gravitational stability of the disks changes with time. On the other hand, if we again consider the ratio of envelope mass to total mass,  $M_{\text{env}}/M_{\text{tot}}$ , or use the age inferred from simple models of protostellar evolution, we do find a trend of decreasing  $Q_{\min}$  for older systems at the  $\sim 2\sigma$  level. This decrease toward gravitationally unstable disks with time may reflect the weak trend of growth of the disk dust mass that was seen in Section 5.2.2.

Ultimately, though the lack of disks that appear to be gravitationally unstable, or perhaps even close, may provide commentary on the importance of gravitational instability in these disks, it is also possible that we do not see a substantial number of gravitationally unstable disks because the timescales for instabilities are short, and mass is drained quickly from the disk when this state is reached (e.g., Stamatellos & Whitworth 2009). We do find some disks that appear to have been caught in the act (e.g., Tobin et al. 2016a, 2018; Reynolds et al. 2021). So it may simply be the case that these disks remain in this state for short periods of time and are therefore difficult to directly observe, or have already formed multiple systems that would have been excluded from our modeling.

## 6. Conclusions

In this work we have fit 97 protostars from the VANDAM: Orion sample with two-dimensional, axisymmetric radiative transfer models. We fit the VANDAM: Orion ALMA 345 GHz visibilities along with the HOPS Survey SEDs simultaneously for each source using MCMC fitting to provide a comprehensive picture of disk and envelope structure. Our main results are as follows:

1. We find a median protostellar disk dust mass of  $5.8^{+4.6}_{-2.7} M_{\oplus}$  and a median disk dust radius of  $29.4^{+4.1}_{-3.2}$  au. Only  $11.3^{+4.6}_{-3.4}\%$  of disks have dust radii larger than 100 au. If we group our sample by observational class, we find median disk dust masses of  $7.1^{+14.3}_{-2.0} M_{\oplus}$ ,  $4.9^{+1.0}_{-2.7} M_{\oplus}$ , and  $14.0^{+1.3}_{-7.0} M_{\oplus}$  and median disk dust radii of  $35.6^{+17.1}_{-10.0}$  au,  $26.9^{+4.5}_{-3.3}$  au, and  $29.5^{+6.1}_{-4.4}$  au for Class 0, Class I, and flat-spectrum sources, respectively.
2. Disk dust masses measured from the frequently used simple flux-based measurement disagree with the disk dust masses measured using our models, and the difference is primarily due to the additional physics included in our radiative transfer modeling that impacts the temperature of the disk. A treatment of the temperature that accounts for disk radius does substantially improve the discrepancy, though there remains substantial scatter. We do note that *disk fluxes* can be accurately recovered with simple methods like two-dimensional Gaussian fitting.
3. We find evidence that Class I disk dust masses can be distinguished from Class 0 disk dust masses and that the envelope dust masses of each class are distinct, though only the distinction between flat spectrum and Class 0 is statistically significant. Otherwise, we find little evidence that most disk and envelope properties evolve with time, whether we consider the Class 0/I/flat-spectrum identification or the bolometric temperature as the tracer of evolution.
4. The distribution of flat-spectrum source inclinations is distinct from both Class 0 and Class I sources, and also that inclination is correlated with bolometric temperature, demonstrating that the bolometric temperature is contaminated by viewing angle and does not directly trace the evolutionary state of these young systems.
5. We use both  $M_{\text{env}}/M_{\text{tot}} = M_{\text{env}}/(M_{*} + M_{\text{disk}} + M_{\text{env}})$  and simple evolutionary tracks in the  $L_{\text{bol}} - M_{\text{env}}$  plane as alternate ways to trace protostellar evolution and find weak evidence that disk dust masses may actually *increase* with time.
6. A comparison of the disk dust masses derived from our radiative transfer modeling with dust masses derived from other, similar radiative transfer modeling studies of disks makes it less clear that Class 0/I/flat-spectrum disks are more massive than Class II disks, though we acknowledge that this difference may be due to environment or systematic differences in modeling details.
7. We compare the bulk properties of the disks in our sample with simulations of protostellar disk formation. We find that simulations that follow the collapse of a molecular cloud to form a population of protostars and



**Figure 14.** Comparison of the minimum value of Toomre's  $Q$  for the disks in our sample as a function of protostellar classification (left), the bolometric temperature (middle), and the ratio of envelope mass to total mass (right).

disks using pure hydrodynamics (Bate 2018) produce disks that are too large in both radius and mass. Simulations that include the effects of magnetic fields produce smaller disks, perhaps in better agreement with our results, though similar simulations of global cloud collapse are as of yet still much smaller in scale and number of protostars (and disks) formed (e.g., Wurster et al. 2019).

8. We use our radiative transfer models to construct the Toomre  $Q$  profile for all of the disks in our sample, and we find that very few disks are gravitationally unstable.

Though it may be expensive, these results demonstrate the importance of detailed radiative transfer modeling in interpreting surveys of protostellar and protoplanetary disks with ALMA. It is clear that further studies employing more uniform techniques across disks in a range of environments and at different ages will be critical for evaluating the evolution of disk properties and their effects on planet formation.

The authors would like to thank the anonymous referee for feedback that helped to focus and improve the manuscript. P.D.S. would also like to thank Leonardo Testi and the referee for pointing him to the protostellar evolutionary tracks used in Section 5.2.3, which added an interesting new piece of analysis to the work. P.D.S. is supported by a National Science Foundation Astronomy & Astrophysics Postdoctoral Fellowship under award No. 2001830. J.J.T. acknowledges support from NSF AST-1814762 and past support from the Homer L. Dodge Endowed Chair at the University of Oklahoma. L.W.L.

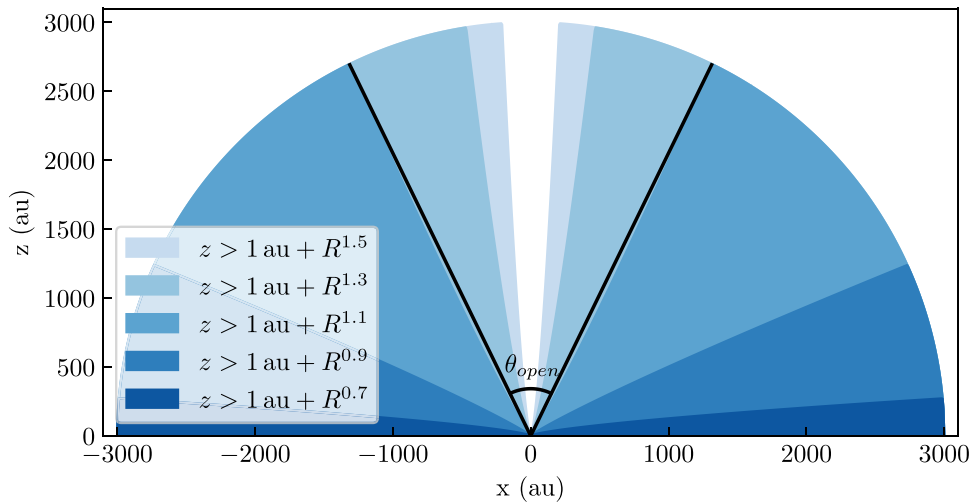
acknowledges support from NSF AST-1910364 and AST-2108794. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant No. ACI-1548562, for the majority of the computing done. Additional computing was performed at the OU Supercomputing Center for Education and Research (OSCE) at the University of Oklahoma (OU), when needed. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2015.1.00041.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), NSC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

*Software:* *pdspy* (Sheehan 2018), *CASA* (McMullin et al. 2007), *RADMC-3D* (Dullemond 2012), *emcee* (Foreman-Mackey et al. 2013), *matplotlib* (Hunter 2007), *corner* (Foreman-Mackey 2016), *GALARIO* (Tazzari et al. 2018), *lifelines* (Davidson-Pilon et al. 2019).

## Appendix Diagram of Envelope Cavity Shape

We show a diagram demonstrating how the  $\xi$  parameter affects the shape of the outflow cavity in our model in Figure 15.





**Figure 15.** Diagram showing the shape of the outflow cavity, as defined by  $z > 1\text{ au} + R^\xi$ , for a range of values of  $\xi$ . The black lines show the definition of  $\theta_{\text{open}}$  for  $\xi = 1.1$ .

### ORCID iDs

Patrick D. Sheehan <https://orcid.org/0000-0002-9209-8708>  
 John J. Tobin <https://orcid.org/0000-0002-6195-0152>  
 Leslie W. Looney <https://orcid.org/0000-0002-4540-6587>  
 S. Thomas Megeath <https://orcid.org/0000-0001-7629-3573>

### References

- Adams, F. C., Ruden, S. P., & Shu, F. H. 1989, *ApJ*, **347**, 959  
 Allen, A., Li, Z., & Shu, F. H. 2003, *ApJ*, **599**, 363  
 Andersen, B. C., Stephens, I. W., Dunham, M. M., et al. 2019, *ApJ*, **873**, 54  
 André, P., Minier, V., Gallais, P., et al. 2008, *A&A*, **490**, L27  
 André, P., & Montmerle, T. 1994, *ApJ*, **420**, 837  
 André, P., Montmerle, T., André, P., & Montmerle, T. 1994, *ApJ*, **420**, 837  
 André, P., Ward-Thompson, D., & Barsony, M. 1993, *ApJ*, **406**, 122  
 Andrews, S. M., Huang, J., Pérez, L. M., et al. 2018, *ApJL*, **869**, L41  
 Andrews, S. M., Rosenfeld, K. A., Kraus, A. L., & Wilner, D. J. 2013, *ApJ*, **771**, 129  
 Andrews, S. M., Wilner, D. J., Zhu, Z., et al. 2016, *ApJL*, **820**, L40  
 Ansdell, M., Haworth, T. J., Williams, J. P., et al. 2020, *AJ*, **160**, 248  
 Ansdell, M., Williams, J. P., Marel, N. v. d., et al. 2016, *ApJ*, **828**, 46  
 Ansdell, M., Williams, J. P., Trapman, L., et al. 2018, *ApJ*, **859**, 21  
 Ballering, N. P., & Eisner, J. A. 2019, *AJ*, **157**, 144  
 Barenfeld, S. A., Carpenter, J. M., Ricci, L., & Isella, A. 2016, *ApJ*, **827**, 142  
 Bate, M. R. 2018, *MNRAS*, **475**, 5618  
 Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Guesten, R. 1990, *AJ*, **99**, 924  
 Birnstiel, T., Dullemond, C. P., & Brauer, F. 2010, *A&A*, **513**, 79  
 Boss, A. P. 1997, *Sci*, **276**, 1836  
 Boss, A. P. 2011, *ApJ*, **731**, 74  
 Boyden, R. D., & Eisner, J. A. 2020, *ApJ*, **894**, 74  
 Brogan, C. L., Pérez, L. M., Hunter, T. R., et al. 2015, *ApJL*, **808**, L3  
 Calvet, N., Hartmann, L., Kenyon, S. J., & Whitney, B. A. 1994, *ApJ*, **434**, 330  
 Cazzoletti, P., Manara, C. F., Baobab Liu, H., et al. 2019, *A&A*, **626**, A11  
 Chabrier, G. 2003, *PASP*, **115**, 763  
 Chen, H., Myers, P. C., Ladd, E. F., & Wood, D. O. S. 1995, *ApJ*, **445**, 377  
 Chiang, E. I., & Goldreich, P. 1999, *ApJ*, **519**, 279  
 Cieza, L. A., Ruiz-Rodríguez, D., Hales, A., et al. 2019, *MNRAS*, **482**, 698  
 Codella, C., Cabrit, S., Gueth, F., et al. 2014, *A&A*, **568**, L5  
 Crapsi, A., Van Dishoeck, E. F., Hogerheijde, M. R., Pontoppidan, K. M., & Dullemond, C. P. 2008, *A&A*, **486**, 245  
 Dapp, W. B., & Basu, S. 2010, *A&A*, **521**, L56  
 Dapp, W. B., Basu, S., & Kunz, M. W. 2012, *A&A*, **541**, A35  
 Davidson-Pilon, C., Kalderstam, J., Zivich, P., et al. 2019, *CamDavidsonPilon/lifelines: v0.21.1*, Zenodo, doi:[10.5281/zenodo.2652543](https://doi.org/10.5281/zenodo.2652543)  
 Dipierro, G., Pinilla, P., Lodato, G., & Testi, L. 2015, *MNRAS*, **451**, 974  
 Draine, B. T. 2003, *ApJ*, **598**, 1026  
 Dullemond, C. P. 2012, RADMC-3D: A multi-purpose radiative transfer tool, Astrophysics Source Code Library, ascl:[1202.015](https://ui.adsabs.org/abs/2012ASCl..1202D..015D)  
 Dunham, M. M., Allen, L. E., Evans, N. J., II, et al. 2015, *ApJS*, **220**, 11  
 Dunham, M. M., Evans, N. J., Terebey, S., Dullemond, C. P., & Young, C. H. 2010, *ApJ*, **710**, 470  
 Dunham, M. M., Vorobyov, E. I., & Arce, H. G. 2014, *MNRAS*, **444**, 887  
 Durisen, R., Boss, A., Mayer, L., et al. 2007, in *Protostars and Planets V*, ed. B. Reipurth et al. (Tucson, AZ: Univ. Arizona Press), 607  
 Eisner, J. A., Arce, H. G., Ballering, N. P., et al. 2018, *ApJ*, **860**, 77  
 Enoch, M. L., Evans, N. J., II, Sargent, A. I., et al. 2008, *ApJ*, **684**, 1240  
 Evans, N. J., Dunham, M. M., Jørgensen, J. K., et al. 2009, *ApJS*, **181**, 321  
 Feiden, G. A. 2016, *A&A*, **593**, A99  
 Fischer, W. J., Megeath, S. T., Furlan, E., et al. 2017, *ApJ*, **840**, 69  
 Fischer, W. J., Megeath, S. T., Furlan, E., et al. 2020, *ApJ*, **905**, 119  
 Foreman-Mackey, D. 2016, *JOSS*, **1**, 24  
 Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306  
 Forgan, D., & Rice, K. 2013, *MNRAS*, **432**, 3168  
 Furlan, E., Fischer, W. J., Ali, B., et al. 2016, *ApJS*, **224**, 5  
 Galli, D., Lizano, S., Shu, F. H., & Allen, A. 2006, *ApJ*, **647**, 374  
 Grant, S. L., Espaillat, C. C., Wendeborn, J., et al. 2021, *ApJ*, **913**, 123  
 Greene, T. P., Gully-Santiago, M. A., & Barsony, M. 2018, *ApJ*, **862**, 85  
 Guszejnov, D., Grudić, M. Y., Hopkins, P. F., Offner, S. S. R., & Faucher-Giguère, C.-A. 2021, *MNRAS*, **502**, 3646  
 Harsono, D., Jørgensen, J. K., Van Dishoeck, E. F., et al. 2014, *A&A*, **562**, A77  
 Hennebelle, P., & Ciardi, A. 2009, *A&A*, **506**, 29  
 Hennebelle, P., Commerçon, B., Chabrier, G., & Marchand, P. 2016, *ApJL*, **830**, L8  
 Hennebelle, P., Commerçon, B., Lee, Y. N., & Charnoz, S. 2020, *A&A*, **635**, A67  
 Hennebelle, P., & Fromang, S. 2008, *A&A*, **477**, 9  
 Hildebrand, R. 1983, *QJRAS*, **24**, 267  
 Hillenbrand, L. A., & Carpenter, J. M. 2000, *ApJ*, **540**, 236  
 Huang, J., Andrews, S. M., Dullemond, C. P., et al. 2018, *ApJL*, **869**, L42  
 Hudgins, D. M., Sandford, S. A., Allamandola, L. J., & Tielens, A. G. G. M. 1993, *ApJS*, **86**, 713  
 Hunter, J. D. 2007, *CSE*, **9**, 90  
 Isella, A., Benisty, M., Teague, R., et al. 2019, *ApJL*, **879**, L25  
 Isella, A., Guidi, G., Testi, L., et al. 2016, *PhRvL*, **117**, 251101  
 Joos, M., Hennebelle, P., & Ciardi, A. 2012, *A&A*, **543**, A128  
 Jørgensen, J. K., Van Dishoeck, E. F., Visser, R., et al. 2009, *A&A*, **507**, 861  
 Kelly, B. C. 2007, *ApJ*, **665**, 1489  
 Kenyon, S. J., Hartmann, L. W., Strom, K. M., & Strom, S. E. 1990, *AJ*, **99**, 869  
 Keppler, M., Benisty, M., Müller, A., et al. 2018, *A&A*, **617**, A44  
 Kratter, K. M., & Matzner, C. D. 2006, *MNRAS*, **373**, 1563  
 Kratter, K. M., Murray-Clay, R. A., & Youdin, A. N. 2010, *ApJ*, **710**, 1375  
 Kristensen, L. E., & Dunham, M. M. 2018, *A&A*, **618**, 158  
 Kuffmeier, M., Frimann, S., Jensen, S. S., & Haugbølle, T. 2018, *MNRAS*, **475**, 2642

- Kuffmeier, M., Goicovic, F. G., & Dullemond, C. P. 2020a, *A&A*, **633**, A3
- Kuffmeier, M., Haugbølle, T., & Nordlund, Å. 2017, *ApJ*, **846**, 7
- Kuffmeier, M., Zhao, B., & Caselli, P. 2020b, *A&A*, **639**, A86
- Lebreuilly, U., Commerçon, B., & Laibe, G. 2020, *A&A*, **641**, A112
- Lebreuilly, U., Hennebelle, P., Colman, T., et al. 2021, *ApJL*, **917**, L10
- Lee, C. F., Li, Z. Y., & Turner, N. J. 2020, *NatAs*, **4**, 142
- Li, Z.-Y., Banerjee, R., Pudritz, R. E., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 173
- Li, Z. Y., Krasnopolsky, R., & Shang, H. 2011, *ApJ*, **738**, 180
- Li, Z. Y., Krasnopolsky, R., & Shang, H. 2013, *ApJ*, **774**, 82
- Lin, Z. Y. D., Li, Z. Y., Yang, H., et al. 2020, *MNRAS*, **496**, 169
- Liu, H. B. 2019, *ApJL*, **877**, L22
- Long, F., Pinilla, P., Herczeg, G. J., et al. 2018, *ApJ*, **869**, 17
- Lynden-Bell, D., & Pringle, J. E. 1974, *MNRAS*, **168**, 603
- Machida, M. N., Inutsuka, S. I., & Matsumoto, T. 2010, *ApJ*, **724**, 1006
- Machida, M. N., Inutsuka, S. I., & Matsumoto, T. 2011, *PASJ*, **63**, 555
- Manara, C. F., Morbidelli, A., & Guillot, T. 2018, *A&A*, **618**, L3
- Manoj, P., Watson, D. M., Neufeld, D. A., et al. 2013, *ApJ*, **763**, 83
- Marchand, P., Masson, J., Chabrier, G., et al. 2016, *A&A*, **592**, A18
- Masson, J., Chabrier, G., Hennebelle, P., Vaytet, N., & Commerçon, B. 2016, *A&A*, **587**, A32
- Matzner, C. D., & McKee, C. F. 2000, *ApJ*, **545**, 364
- Maury, A. J., André, P., Testi, L., et al. 2019, *A&A*, **621**, A76
- Maury, A. J., Belloche, A., André, P., et al. 2014, *A&A*, **563**, L2
- Maury, A. J., Girart, J. M., Zhang, Q., et al. 2018, *MNRAS*, **477**, 2760
- McKee, C. F., & Offner, S. S. 2010, *ApJ*, **716**, 167
- McKee, C. F., & Tan, J. C. 2002, *ApJ*, **585**, 850
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in *ASP Conf. Ser. 376, Astronomical Data Analysis Software and Systems XVI*, ed. R. A. Shaw et al. (San Francisco, CA: ASP), 127
- Mellon, R. R., & Li, Z. 2008, *ApJ*, **681**, 1356
- Min, M., Hovenier, J. W., & de Koter, A. 2005, *A&A*, **432**, 909
- Molinari, S., Pezzuto, S., Cesaroni, R., et al. 2008, *A&A*, **481**, 345
- Murillo, N. M., Lai, S. P., Bruderer, S., Harsono, D., & Van Dishoeck, E. F. 2013, *A&A*, **560**, A103
- Myers, P. C., Fuller, G. A., Mathieu, R. D., et al. 1987, *ApJ*, **319**, 340
- Najita, J. R., & Kenyon, S. J. 2014, *MNRAS*, **445**, 3315
- Offner, S. S., & Arce, H. G. 2014, *ApJ*, **784**, 61
- Pascucci, I., Testi, L., Herczeg, G. J., et al. 2016, *ApJ*, **831**, 125
- Pezzuto, S., Benedettini, M., Di Francesco, J., et al. 2021, *A&A*, **645**, A55
- Pinte, C., Price, D. J., Ménard, F., et al. 2018, *ApJL*, **860**, L13
- Pinte, C., Price, D. J., Ménard, F., et al. 2020, *ApJL*, **890**, L9
- Reynolds, N. K., Tobin, J. J., Sheehan, P., et al. 2021, *ApJL*, **907**, L10
- Ribas, A., Espaillat, C. C., Macías, E., & Sarro, L. M. 2020, *A&A*, **642**, A171
- Ricci, L., Testi, L., Natta, A., et al. 2010, *A&A*, **512**, A15
- Robitaille, T. P., Whitney, B. A., Indebetouw, R., Wood, K., & Denzmore, P. 2006, *ApJS*, **167**, 256
- Sadavoy, S. I., Di Francesco, J., Andre, P., et al. 2014, *ApJL*, **787**, L18
- Safron, E. J., Fischer, W. J., Megeath, S. T., et al. 2015, *ApJL*, **800**, L5
- Sallum, S., Follette, K. B., Eisner, J. A., et al. 2015, *Natur*, **527**, 342
- Sanchis, E., Testi, L., Natta, A., et al. 2021, *A&A*, **649**, A19
- Santos-Lima, R., De Gouveia Dal Pino, E. M., & Lazarian, A. 2012, *ApJ*, **747**, 21
- Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M., & Molinari, S. 1996, *A&A*, **309**, 827
- Segura-Cox, D. M., Harris, R. J., Tobin, J. J., et al. 2016, *ApJL*, **817**, L14
- Segura-Cox, D. M., Looney, L. W., Tobin, J. J., et al. 2018, *ApJ*, **866**, 161
- Seifried, D., Banerjee, R., Pudritz, R. E., & Klessen, R. S. 2012, *MNRAS*, **423**, L40
- Sheehan, P. 2018, psheehan/pdspy: pdspy: A MCMC Tool for Continuum and Spectral Line Radiative Transfer Modeling, Zenodo, doi:10.5281/zenodo.2455079
- Sheehan, P. D., & Eisner, J. A. 2014, *ApJ*, **791**, 19
- Sheehan, P. D., & Eisner, J. A. 2017, *ApJ*, **851**, 45
- Stamatellos, D., & Whitworth, A. P. 2009, *MNRAS*, **392**, 413
- Stephens, I. W., Dunham, M. M., Myers, P. C., et al. 2018, *ApJS*, **237**, 22
- Stutz, A. M., Tobin, J. J., Stanke, T., et al. 2013, *ApJ*, **767**, 36
- Tazzari, M., Beaujean, F., & Testi, L. 2018, *MNRAS*, **476**, 4527
- Tazzari, M., Clarke, C. J., Testi, L., et al. 2021, *MNRAS*, **506**, 5117
- Teague, R., Bae, J., Bergin, E. A., Birnstiel, T., & Foreman-Mackey, D. 2018, *ApJL*, **860**, L12
- Tobin, J. J., Hartmann, L., Calvet, N., et al. 2012, *Natur*, **492**, 83
- Tobin, J. J., Kratter, K. M., Persson, M. V., et al. 2016a, *Natur*, **538**, 483
- Tobin, J. J., Looney, L. W., Li, Z.-Y., et al. 2016b, *ApJ*, **818**, 73
- Tobin, J. J., Looney, L. W., Li, Z.-Y., et al. 2018, *ApJ*, **867**, 43
- Tobin, J. J., Sheehan, P. D., Megeath, S. T., et al. 2020a, *ApJ*, **890**, 130
- Tobin, J. J., Sheehan, P. D., Reynolds, N., et al. 2020b, *ApJ*, **905**, 162
- Tomida, K., Machida, M. N., Hosokawa, T., Sakurai, Y., & Lin, C. H. 2017, *ApJL*, **835**, L11
- Trapman, L., Facchini, S., Hogerheijde, M. R., van Dishoeck, E. F., & Bruderer, S. 2019, *A&A*, **629**, A79
- Tripathi, A., Andrews, S. M., Birnstiel, T., & Wilner, D. J. 2017, *ApJ*, **845**, 44
- Tsukamoto, Y., Iwasaki, K., Okuzumi, S., Machida, M. N., & Inutsuka, S. 2015, *ApJL*, **810**, L26
- Tsukamoto, Y., Okuzumi, S., Iwasaki, K., Machida, M. N., & Inutsuka, S. I. 2017a, *PASJ*, **69**, 95
- Tsukamoto, Y., Okuzumi, S., & Kataoka, A. 2017b, *ApJ*, **838**, 151
- Tychoniec, L., Manara, C. F., Rosotti, G. P., et al. 2020, *A&A*, **640**, A19
- Tychoniec, L., Tobin, J. J., Karska, A., et al. 2018, *ApJS*, **238**, 19
- Ulrich, R. K. 1976, *ApJ*, **210**, 377
- van der Marel, N., Dong, R., di Francesco, J., Williams, J. P., & Tobin, J. 2019, *ApJ*, **872**, 112
- Van Terwisga, S. E., Van Dishoeck, E. F., Mann, R. K., et al. 2020, *A&A*, **640**, A27
- Vorobyov, E. I., & Basu, S. 2007, *MNRAS*, **381**, 1009
- Vorobyov, E. I., & Basu, S. 2008, *MNRAS*, **393**, 822
- Weidenschilling, S. J. 1977, *MNRAS*, **180**, 57
- Whitney, B. A., Wood, K., Bjorkman, J. E., & Cohen, M. 2003, *ApJ*, **598**, 1079
- Williams, J. P., Cieza, L., Hales, A., et al. 2019, *ApJL*, **875**, L9
- Woitke, P., Min, M., Pinte, C., et al. 2016, *A&A*, **586**, A103
- Wurster, J., & Bate, M. R. 2019, *MNRAS*, **486**, 2587
- Wurster, J., Bate, M. R., & Price, D. J. 2019, *MNRAS*, **489**, 1719
- Wurster, J., & Lewis, B. T. 2020a, *MNRAS*, **495**, 3807
- Wurster, J., & Lewis, B. T. 2020b, *MNRAS*, **495**, 3795
- Xu, W., & Kunz, M. W. 2021, *MNRAS*, **502**, 4911
- Yen, H.-W., Koch, P. M., Takakuwa, S., et al. 2017, *ApJ*, **834**, 178
- Yoo, H., Lee, J.-E., Mairs, S., et al. 2017, *ApJ*, **849**, 69
- Zhao, B., Caselli, P., Li, Z. Y., et al. 2020, *MNRAS*, **492**, 3375
- Zhu, Z., Hartmann, L., & Gammie, C. 2010, *ApJ*, **713**, 1143
- Zhu, Z., Hartmann, L., Nelson, R. P., & Gammie, C. F. 2012, *ApJ*, **746**, 110
- Zhu, Z., Zhang, S., Jiang, Y.-F., et al. 2019, *ApJL*, **877**, L18
- Zubko, V. G., Mennella, V., Colangeli, L., & Bussoletti, E. 1996, *MNRAS*, **282**, 1321