The Astrophysical Journal, 991:176 (14pp), 2025 October 1 https://doi.org/10.3847/1538-4357/adfa15

Radiative Transfer Modeling of a Shadowed Protoplanetary Disk Assisted by a Neural Network

Jonathan P. Williams[1], Myriam Benisty[2], Christian Ginski[3], Giuseppe Lodato[4], and Maria Vincent[5]
[1] Institute for Astronomy, University of Hawai'i at Mānoa, 2680 Woodlawn Drive, Honolulu, HI 96822, USA; jw@hawaii.edu
[2] Max-Planck Institute for Astronomy (MPIA), Königstuhl 17, 69117 Heidelberg, Germany
[3] School of Natural Sciences, Center for Astronomy, University of Galway, Galway, H91 CF50, Ireland
[4] Dipartimento di Fisica, Universitá degli Studi di Milano, Via Celoria 16, I-20133 Milano, Italy
[5] Institute for Astronomy, University of Hawai'i, 640 N. Aohoku Pl., Hilo, HI 96720, USA

Received 2025 June 10; revised 2025 August 4; accepted 2025 August 7; published 2025 September 26

Abstract

We present observations and detailed modeling of a protoplanetary disk around the T Tauri star V1098 Sco. Millimeter-wavelength data from the Atacama Large Millimeter/submillimeter Array (ALMA) show a ring of large dust grains with a central cavity that is filled with molecular gas. Near-infrared data from the Very Large Telescope (VLT) detect the scattered starlight from the disk surface and reveal a large shadow that extends over its entire southern half. We model the ALMA continuum and line data to determine the outer-disk geometry and the central stellar mass. Using radiative transfer models, we demonstrate that a misaligned inner disk, tilted in both inclination and position angle with respect to the outer disk, can reproduce the salient scattered-light features seen with VLT. Applying an image threshold algorithm to compare disk morphologies and training a neural network on a set of high-signal-to-noise-ratio models, we forward model the data and determine the inner-disk geometry. We find that the rotation axes of the inner and outer disks are misaligned by 38° and constrain the mass and location of a perturbing planetary or substellar companion. The technique of simulation-based inference that is illustrated here is broadly applicable for radiative transfer modeling of other objects.

Unified Astronomy Thesaurus concepts: Protoplanetary disks (1300); Radiative transfer (1335); Neural networks (1933)

1. Introduction

The varied orbital architectures of the multitude of planetary systems discovered to date are due to a combination of their formation properties and subsequent interactions with the surrounding disk, together with longer-term dynamical effects with the host star and planetary siblings. Disk properties determine the locations of the initial seeds, which then symbiotically evolve through gravitational torques that drive gas accretion, gap clearing, and planetary migration.

Planetary embryos form in the dense, flat disk midplane, but fully formed planets can, in some cases, orbit at high inclinations relative to the stellar rotation axis (A. H. M. J. Triaud 2018). This may be caused after the disk has been dispersed by planet–planet scattering (E. B. Ford & F. A. Rasio 2008) or interactions with a companion star via the Kozai mechanism (Y. Wu & N. Murray 2003). However, orbital obliquity might also be imprinted during formation (M. R. Bate et al. 2010), as some protoplanetary disks (defined as disks around optically visible pre-main-sequence stars) also show indications of the early evolution of orbital inclination through the asymmetric shadows cast by a misaligned inner disk on scattered light from the outer disk (S. Marino et al. 2015), from direct imaging of edge-on disks (M. Villenave et al. 2024), and as dips in the starlight from dust along the line of sight toward systems that appear more face-on (M. Ansdell et al. 2020).

For over a decade now, we have had the tools to image protoplanetary disk structure at the scales of planet–disk interactions. The Atacama Large Millimeter/submillimeter Array (ALMA) and the SPHERE instrument on the Very Large Telescope (VLT) have been at the forefront of such work (J. Bae et al. 2023; M. Benisty et al. 2023). Both resolve disk structures at scales of a few astronomical units in nearby disks and complement each other, as the long wavelengths of ALMA reveal the cold molecular gas and relatively large (approximately millimeter-sized) dust grains in the disk midplane, whereas SPHERE detects the near-infrared starlight scattered by much smaller (approximately micron-sized) grains in the upper disk atmosphere (S. M. Andrews 2020).

In this paper, we present ALMA and SPHERE images of the protoplanetary disk around the T Tauri star V1098 Sco (2MASS J16140792-1938292). V1098 Sco is of spectral type K1 and lies at a distance of 160 pc (Gaia Collaboration 2020) in the Upper Sco region of the Sco–Cen OB association, where the median age is estimated to be 5 ± 2 Myr (M. J. Pecaut & E. E. Mamajek 2016). It was first recognized as a T Tauri star by J. Gregorio-Hetem et al. (1992), who carried out a spectroscopic survey of color-selected IRAS sources that showed it to have weak and variable Hα, indicative of low accretion, and moderate lithium, indicative of youth. More recently, it was classified as a quasiperiodic dipper from the analysis of K2 light curves by A. M. Cody & L. A. Hillenbrand (2018). However, this otherwise unremarkable star has a quite striking appearance in scattered light, as one half of the disk appears to be missing, due to a large shadow, and millimeter wavelength imaging reveals a large central cavity.

We present the observations in Section 2, fit the ALMA data in Section 3, and model the SPHERE scattered-light image in Section 4. We set up a large grid of radiative transfer models and employ a novel method, using a neural network to interpolate model images for parameter values between grid points, which allows us to efficiently fit the data using a

Markov Chain Monte Carlo (MCMC) approach. We find that the outer-disk shadow is caused by a large misaligned inner disk, and we discuss mechanisms for this particular geometry in Section 5. We then summarize our results in Section 6.

2. Observations

2.1. SPHERE Near-infrared Observations

V 1098 Sco was observed as part of the ESO Disk Evolution Study Through Imaging of Nearby Young Stars (DESTINYS) large program (C. Ginski et al. 2020; Project ID: 1104.C-0415 (D)) on 2022 July 30, using the IRDIS subsystem of SPHERE. The observations were obtained in dual polarization imagings mode (R. G. van Holstein et al. 2020) with the broadband H filter, with an effective wavelength of 1.6 μm. The central star was obscured by a coronagraphic mask with an effective inner working angle (defined as a 50% drop in throughput) of 92.5 mas. The weather conditions during the observations were excellent, with optical seeing between $0''.4$ and $0''.8$ and a coherence time of the atmosphere between 5 ms and 11 ms. We recorded a total of 26 full polarimetric cycles, each comprised of four images with different half-wave plate orientations, as described in C. Ginski et al. (2020). Each image had an individual exposure time of 32 s, giving us a total integration time of 55.5 minutes. In addition to the science data, we recorded dedicated flux, center, and sky calibration frames at the beginning and end of the science sequence.

The data were reduced using the public IRDIS Data reduction for Accurate Polarimetry (IRDAP) pipeline.[6] To briefly summarize, IRDAP first performed general image-processing tasks, such as sky subtraction, flat-fielding, and bad-pixel correction, before each frame was centered using the stellar coordinates measured in the center calibration frames. Using center calibration observations at the beginning and end of the science sequence, we find that the determined stellar center position deviates by less than 0.05 pixels between these two, indicating the excellent stability of the stellar position behind the coronagraphic mask. After the initial data processing, IRDAP then performed polarization differential imaging, which cancels the unpolarized stellar light and retains the partially linearly polarized scattered light from the circumstellar disk. IRDAP accounts for instrumental polarization by means of a detailed instrument model, described in R. G. van Holstein et al. (2020).

2.2. ALMA Millimeter Observations

We downloaded ALMA observations of V1098 Sco from the archive (Project 2018.1.00564.S, recently published in J. M. Carpenter et al. 2025) and reduced it using the same procedures[7] as for the eDisk project (N. Ohashi et al. 2023). The final data products are a Band 7 (340 GHz) continuum and a CO 3–2 spectral line. The beam and rms noise levels are $0''.41 \times 0''.32$ at a position angle of $-76°$ and 0.13 mJy beam^{-1} for the continuum and $0''.38 \times 0''.31$ at $-75°$ and 5 mJy beam^{-1} per 0.5 km s^{-1} channel for the line.

2.3. Disk Image Comparison

Images of the disk in scattered light at 1.6 μm and thermal continuum at 0.88 mm are shown in Figure 1. The SPHERE image shows a large shadow and faint arc of emission to the south, and two annular rings in the north. The ALMA image is strikingly different, with a full circle of emission around a large central cavity. It is apparent that the larger, approximately millimeter-sized, dust grains that produce the ALMA emission are confined to a ring in the outer disk, whereas the smaller, approximately micron-sized, grains that scatter infrared light are more broadly distributed, but also there must be some inner-disk feature that blocks the starlight from hitting one side of the disk.

Given the very different appearances of the two images, which differ by a factor of >500 in wavelength, we model each data set independently. We begin by analyzing the interferometer visibilities, to determine the properties of the outer disk; then we study the kinematics, to measure the central stellar mass; and finally we perform the radiative transfer models, to understand the structure of the inner disk and, in particular, how it causes the shadow–arc combination seen in the south.

3. Geometry of the Outer Disk

3.1. Continuum

The resolution of the ALMA data is about an order of magnitude lower than that of the SPHERE image, but it is still sufficient to measure the outer-disk properties. We carry out the analysis in the visibility plane, as these are direct interferometric measurements, and we avoid additional modeling steps, due to the incomplete sampling of the Fourier plane.

We model the radial intensity profile as an azimuthally symmetric Gaussian ring centered at radius $R_{\rm ring}$ with FWHM $= 2.355\sigma_{\rm ring}$:

$$I(R) = I_0 \exp\left[-\frac{(R - R_{\rm ring})^2}{2\sigma_{\rm ring}^2}\right]. \quad (1)$$

Using the `galario` package (M. Tazzari et al. 2018), we then fit the real and imaginary parts of the complex visibilities to constrain the ring parameters and disk inclination and position angle. The fit to the data is shown in Figure 2, and the inferred parameters are listed in Table 1.

The relatively large detected grains that emit most strongly at these wavelengths are distributed in a wide ring, and there is no detectable millimeter emission within ∼40 au of the star. This is indicative of a strong dust trap, similar to those found in ∼10% of young protoplanetary disks (N. van der Marel et al. 2018).

3.2. Kinematics

Although the large dust grains are confined to the outer disk, we detect CO emission within the millimeter continuum cavity. This allows us to map the disk rotation and determine the stellar mass. The first moment of the CO data cube in Figure 3 shows a velocity shift of about ±2 km s^{-1} from north to south relative to the systemic velocity.

We use the `eddy` package (R. Teague 2019) to fit a Keplerian profile to the 2D velocity field. We fix the inclination equal to $40^\circ.1$, as derived from the galario fit to the continuum visibilities, and we find a best-fit stellar mass of 0.99 $M_\odot$. The systemic velocity is 7.3 km s^{-1} and the position angle is $-7°$ (defined as the angle, east of north, of the redshifted side of the major axis). The fit and residuals are shown in Figure 4.

---

[6] https://irdap.readthedocs.io/en/latest/

[7] https://github.com/jjtobin/edisk

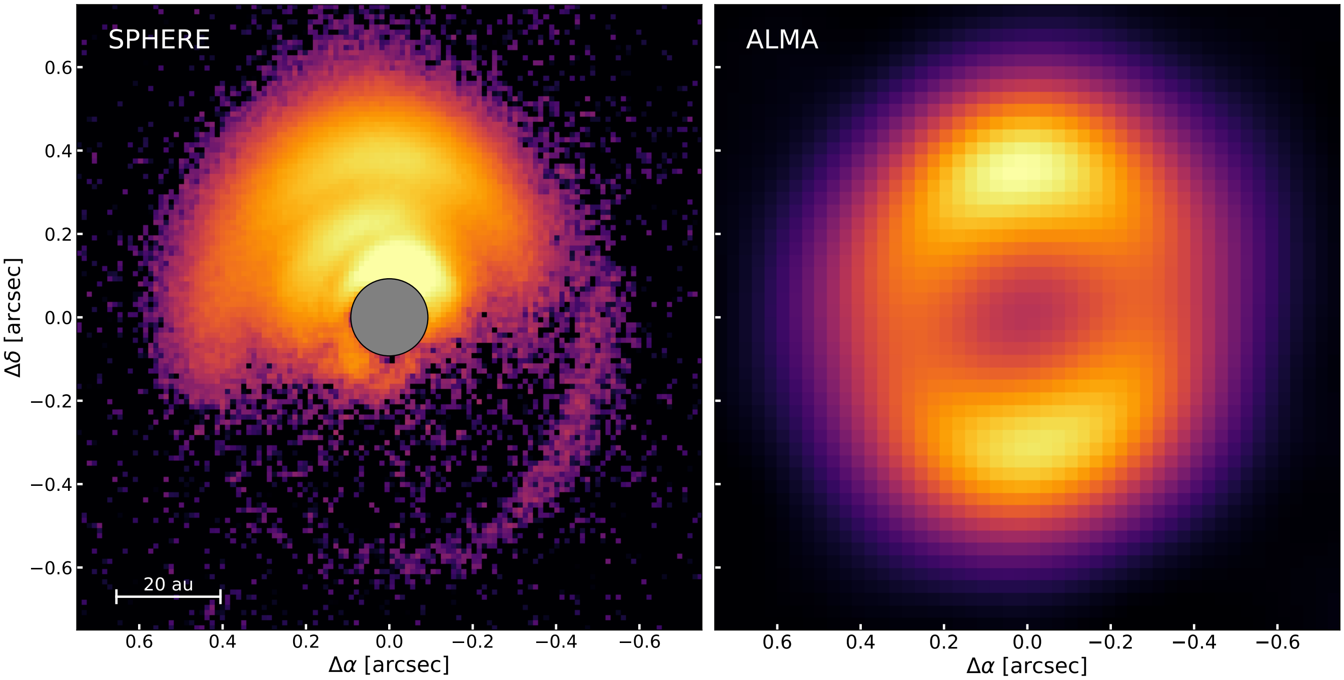

Figure 1. SPHERE and ALMA images of the disk around V1098 Sco showing, respectively, scattered light from the disk surface at 1.6 μm (left panel) and thermal emission from larger dust grains at 0.8 mm (right panel). The SPHERE image is shown on a log scale with a dynamic range of 100 and a coronagraph mask of 92.5 mas. The ALMA image is shown on a linear scale from 0 to the peak value. The resolution of the SPHERE image is about 40 mas and is about 10 times larger, $\sim 0\farcs4 \times 0\farcs3$, for the ALMA image.

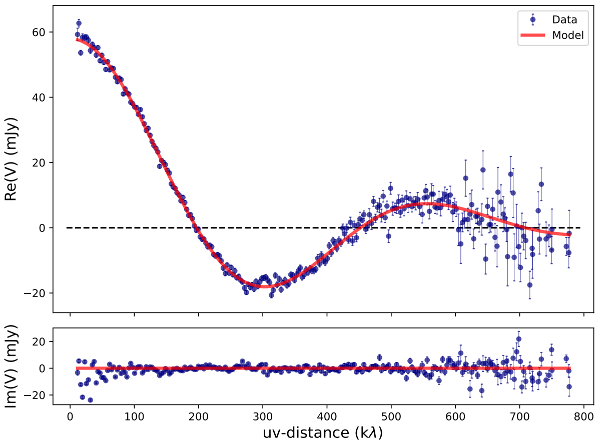

Figure 2. Observations and model fits to the real and imaginary parts of the complex ALMA visibilities.

Table 1
ALMA Continuum Fit

Parameter	Value	Error	Unit
$R_{\rm ring}$	63.0	0.1	au
FWHM	29.0	0.4	au
Incl.	40.1	0.25	deg
Position Angle	−2.0	0.4	deg

The derived stellar mass is consistent with the K1 spectral type and 5 Myr median age of Upper Sco. The position angle is consistent with the `galario` value, given the different scale heights of the millimeter continuum and molecular-gas-emitting regions and the small number of resolution elements across the disk. The formal errors in the `eddy` fit are small (less than 1%), but the true uncertainties are closer to ∼10%, due to systematics and additional effects not modeled here, including disk pressure and self-gravity (C. Longarini et al. 2025). The residual map shows that the fit is poor toward the center, suggestive of non-Keplerian motions or geometrical effects in this region. The resolution of these data is too poor to examine this further, but we surmise that it is caused by a tilted inner disk that produces the shadow in the scattered light.

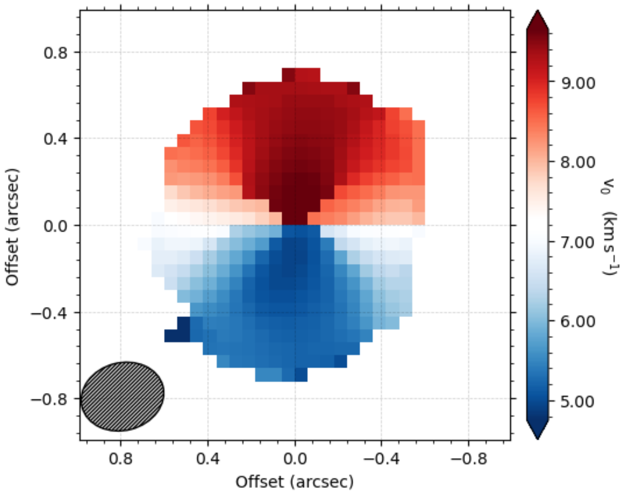

Figure 3. Disk rotation map determined from the central velocity of the CO 2–1 emission. The $0\farcs38 \times 0\farcs30$ beam size is shown in the lower left corner.

4. Radiative Transfer Modeling

We model the thermal emission and scattering from the dust, using the versatile `radmc3d` code (C. P. Dullemond et al. 2012) to fit the ALMA and SPHERE images in turn. This requires a dust density and opacity prescription. We set up the

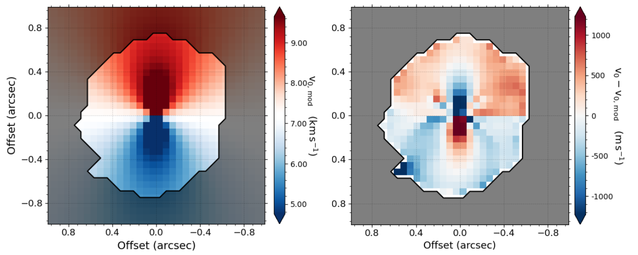

Figure 4. Model fit (left panel) and residuals (right panel) to the CO rotation map assuming a flat disk Keplerian profile.

density distribution using a power law in surface density and vertical scale height in cylindrical coordinates, (R, Z), centered on the star:

$$\Sigma_{\mathrm{dust}}(R) = \Sigma_1\left(\frac{R}{1\ \mathrm{au}}\right)^{-\alpha};\ R_{\mathrm{in}} < R < R_{\mathrm{out}}, \tag{2}$$

$$H(R) = H_1\left(\frac{R}{1\ \mathrm{au}}\right)^{\beta}. \tag{3}$$

Assuming a vertical isothermal temperature profile, the dust volume density is then given by

$$\rho_{\mathrm{dust}}(R, Z) = \Sigma_{\mathrm{dust}}(R)\exp\left[-\frac{Z^2}{2H(R)^2}\right]. \tag{4}$$

We chose this simple truncated power-law formulation rather than an exponential taper edge for simplicity. In practice, since the scattered light falls off as the inverse square of radius and the resolution of the ALMA data is low, we are not able to constrain the falloff in surface density at the outer edge of the disk.

We model the dust as consisting of two size populations (T. Birnstiel et al. 2012), with 0.1 μm and 300 μm grains and opacities calculated for a composition of amorphous olivine with 50% Mg and 50% Fe (C. Jaeger et al. 1994; J. Dorschner et al. 1995), using the `makeDustOpac` module in `radmc3d`. Each population follows the same prescription described above, but with different parameter sets $\{\Sigma_1, \alpha, R_{\mathrm{in}}, R_{\mathrm{out}}, H_1, \beta\}$, which we preface with the superscripts "small" or "big."

4.1. Large Grains in the Outer Disk

The small grains have negligible thermal emission at millimeter wavelengths, and the large grains do not efficiently scatter the short-wavelength starlight. We set the mass fraction to be 10% small grains and 90% large grains, based on models that match the infrared–millimeter spectral energy distribution (SED) of typical disks (P. D'Alessio et al. 2006), though the actual ratio is not critical to this work, as we are mainly interested in determining the geometry, rather than the mass or composition of the disk. We then adjust the density prescription parameters for the large grains to find a representative fit to the ALMA map, then use that to inform the outer-disk parameters when carrying out a more detailed fit to the SPHERE map in the following section.

To model the ALMA continuum image, we fixed the inclination to 40° and the position angle to $-5°$, based on the `galario` and `eddy` fits, and we ran a grid of models varying $\{\Sigma_1, \alpha, R_{\mathrm{in}}, R_{\mathrm{out}}\}^{\mathrm{big}}$. The quality of each model was assessed by calculating the sum of the squared intensity residuals, with the lowest value at $\Sigma_1^{\mathrm{big}} = 2.2\ \mathrm{g\ cm^{-2}}$, $\alpha^{\mathrm{big}} = 1$, and a radial extent from $R_{\mathrm{in}}^{\mathrm{big}} \simeq 40$ au to $R_{\mathrm{out}}^{\mathrm{big}} \simeq 85$ au. The total dust mass is $42\,M_\oplus$. As the emission at these wavelengths is optically thin and the disk is seen fairly face-on, our models are insensitive to the vertical scale-height parameters, H_1^{big} and β^{big}, at least at the resolution of the data. A comparison of the data and model is shown in Figure 5.

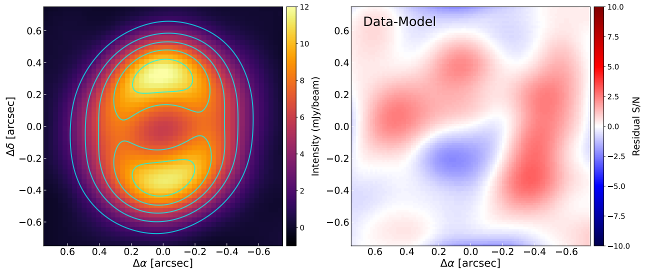

Figure 5. Comparison of radiative transfer model to the ALMA continuum image. The left panel shows the observed intensities on a linear stretched color scale, with the model overlaid in contours from 2 to 12 mJy beam^{-1}. The right panel plots the residuals in units of the rms noise level in the data.

Given the moderate resolution and dynamic range of the ALMA data, we do not perform a statistically rigorous fit, as these constraints on the outer disk are sufficient for the main goal of measuring the properties of the inner disk.

4.2. Geometry of the Tilted Inner Disk

We now proceed to model the shadow seen in the SPHERE scattered-light image. Given the symmetry of the ALMA continuum and CO maps, and the residuals to the Keplerian fit, we consider a misaligned inner disk. Disk shadows have been seen in scattered-light images of several protoplanetary disks and provide a window into the inner disk, often at scales that are hard to image in other ways (A. J. Bohn et al. 2022). For V1098 Sco, the shadow extends across an entire half of the disk, similar to HD 143006 (M. Benisty et al. 2018) and HD 139614 (G. A. Muro-Arena et al. 2020), and we consider a similar geometric arrangement here, schematically illustrated in Figure 6.

In this case, we see an additional feature, in the form of a faint arc to the south in the otherwise shadowed region. We rule out a scenario where this comes from the outer radius of the front side of the disk, just beyond a shadowed region, as the arc does not extend all the way around the disk, nor is it contiguous with the outer parts of the unshadowed northern disk. Instead, we consider it to arise from the backside of the disk. By symmetry, a tilted inner disk that shadows the southern half of the front side of a flared disk would shadow the northern half of the backside but leave the southern half fully illuminated. As in IM Lup (H. Avenhaus et al. 2018), we see the extincted outer edge of this backside but in this case only as a partial arc. This fortuitous arrangement provides an extra, critical, constraint on the inner-disk geometry.

We fix the parameters for the large-grain population based on the ALMA fit shown above and use the same surface density normalization and power-law index. As additional parameters are necessary to describe the inner disk, we first define the overall radial extent. The models are insensitive to the inner edge, $R_{\mathrm{in}}^{\mathrm{small}}$, as long as it lies within the coronagraph and we fix it to 0.5 au. The outer radius, $R_{\mathrm{out}}^{\mathrm{small}}$, simply scales the extent of the emission and we set it to 96 au (0.″6) by eye.

To create the shadow, we break the disk at a tilt radius, R_{tilt}, and incline and rotate the inner disk at angles i_{tilt} and $\mathrm{PA}_{\mathrm{tilt}}$,

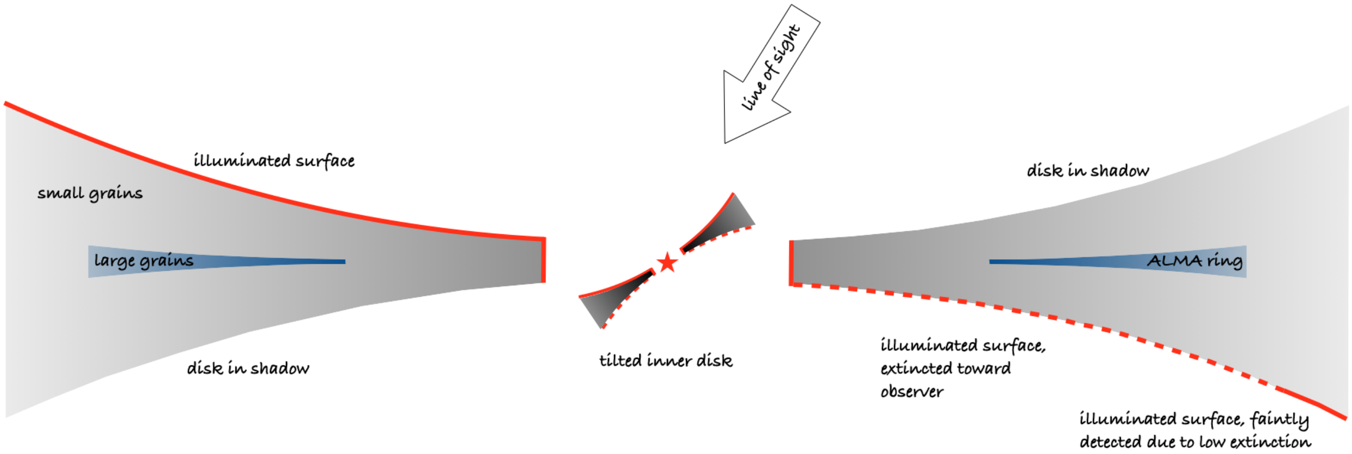

Figure 6. Edge-on cutaway of the inner- and outer-disk geometry. Our viewing angle is from the top right corner, such that we see scattered light from the illuminated northern (left) side but not from the shadowed southern (right) side, except for faint emission from the outer parts of the highly extincted backside of the disk. The flattened thick ring of large dust grains seen in the ALMA continuum image is illustrated by the blue region in the midplane.

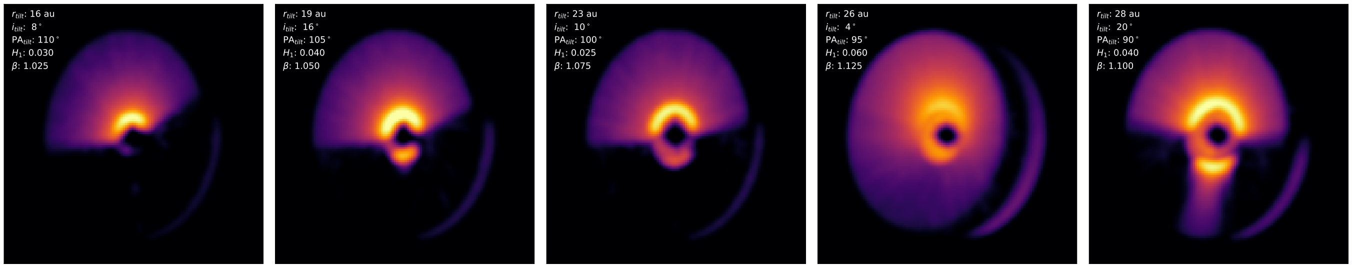

Figure 7. A representative set of images showcasing the range of scattered-light morphologies in the model grid. The parameter values for each model are shown in the top left corner of each panel. The central region is masked out, to mimic the SPHERE coronagraph. Additional plots showing the systematic effects of changing a single parameter one at a time are shown in the Appendix.

respectively. This requires a change to a 3D spherical coordinate system (r, θ, ϕ). The tilt radius is constrained to lie within the large-grain cavity of 40 au, so that the tilted inner disk is entirely composed of small grains. We note that the SPHERE image shows a bright ring of emission around the coronagraph that extends into the otherwise shadowed southern half, though with a marked asymmetry, which is likely to be either the front side of the tilted inner disk or the inner edge of the outer disk illuminated, due to light coming through the dust sublimation hole in the inner disk. The coronagraph hole corresponds to 7.5 au in radius, and we find that our models require a larger $R_{\rm tilt}$ to match the extent and sharpness of the shadow, which implies that we can indeed directly see the inner disk.

The five free parameters[8] $\{R_{\rm tilt}, i_{\rm tilt}, {\rm PA}_{\rm tilt}, H_1$, and $\beta\}$ that describe the inner-disk geometry and disk vertical structure each affect the scattered-light image in distinct ways. A small set of representative models that vary multiple parameters at once is shown in Figure 7. Figures A1–A4 in the Appendix contain multiple plots where only one parameter is changed at a time, to more clearly illustrate their individual effects.

Some general trends are that $R_{\rm tilt}$ affects the location of the bright ring of emission closer to the star and how clearly it is seen on the southern side. Low $i_{\rm tilt}$ produce a blurry shadow boundary, while high $i_{\rm tilt}$ produce a tongue of scattered light in the shadowed region where starlight escapes through the central hole (we have verified that this effect is seen for different $R_{\rm in}$ greater than the dust sublimation radius of $\sim$0.1 au and essentially independent of its precise value). PA$_{\rm tilt}$ simply rotates the shadow boundary and arc. The flaring parameters, H_1 and β, affect the illumination on the northern side of the disk, the angle of the shadow boundary, and the brightness of the arc.

We find that we can produce models that meet the "eyeball test" of qualitatively matching the data. To find the best set of parameters that match the observations, we run a large grid of models stepping through each parameter. The grid ranges and step sizes are shown in Table 2 and amount to 34,560 models in total.

Whereas we could readily calculate a likelihood for the model fit to the ALMA data, based on the squared difference of the intensities, this is much harder to do for the SPHERE image, for several reasons. First, the disk shows substructure, in the form of a second ring farther out from the presumed inner–outer disk boundary. Though interesting in its own right, it would require additional parameters to model, which would add to the complexity without aiding the solution of what is causing the disk shadow. Second, the scattering is angle-dependent, but modeling the phase function requires additional parameters for the dust opacities, and again this complicates rather than aids the solution. Finally, the faint arc from the backside provides important information, but it is about 100 times fainter than the disk from the front side and would therefore be comparably downweighted in any intensity-based

---

[8] We drop the suffix "small," for readability.

Table 2
Model Grid Parameters

Parameter	Range	Step	Unit
$R_{\rm tilt}$	[15, 30]	1	au
$i_{\rm tilt}$	[4, 20]	2	deg
$\mathrm{PA}_{\rm tilt}$	[90, 110]	5	deg
H_1	[0.025, 0.060]	0.005	au
β	[1.025, 1.150]	0.025	⋯

comparison. Moreover, the brightness is a combination of backscattering from the dust surface and extinction by dust in the disk interior, so, again, it is sensitive to disk properties that are unrelated to our principal goal.

Instead, we focus on matching the scattered-light morphology, to compare the models to the observations. We define boundaries for the northern disk and southern arc based on contouring the SPHERE image. We then use a local threshold algorithm to define the regions of emission in the model. In this way, we reduce the data and models to binary maps and can readily quantify the goodness of fit by maximizing the likelihood defined by the sum of the squared-difference map (equivalent to the sum of the absolute differences). The procedure is graphically illustrated in Figure 8. Various thresholding algorithms are implemented in the Python image-processing package `scikit-image`, and by experimentation we settled on the triangle method, where the threshold is automatically determined from the normalized histogram of pixel values (G. W. Zack et al. 1977).

Using this technique, we are able to quantify the model fits to the SPHERE image and constrain the range of parameters that provide good fits. This shows that the model design of a tilted inner disk works well and that we can constrain its properties. However, to determine the best-fit parameter uncertainties and understand their interdependences requires a new approach, since the radiative transfer model is computationally expensive. This motivates a machine learning approach.

4.2.1. Model Image Interpolation Using an Artificial Neural Network

Each `radmc3d` image is produced by propagating photons through the disk in random directions and modeling the scattering with dust particles, by sampling from the allowed range of quantum outcomes. To reduce the inherent Poisson noise in this method requires a large number of photons, with a correspondingly high computational demand. For the initial parameter exploration, we use 10^6 photon packets, and each model takes almost 2 minutes to run. Even with multiple cores, the grid of ~35,000 models takes many days to produce. To carry out an MCMC parameter search would likely require at least this many and perhaps as many as an order of magnitude more models to achieve statistical convergence. Furthermore, any tweaks to the model design would require a restart. In short, this approach is extremely time-intensive and, consequently, a Bayesian approach to radiative transfer modeling is rarely used (P. D. Sheehan & J. A. Eisner 2017; G. Columba et al. 2024).

An alternative approach is to interpolate model images at intermediate parameter values between the grid points. J. P. Williams & C. McPartland (2016) carried out an ad hoc formulaic approach on a set of simulated ALMA CO data cubes, but more sophisticated techniques are now available that can better handle the complexities of multiparameter image production and, equally importantly, test and verify the outcome (simulation-based inference: K. Cranmer et al. 2020). Here, we train an artificial neural network on radiative transfer images and use it to produce new models at any set of parameter values. Our approach is similar to Á. Ribas et al. (2020), T. Kaeufer et al. (2023), and A. Ruzza et al. (2024), who used neural networks trained on SED models and hydrodynamical simulations, respectively.

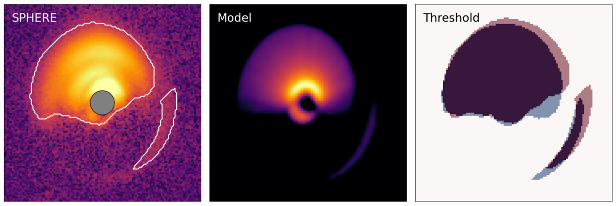

Figure 8. Illustration of how the models are compared to the observations. The left panel shows the SPHERE image, with borders defined by contouring, to highlight the illuminated northern side of the disk and the faint southern arc. A representative model is shown in the middle panel. An autothresholding algorithm then defines the regions of emission. The right panel shows binary images of the emission for the observations (red) and model (blue), which are used to quantify the goodness of fit.

The network architecture is diagrammed in Figure 9. The input is a vector composed of the five model parameters $\{R_{\rm tilt}, i_{\rm tilt}, \mathrm{PA}_{\rm tilt}, H_1, \beta\}$, and the output is a vector with 1600 elements that we reshape into a 40 × 40 image. The input and output layers are connected through two hidden layers with increasing numbers of neurons, to enable the scaling from five parameters to 1600 pixel values. Each neuron in a layer is related to each neuron in the preceding layer by a multiplication factor (weight), an offset (bias), and a nonlinear activation function (in this case, "leaky ReLu" filters). The network is trained (i.e., weights and biases are iteratively determined) by comparing the outputs (predictions) to an image data set (targets) produced with known input parameters. The comparison is quantified via a loss function that calculates the mean squared error between the network predictions and training targets.

Rather than train the network on the existing grid of models, we sample the parameter space over the same range in Table 2, using a Sobol sequence, as in T. Kaeufer et al. (2023). This ensures that the parameter set in each model is different from every other model, which is both efficient and mitigates possible overlearning errors from duplicate parameter subsets. Indeed, we found that we could achieve a substantially lower space-filling discrepancy than the large grid with just a few hundred Sobol samples. We took advantage of this to increase the signal-to-noise ratio of each radiative transfer model, by running `radmc3d` with 5×10^6 photon packets, fives times more than the grid models, producing a training set of 1024 images.

We found that the loss function decreased rapidly over a few hundred iterative epochs but then more slowly after about a thousand epochs. However, there were occasional spikes, which we found to be due to the amplification of noisy pixels in the radiative transfer models toward the southern (shadowed) side of the disk. The number of these spikes in the loss function ticked up slightly beyond 2000 iterations. We therefore used the network weights and biases as determined from 2000 epochs for our model fitting.

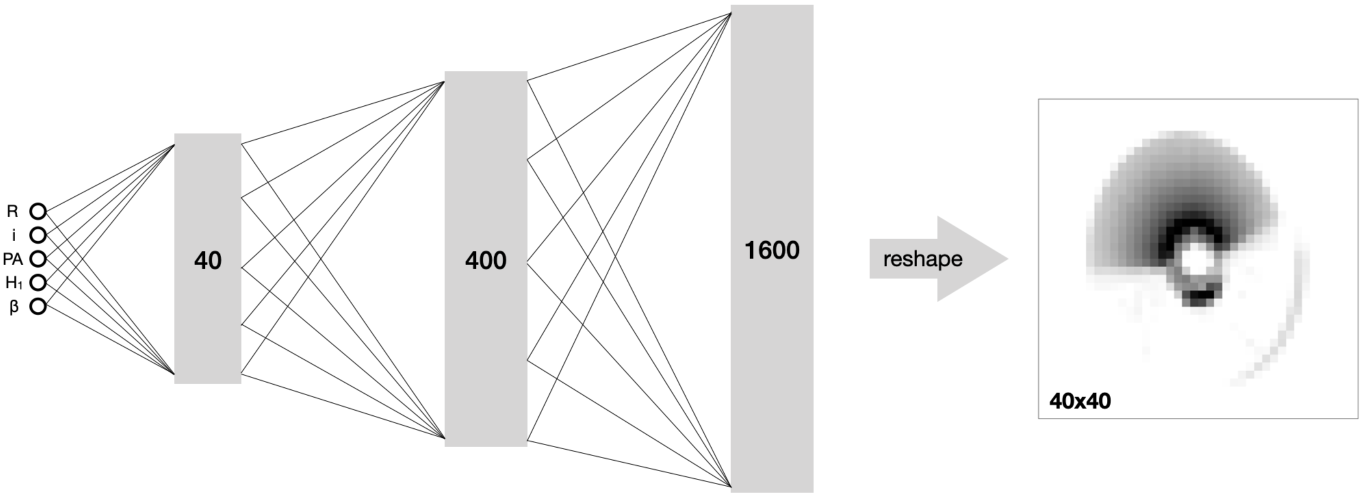

Figure 9. The architecture of the neural network used to produce images from the five input parameters of the model. There are three layers in all, represented by the gray rectangles and labeled by the number of output nodes. The interconnections between them are schematically shown (in practice, each output node of one layer connects to every input node of the next layer). The final layer produces a vector with 1600 elements, which we reshape into a 40 × 40 pixel image.

To demonstrate the network performance, we carried out a validation test, by comparing an independent set of 128 radiative transfer models with neural network images from a different Sobol sequence. The results of this validation procedure are shown in Figure 10. Here, the mean and standard deviation are calculated over the 40 × 40 pixels in each image, which is about a factor of 2 times greater than the area of the physical disk projection. In a few cases, one or a small group of "hot" pixels caused by the aforementioned amplification of noise in the southern disk were clipped in the calculation. Higher-signal-to-noise-ratio radiative transfer models might be required to train neural networks for detailed intensity modeling across a high dynamic range, but the effect is mitigated here, by the autothresholding algorithm illustrated in Figure 8 that we use to fit the data.

The accuracy of an extrapolation from five inputs to a 1600-element vector seems at first surprising, but there is a high correlation between neighboring pixels, due to the smoothness of the physical disk structure. As a direct visual comparison, Figure 11 plots the neural network predictions for the representative set of grid images shown in Figure 7. Allowing for the lower resolution of the neural network output, the similarity with the `radmc3d` images across the full range of disk morphologies found in the grid is quite striking, although an example of a "hot" pixel is seen in the fifth panel, where the network is not as well trained for the relatively rare cases where there is a narrow region of scattered light in the southern disk.

We experimented with the network design by changing the number of nodes in each layer, the number of layers, and the activation function. We found that additional layers produced similarly low loss functions but with a slightly higher proportion of (still occasional) spikes and no significant differences in the validation tests. We therefore settled on the relatively simple architecture in Figure 9.

Modern Python packages such as `PyTorch` (used here) and `TensorFlow` provide straightforward setups of artificial neural networks, opening up a new and efficient pathway to fit radiative transfer models to observations. The computational time required to produce high-quality images decreases from minutes to milliseconds. This then allows a Bayesian approach to parameter estimation.

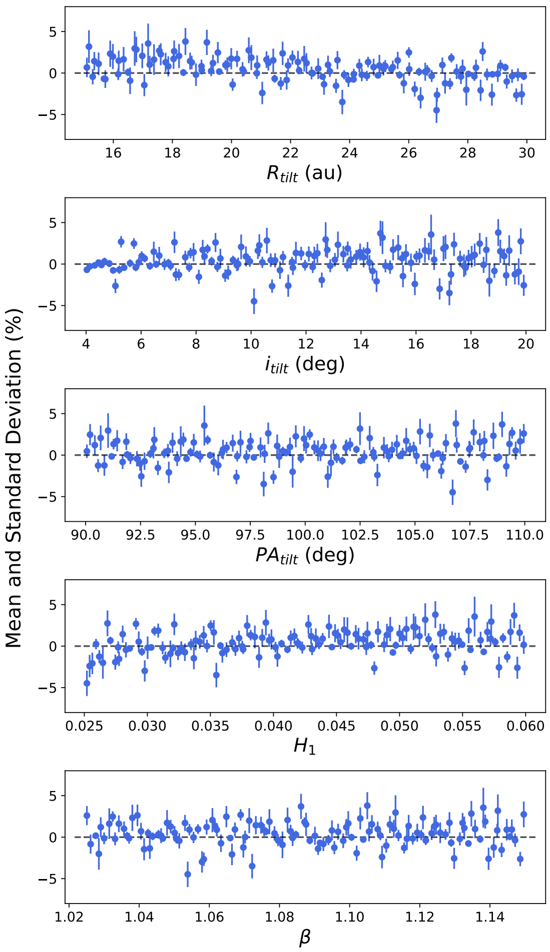

Figure 10. Validation test of the interpolation procedure. The mean differences between the images produced by radiative transfer and the neural network are shown for 128 models at random parameter values within the grid boundaries.

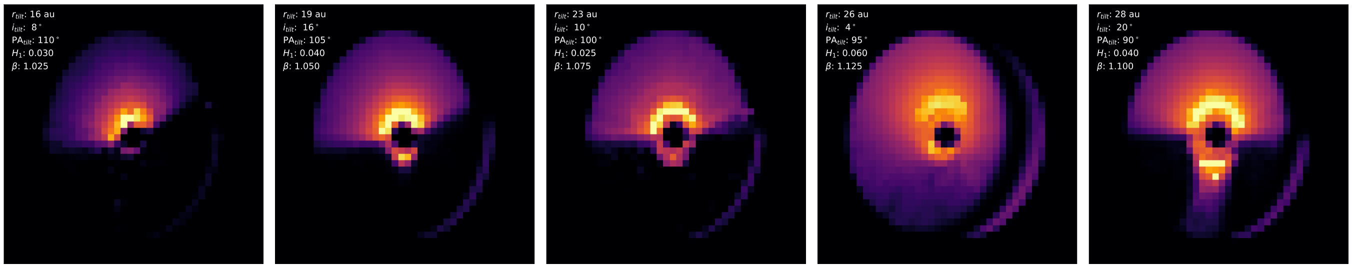

Figure 11. A set of images produced by the neural network for the same parameter sets used in Figure 7, showcasing the range of disk morphologies in the large model grid. The size of each image is 40 × 40 pixels, three times lower than the `radmc3d` images, but the agreement is otherwise very good.

4.2.2. MCMC Model Fitting

With the ability to quickly produce model images for any set of physical parameters and to use image-processing techniques that identify areas of emission and allow quantitative comparisons with data, we can search for the best fit through an MCMC analysis. We use the `emcee` package (D. Foreman-Mackey et al. 2013), uniformly sampling across the parameter extents of our model grid in Table 2, with 20 walkers.

The MCMC chains converged rapidly, and we show the results for 10,000 steps after discarding the first 2000 from the burn-in phase as a corner plot in Figure 12. The contour plots show a correlation between the scale height and flaring index, and a slight dependence with the position angle of the inner-disk tilt. The histograms of the individual parameters are nearly symmetric, and we use the median and maximum difference with the 16th and 84th percentiles to provide an approximate mean and standard deviation for each parameter in Table 3. The precision with which each parameter is determined belies the error in matching the observations, which is mainly due to the validity of the model description rather than the statistical uncertainties in the fit.

Figure 13 shows the maximum likelihood model in comparison with the SPHERE image. The parameters of this model —$R_{\rm tilt} = 23.4$ au, $i_{\rm tilt} = 13\overset{\circ}{.}0$, $\mathrm{PA}_{\rm tilt} = 103\overset{\circ}{.}0$, $H_1 = 0.048$ au, and $\beta = 1.090$—are very similar to the median values from the marginalized histograms in Table 3. With the exception of the disk substructure and scattering phase-angle dependence, neither of which we attempted to model, the agreement is very good, showing that the five parameters for disk size, flaring, and orientation are sufficient to capture the essential features in the image and allow us to determine the geometric properties of the inner disk.

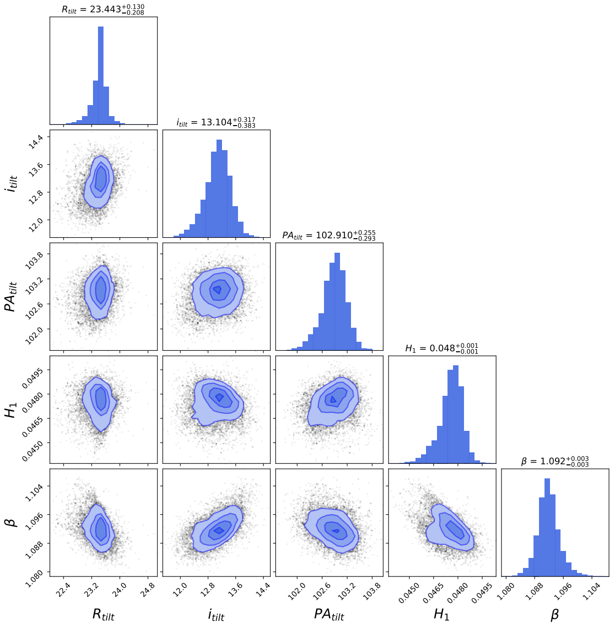

Figure 12. Corner plot of the parameter distributions from the MCMC fitting procedure. The histograms at the top of each row show the distributions of a single parameter for all sets of other parameter values. The contour plots show how the fits for each pair of parameters depend on each other.

Table 3
Inner-disk Fit

Parameter	Value	Error	Unit
$R_{\rm tilt}$	23.4	0.2	au
$i_{\rm tilt}$	13.1	0.4	deg
PA$_{\rm tilt}$	102.9	0.3	deg
H_1	0.048	0.001	au
β	1.092	0.003	⋯

5. Discussion

We have carried out a step-by-step modeling procedure on ALMA and SPHERE observations of V1098 Sco to explain the remarkable shadow feature that obscures the southern half of the disk. The ALMA continuum observations show a ring centered at a radius of 63 au around a central cavity depleted of large dust grains within about a 40 au radius. The CO map constrains the stellar mass to about 1 $M_\odot$ but shows residuals near the center that we attribute to a tilted inner disk. Using these results, we fix the inclination and position angle for the outer disk to $i_{\rm outer} = 40°$ and $\mathrm{PA}_{\rm outer} = -5°$, respectively, and we were then able to match the morphology of the scattered-light emission with a tilted and misaligned inner disk.

The inferred size of the inner disk has a radius of 23 au, which is large enough that the inner edge of the outer disk lies beyond the SPHERE coronagraph (7.5 au radius), appearing as a bright ring in the north and a fainter ring in the south, but also lies comfortably within the central hole seen in the ALMA image. The inner disk is inclined by about 13°, which is more face-on than the outer disk but is most different in position angle, with a value of 103° for the angle of the major axis relative to north (with a 180° geometric degeneracy). This strong inner-disk misalignment explains why the shadow boundary runs from east–west almost orthogonal to the near-north–south line of the outer-disk major axis. If the inner disk were tilted but had the same position angle as the outer disk, the shadow boundary would be parallel to the major axis.

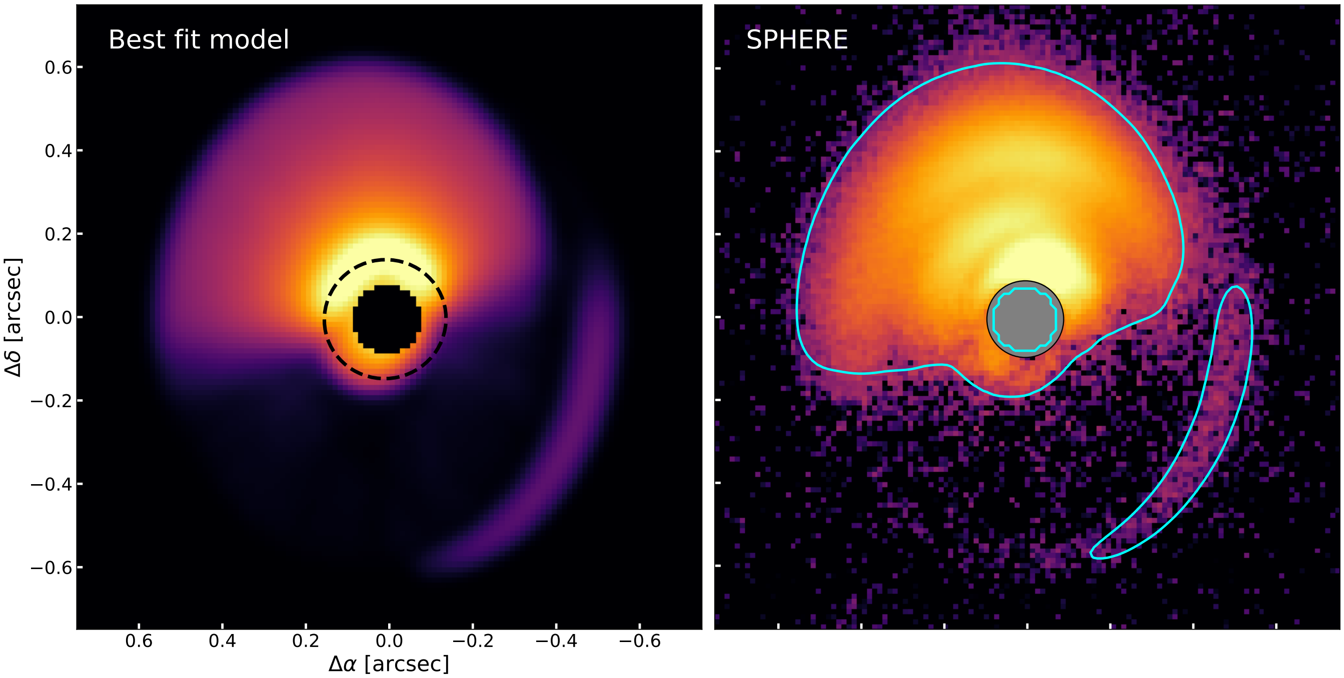

Figure 13. Model image for the maximum likelihood model from the MCMC fit (left panel). The dashed ellipse in the model image shows the break radius between the inner and outer disks at 20 au. The SPHERE image is overlaid with a low-level contour of the model, to show the level of agreement between the extent of the emitting regions.

The more face-on orientation of the inner disk implies smaller projected rotation velocities, explaining why the Keplerian fit residuals are opposite to the sense of the overall disk rotation in Figure 4. Under the reasonable assumption that the inner and outer disk rotate in the same direction, the position angle for the redshifted side of the inner disk is then −77°, using the same definition as that used for the outer disk. We then conclude that the difference in the projected angles of the inner- and outer-disk rotation axes is $\Delta\mathrm{PA} = 72°$, as illustrated in Figure 14. From M. Min et al. (2017), we calculate the misalignment angle between the outer disk and tilted inner disk to be

$$\mu = \cos^{-1}[\sin(i_{\mathrm{outer}})\sin(i_{\mathrm{tilt}})\cos(\Delta\mathrm{PA}) + \cos(i_{\mathrm{outer}})\cos(i_{\mathrm{tilt}})] = 38°. \quad (5)$$

A misalignment between the inner- and outer-disk rotation axes requires a gravitational torque from a second source in addition to the central star. Binary stars can warp a circumstellar or circumbinary disk (K. Batygin 2012; S. Facchini et al. 2013) and, for the latter case, a small inner disk around one of the stars can cast shadows on the outer disk (D. J. Price et al. 2018). Simulations show that a range of scattered-light morphologies can be produced including large shadows that closely resemble our SPHERE observations (S. Facchini et al. 2018).

There are no clear signatures of a binary companion. V1098 Sco lies within a fairly crowded star field, and there are several faint optical sources within an arcminute, but there are no Gaia sources at the same distance. The large size of the outer disk argues against a companion within several hundred astronomical units ($\lesssim 2''$). The Gaia DR3 renormalized unit weight error of 1.648 is slightly higher than typical values for single stars (A. Castro-Ginard et al. 2024), but this could be attributed to the presence of scattered light from the disk (S. Fitton et al. 2022). The CO velocities analyzed in Section 3.2 show that the central mass is $1\,M_{\odot}$, which is consistent with the K1 spectral type and an age of ∼5 Myr of the star, and rules out a close equal-mass binary.

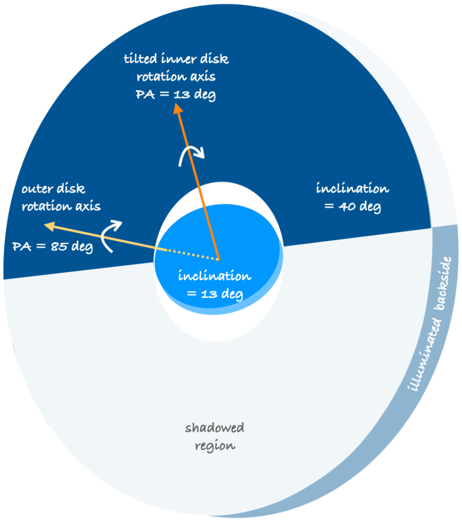

Figure 14. Geometry of the disk system inferred from the ALMA and SPHERE modeling. The outer disk is moderately inclined, with the major axis running approximately north–south. The inner disk is more face-on and almost orthogonally oriented, such that the shadow boundary runs east–west. The inferred misalignments between the rotation axes of the inner and outer disk are shown by the orange and yellow arrows, respectively.

Nevertheless, the observations allow for a low-mass close companion that is too faint to change the spectral type and either lies behind the coronagraph, is enshrouded in dust, or is otherwise undetectable in the SPHERE image. C. Nixon et al. (2013) model the evolution of misaligned gas disk around a binary and show that the break radius between the inner and outer disk depends on the binary separation and mass ratio.

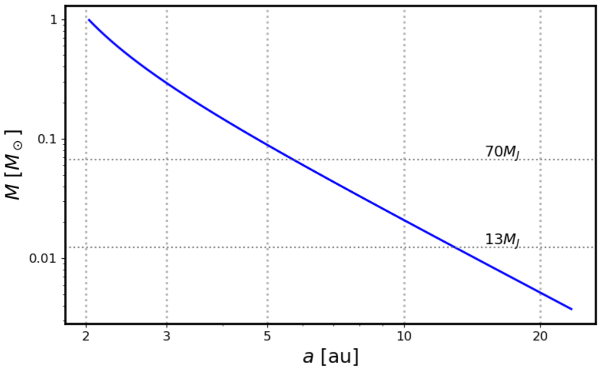

Figure 15. The relation between the mass and orbital radius of a companion that would break the disk at 23 au. The horizontal dotted lines mark the substellar and planet boundaries at 70 and 13 Jupiter masses, respectively.

Although they focus on disks around supermassive black holes, they also provide a formula for a pressure-dominated disk in an appendix that is applicable to protoplanetary disks. In our notation, this translates to a constraint on the putative companion orbital radius, a, and mass, M:

$$a\left(\frac{M}{1\,M_\odot + M}\right)^{1/2} \gtrsim R_{\rm tilt}\left(0.75\sin 2\mu\,\frac{R}{H}\right)^{-1/2}, \qquad (6)$$

where we have substituted the mass of the primary from Section 3.2. Our modeling indicates that the aspect ratio at the edge of the inner disk is small, $H/R \simeq 2.7 \times 10^{-3}$, and, together with our measurement of the misalignment angle, $\mu = 38°$, the right-hand side equates to 1.4 au. We plot the relation between the companion mass and orbital radius in Figure 15. This holds as long as we are in the wave-like regime of warp propagation, which requires a low viscosity, $\alpha < H/R$.

Once a companion opens up a gap, the disk breaks into an inner and outer region, which then independently dynamically evolve (R. Nealon et al. 2018). A nonzero planet obliquity can then drive a tilt mismatch between the inner and outer disks on Myr timescales, which then precess at different rates and become misaligned (J. E. Owen & D. Lai 2017; Z. Zhu 2019). Future higher-resolution studies of the CO kinematics can provide stronger constraints on the mass and location of any perturbing object or perhaps lead to consideration of multiple objects.

In general, shadowed disks are interesting sources for studying disk physics and hunting for protoplanets. They systematically have high near-infrared excesses indicative of large inner disks (A. Garufi et al. 2018), and ALMA imaging often shows an inner cavity and outer continuum ring that can be explained by a super-Jupiter-mass object that prevents millimeter-sized grains from drifting inward (e.g., A. Pohl et al. 2017; M. Villenave et al. 2019). Detailed multiwavelength radiative transfer models can help us better quantify the properties of these complex sources and guide searches for young planets.

6. Summary

We have imaged the scattered light from the disk around the T Tauri star V1098 Sco and found a large shadow that covers its entire southern half. Using ALMA archival data, radiative transfer models, and machine learning to speed up the image production, we show that the morphology of the scattered-light emission can be explained by a disk that is broken into inner and outer regions, which are misaligned in both inclination and position angle. The main results of this work are as follows:

1. The ALMA continuum image shows a central cavity with radius ∼40 au that is devoid of large (approximately millimeter-sized) dust grains. The CO velocity map constrains the mass of the central star to be 1 $M_\odot$. We use these ALMA observations to fix the inclination and position angle of the outer disk.
2. We produce a grid of radiative transfer models and show that a tilted inner disk of small (approximately micron-sized) dust grains can qualitatively reproduce the large shadow and faint arc features in the SPHERE scattered-light image of the disk.
3. We use an autothresholding technique borrowed from text recognition algorithms to identify the salient features in the model images and thereby quantitatively compare with observations without the extraneous details of the disk substructure and scattering phase function.
4. We design a neural network to efficiently produce model images for arbitrary parameter values in between the discrete model grid points. The tremendous increase in speed allows forward modeling of the SPHERE image.
5. We constrain the size, inclination, and position angle of the inner disk. Its radius of 23 au is about half that of the ALMA cavity size. The inner disk is more face-on and twisted, such that its rotation axis is offset by 38° from that of the outer disk.

The scattering of starlight from a circumstellar disk depends not only on the spatial distribution of the dust, but also on its optical properties. The former is more meaningful for understanding planet formation. We have shown that simple image-processing techniques that delineate morphology, coupled with machine learning tools that speed up radiative transfer modeling, provide a Bayesian pathway to quantitative modeling of disk observations with SPHERE. We hope that the methodology described here proves useful for the analysis of other data in the future.

Acknowledgments

We first acknowledge the referee for an insightful report and the suggestion to use Sobol sequences to sample the parameter space when training the neural network. We thank Kees Dullemond for providing the code to warp the inner disk that we used in the radmc3d modeling. Thanks also to Richard Nelson and Jaehan Bae for helpful discussions on the various dynamical processes that misalign disks; these conversations occurred at the Kavli Institute for Theoretical Physics (KITP) and were supported by grant No. NSF PHY-2309135. This work was also supported by NASA through grant No. 23-XRP23 2-0148. J.P.W. thanks NASA and the NSF for supporting basic research that actually does help make America great. M.B. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (PROTOPLANETS, grant agreement No. 101002188). This paper is based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO program 1104.C-0415(D) (DESTINYS Large Program). G.L. acknowledges support from PRIN-MUR 20228JPA3A and from the European Union Next Generation EU, CUP: G53D23000870006. This paper makes use of the following

ALMA data: ADS/JAO.ALMA#2018.1.00564.S. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSTC 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. We thank the VLT and ALMA staff astronomers for making these observations possible.

Author Contributions

J.P.W. carried out the analysis and wrote the manuscript. M.B. initiated the project and provided feedback throughout. C.G. coordinated the SPHERE observations and reduced the data. G. L. provided theoretical input on the interpretation of the modeling results. M.V. helped with the early stages of the project.

Facilities: VLT: Melipal, ALMA.

Software: galario (M. Tazzari et al. 2018), eddy (R. Teague 2019), radmc3d (C. P. Dullemond et al. 2012), emcee (D. Foreman-Mackey et al. 2013), astropy (Astropy Collaboration et al. 2013, 2018, 2022), matplotlib (J. D. Hunter 2007), scipy (P. Virtanen et al. 2020), scikit (F. Pedregosa et al. 2011), PyTorch (A. Paszke et al. 2019).

Appendix
Model Grid Images

We present a gallery of simulated SPHERE images from the model grid, as in Figure 7, but we change one parameter at a time (two for the scale-height variation), while keeping the others fixed close to the interpolated best-fit values in Table 3.

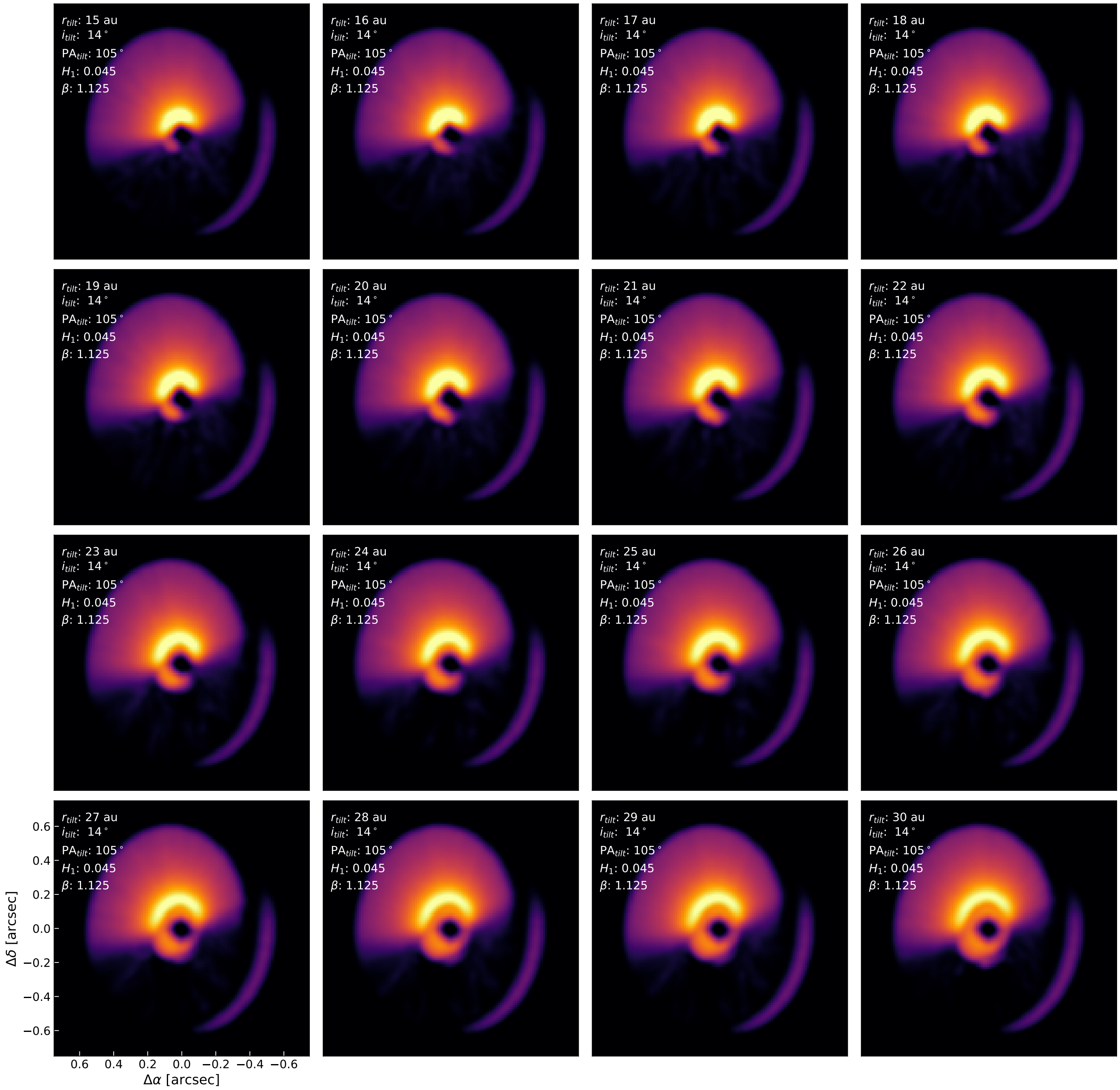

Figure A1. Variation of the inner-disk radius, $R_{\rm tilt}$.

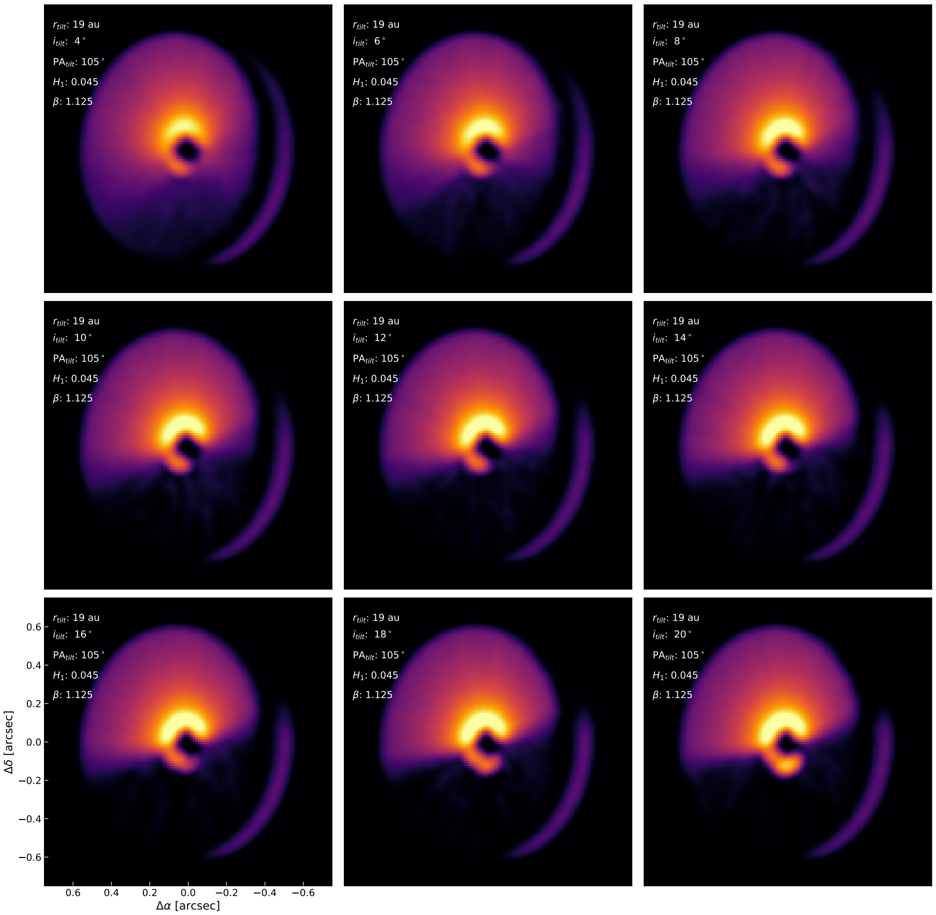

Figure A2. Variation of inner-disk inclination, i_{tilt}.

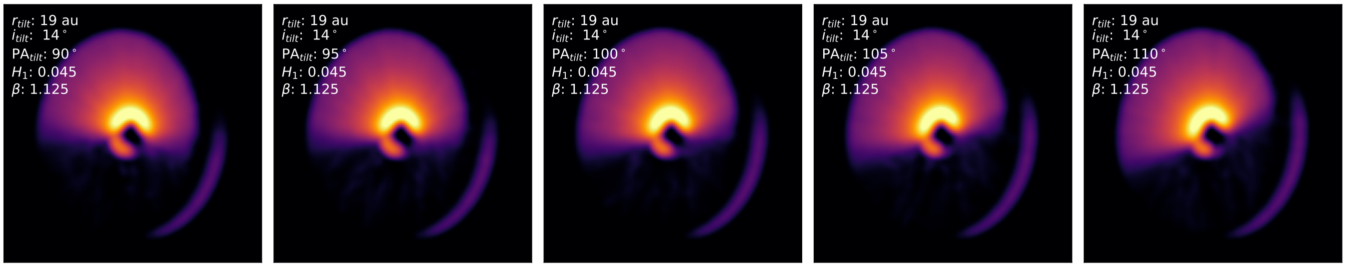

Figure A3. Variation of inner-disk position angle, PA$_{\text{tilt}}$.

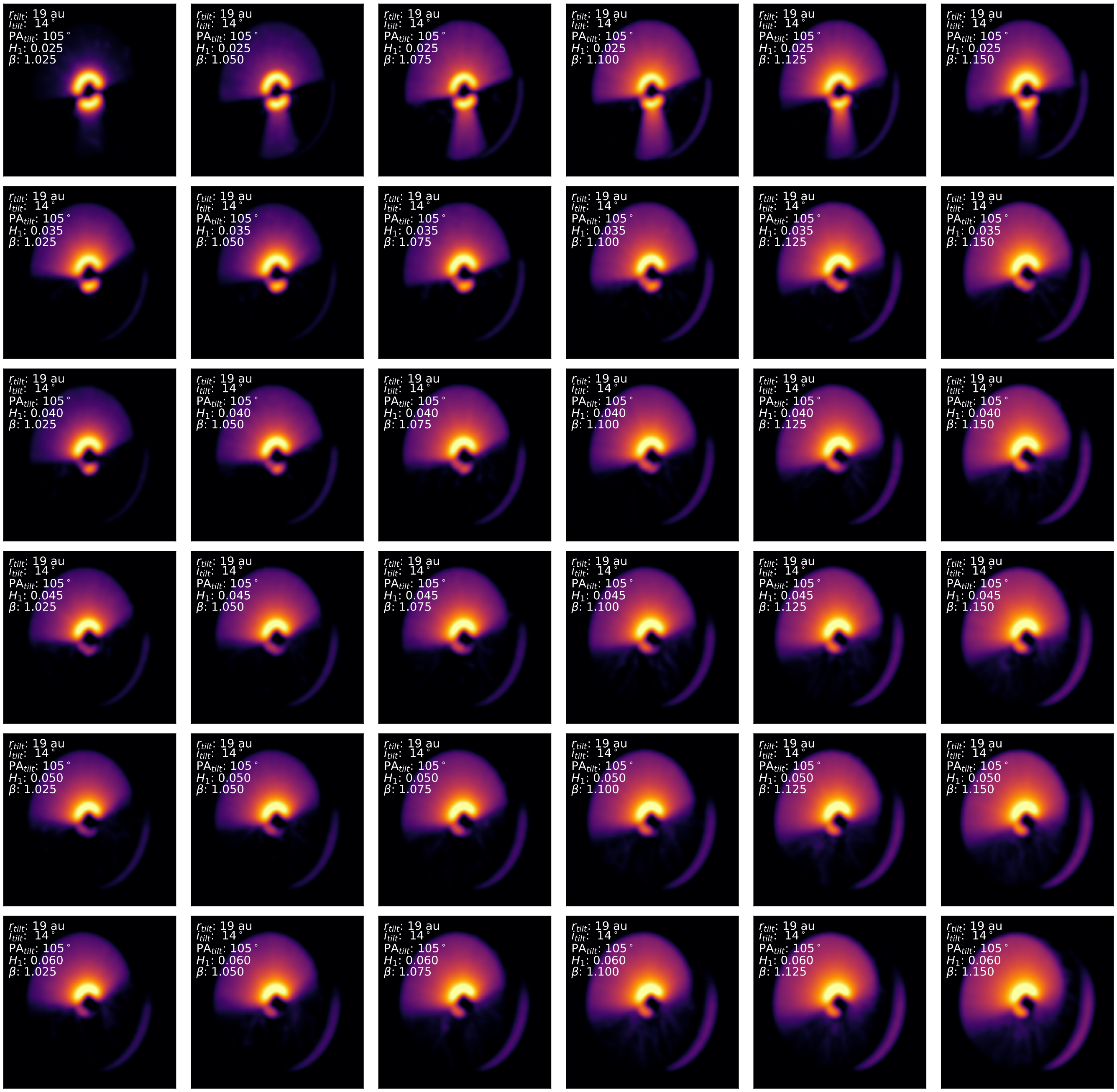

Figure A4. Variation of scale-height parameters, H_1 and β.

ORCID iDs

Jonathan P. Williams https://orcid.org/0000-0001-5058-695X
Myriam Benisty https://orcid.org/0000-0002-7695-7605
Christian Ginski https://orcid.org/0000-0002-4438-1971
Giuseppe Lodato https://orcid.org/0000-0002-2357-7692
Maria Vincent https://orcid.org/0000-0001-5763-378X

References

Andrews, S. M. 2020, ARA&A, 58, 483
Ansdell, M., Gaidos, E., Hedges, C., et al. 2020, MNRAS, 492, 572
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipocz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Avenhaus, H., Quanz, S. P., Garufi, A., et al. 2018, ApJ, 863, 44
Bae, J., Isella, A., Zhu, Z., et al. 2023, in ASP Conf. Ser. 534, Protostars and Planets VII, ed. S. Inutsuka et al. (San Francisco, CA: ASP), 423
Bate, M. R., Lodato, G., & Pringle, J. E. 2010, MNRAS, 401, 1505
Batygin, K. 2012, Natur, 491, 418
Benisty, M., Dominik, C., Follette, K., et al. 2023, in ASP Conf. Ser. 534, Protostars and Planets VII, ed. S. Inutsuka et al. (San Francisco, CA: ASP), 605
Benisty, M., Juhász, A., Facchini, S., et al. 2018, A&A, 619, A171
Birnstiel, T., Klahr, H., & Ercolano, B. 2012, A&A, 539, A148
Bohn, A. J., Benisty, M., Perraut, K., et al. 2022, A&A, 658, A183
Carpenter, J. M., Esplin, T. L., Luhman, K. L., Mamajek, E. E., & Andrews, S. M. 2025, ApJ, 978, 117
Castro-Ginard, A., Penoyre, Z., Casey, A. R., et al. 2024, A&A, 688, A1
Cody, A. M., & Hillenbrand, L. A. 2018, AJ, 156, 71
Columba, G., Rigliaco, E., Gratton, R., et al. 2024, A&A, 681, A19
Cranmer, K., Brehmer, J., & Louppe, G. 2020, PNAS, 117, 30055

D’Alessio, P., Calvet, N., Hartmann, L., Franco-Hernández, R., & Servín, H. 2006, ApJ, 638, 314

Dorschner, J., Begemann, B., Henning, T., Jaeger, C., & Mutschke, H. 1995, A&A, 300, 503

Dullemond, C. P., Juhasz, A., Pohl, A., et al., 2012 RADMC-3D: A multi-purpose radiative transfer tool, Astrophysics Source Code Library, ascl:1202.015

Facchini, S., Juhász, A., & Lodato, G. 2018, MNRAS, 473, 4459

Facchini, S., Lodato, G., & Price, D. J. 2013, MNRAS, 433, 2142

Fitton, S., Tofflemire, B. M., & Kraus, A. L. 2022, RNAAS, 6, 18

Ford, E. B., & Rasio, F. A. 2008, ApJ, 686, 621

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Gaia Collaboration 2020, yCat, 1350

Garufi, A., Benisty, M., Pinilla, P., et al. 2018, A&A, 620, A94

Ginski, C., Ménard, F., Rab, C., et al. 2020, A&A, 642, A119

Gregorio-Hetem, J., Lepine, J. R. D., Quast, G. R., Torres, C. A. O., & de La Reza, R. 1992, AJ, 103, 549

Hunter, J. D. 2007, CSE, 9, 90

Jaeger, C., Mutschke, H., Begemann, B., Dorschner, J., & Henning, T. 1994, A&A, 292, 641

Kaeufer, T., Woitke, P., Min, M., Kamp, I., & Pinte, C. 2023, A&A, 672, A30

Longarini, C., Lodato, G., Rosotti, G., et al. 2025, ApJL, 984, L17

Marino, S., Perez, S., & Casassus, S. 2015, ApJL, 798, L44

Min, M., Stolker, T., Dominik, C., & Benisty, M. 2017, A&A, 604, L10

Muro-Arena, G. A., Benisty, M., Ginski, C., et al. 2020, A&A, 635, A121

Nealon, R., Dipierro, G., Alexander, R., Martin, R. G., & Nixon, C. 2018, MNRAS, 481, 20

Nixon, C., King, A., & Price, D. 2013, MNRAS, 434, 1946

Ohashi, N., Tobin, J. J., Jørgensen, J. K., et al. 2023, ApJ, 951, 8

Owen, J. E., & Lai, D. 2017, MNRAS, 469, 2834

Paszke, A., Gross, S., Massa, F., et al. 2019, arXiv:1912.01703

Pecaut, M. J., & Mamajek, E. E. 2016, MNRAS, 461, 794

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825

Pohl, A., Benisty, M., Pinilla, P., et al. 2017, ApJ, 850, 52

Price, D. J., Cuello, N., Pinte, C., et al. 2018, MNRAS, 477, 1270

Ribas, Á., Espaillat, C. C., Macìas, E., & Sarro, L. M. 2020, A&A, 642, A171

Ruzza, A., Lodato, G., & Rosotti, G. P. 2024, A&A, 685, A65

Sheehan, P. D., & Eisner, J. A. 2017, ApJ, 851, 45

Tazzari, M., Beaujean, F., & Testi, L. 2018, MNRAS, 476, 4527

Teague, R. 2019, JOSS, 4, 1220

Triaud, A. H. M. J. 2018, in The Rossiter-McLaughlin Effect in Exoplanet Research, ed. H. J. Deeg & J. A. Belmonte (Berlin: Springer), 2

van der Marel, N., Williams, J. P., Ansdell, M., et al. 2018, ApJ, 854, 177

van Holstein, R. G., Girard, J. H., de Boer, J., et al. 2020, A&A, 633, A64

Villenave, M., Benisty, M., Dent, W. R. F., et al. 2019, A&A, 624, A7

Villenave, M., Stapelfeldt, K. R., Duchêne, G., et al. 2024, ApJ, 961, 95

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261

Williams, J. P., & McPartland, C. 2016, ApJ, 830, 32

Wu, Y., & Murray, N. 2003, ApJ, 589, 605

Zack, G. W., Rogers, W. E., & Latt, S. A. 1977, JHC, 25, 741

Zhu, Z. 2019, MNRAS, 483, 4221