The Astrophysical Journal Supplement Series, 257:2 (19pp), 2021 November https://doi.org/10.3847/1538-4365/ac1430

Molecules with ALMA at Planet-forming Scales (MAPS). II. CLEAN Strategies for Synthesizing Images of Molecular Line Emission in Protoplanetary Disks

Ian Czekala[1,2,3,4,5,25], Ryan A. Loomis[6], Richard Teague[7], Alice S. Booth[8,9], Jane Huang[7,10,25], Gianni Cataldi[11,12], John D. Ilee[9], Charles J. Law[7], Catherine Walsh[9], Arthur D. Bosman[10], Viviana V. Guzmán[13], Romane Le Gal[7,14,15,16], Karin I. Öberg[7], Yoshihide Yamato[11], Yuri Aikawa[11], Sean M. Andrews[7], Jaehan Bae[17,18,25], Edwin A. Bergin[10], Jennifer B. Bergner[19,25], L. Ilsedore Cleeves[20], Nicolas T. Kurtovic[21,22], François Ménard[14], Hideko Nomura[12], Laura M. Pérez[21], Chunhua Qi[7], Kamber R. Schwarz[23,25], Takashi Tsukagoshi[12], Abygail R. Waggoner[20], David J. Wilner[7], and Ke Zhang[10,24,26]

[1] Department of Astronomy and Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA; iczekala@psu.edu
[2] Center for Exoplanets and Habitable Worlds, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[3] Center for Astrostatistics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[4] Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA
[5] Department of Astronomy, 501 Campbell Hall, University of California, Berkeley, CA 94720-3411, USA
[6] National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA 22903, USA
[7] Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA
[8] Leiden Observatory, Leiden University, 2300 RA Leiden, the Netherlands
[9] School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK
[10] Department of Astronomy, University of Michigan, 323 West Hall, 1085 S. University Ave., Ann Arbor, MI 48109, USA
[11] Department of Astronomy, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan
[12] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[13] Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
[14] Université Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France
[15] IRAP, Université de Toulouse, CNRS, CNES, UT3, 31400 Toulouse, France
[16] IRAM, 300 rue de la piscine, F-38406 Saint-Martin d'Hères, France
[17] Earth and Planets Laboratory, Carnegie Institution for Science, 5241 Broad Branch Rd. NW, Washington, DC 20015, USA
[18] Department of Astronomy, University of Florida, Gainesville, FL 32611, USA
[19] University of Chicago Department of the Geophysical Sciences, Chicago, IL 60637, USA
[20] Department of Chemistry, University of Virginia, Charlottesville, VA 22904, USA
[21] Departamento de Astronomía, Universidad de Chile, Camino El Observatorio 1515, Las Condes, Santiago, Chile
[22] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany
[23] Lunar and Planetary Laboratory, University of Arizona, 1629 E. University Blvd., Tucson, AZ 85721, USA
[24] Department of Astronomy, University of Wisconsin–Madison, 475 N. Charter St., Madison, WI 53706, USA
Received 2021 February 14; revised 2021 May 26; accepted 2021 May 29; published 2021 November 3

Abstract

The Molecules with ALMA at Planet-forming Scales Large Program (MAPS LP) surveyed the chemical structures of five protoplanetary disks across more than 40 different spectral lines at high angular resolution (0.″15 and 0.″30 beams for Bands 6 and 3, respectively) and sensitivity (spanning 0.3–1.3 mJy beam^{-1} and 0.4–1.9 mJy beam^{-1} for Bands 6 and 3, respectively). In this article, we describe the multistage workflow—built around the CASA `tclean` image deconvolution procedure—that we used to generate the core data product of the MAPS LP: the position–position–velocity image cubes for each spectral line. Owing to the expansive nature of the survey, we encountered a range of imaging challenges: some are familiar to the submillimeter protoplanetary disk community, like the need to use an accurate CLEAN mask, and others are less well known, like the incorrect default flux scaling of the CLEAN residual map first described by Jorsater & van Moorsel (the "JvM effect"). We distill lessons learned into recommended workflows for synthesizing image cubes of molecular emission. In particular, we describe how to produce image cubes with accurate fluxes via "JvM correction," a procedure that is generally applicable to any image synthesized via CLEAN deconvolution but is especially critical for low signal-to-noise ratio (S/N) emission. We further explain how we used visibility tapering to promote a common, fiducial beam size and contextualize the interpretation of S/N when detecting molecular emission from protoplanetary disks. This paper is part of the MAPS special issue of the Astrophysical Journal Supplement.

Unified Astronomy Thesaurus concepts: Protoplanetary disks (1300); Submillimeter astronomy (1647); Radio interferometry (1346); Deconvolution (1910)

1. Introduction

Submillimeter interferometers like the Atacama Large Millimeter/submillimeter Array (ALMA) enable high spatial and spectral resolution observations of protoplanetary disks. The Molecules with ALMA at Planet-forming Scales Large Program (MAPS LP) used ALMA to survey the chemical structures of five protoplanetary disks across more than 40 different spectral lines; an overview of the program and references to the full suite of MAPS papers are provided in Öberg et al. (2021). In this paper, we describe the imaging strategies we employed to synthesize position–position–velocity image cubes from the interferometric visibilities.

[25] NASA Hubble Fellowship Program Sagan Fellow.
[26] NASA Hubble Fellow.

These image cubes form a core data product of the MAPS LP from which other value-added data products like moment maps, radial intensity profiles, and emission surfaces are derived (Law et al. 2021a, 2021b).

Consider an astronomical source whose spatial sky brightness distribution is described by I. Over small angular extents, the distribution is indexed $I(l, m)$ by the direction cosines $l = \sin(\Delta\alpha \cos\delta)$ and $m = \sin(\Delta\delta)$ relative to some phase center, where α is R.A. and δ is decl. and l increases to the east and m increases to the north. The visibility function of the astronomical source is the Fourier transform of the sky brightness distribution, given by

$$\mathcal{V}(u, v) = \iint I(l, m) \exp\{-2\pi i(ul + vm)\}\, dl\, dm, \quad (1)$$

and is a complex quantity having real and imaginary components with units of flux, e.g., janskys (for a full discussion of the conditions that must be met in order for Equation (1) to be accurate, see Thompson et al. 2017, Chapter 3). Interferometers sample $\mathcal{V}$ at a discrete set of spatial frequencies fundamentally dictated by the array configuration and observing frequency (Thompson et al. 2017). The spatial frequencies (u, v) are measured in multiples[27] of the observing wavelength λ or kλ, which can also be converted to length (e.g., meters), directly corresponding to the instantaneous, projected baselines of the array.[28] For radio interferometers like ALMA, the fundamental data product is then—for every spectral channel—a set of calibrated visibility measurements at various (u_k, v_k) coordinates. These visibilities are complex-valued numbers measured in the presence of Gaussian noise:

$$V_{\mathrm{data},k} = \mathcal{V}(u_k, v_k) + \varepsilon_k. \quad (2)$$

The distribution from which the noise ε_k is drawn is usually well described by a Gaussian distribution with thermal weights[29] equal to the inverse variance $w_k = \sigma_k^{-2}$. Observational data is often interpreted in the context of a model. Fortunately, this measurement process defines a straightforward likelihood function for any set of model visibilities $\boldsymbol{V}_{\mathrm{model}}$,

$$\ln \mathcal{L} = \ln p(\boldsymbol{V}_{\mathrm{data}}|\boldsymbol{V}_{\mathrm{model}}) \propto -\frac{\chi^2}{2}, \quad (3)$$

where

$$\chi^2 = \sum_k^N w_k\, |V_{\mathrm{data},k} - V_{\mathrm{model},k}|^2 \quad (4)$$

and N is the number of visibility measurements. Throughout this work, we use bold notation to signify vector quantities. Model visibilities can be generated from analytic forward models $I(l, m) \rightleftharpoons \mathcal{V}(u, v)$ (e.g., an axisymmetric intensity profile describing thermal emission from dust rings in a protoplanetary disk; Zhang et al. 2016; Guzmán et al. 2018; Jennings et al. 2020) or by fast-Fourier-transforming complete radiative transfer models of $I(l, m)$ (e.g., complicated dust morphologies (Tazzari et al. 2018) or kinematic models of spatially resolved CO emission (Czekala et al. 2019)). Regardless of how model visibilities are generated, model fitting with the visibility likelihood function (Equation 3) and judicious choices of prior probability distributions creates a well-motivated posterior probability distribution that can be used for Bayesian parameter inference (e.g., Hogg & Foreman-Mackey 2018; Speagle 2019).

Most of the difficulty in synthesizing images representative of the sky brightness distribution $I(l, m)$ stems from the fact that the visibility function is not adequately sampled at all of the spatial frequencies where it has significant power. Because the Fourier transform is a linear operator, one maximum likelihood image solution is simply the inverse Fourier transform of the visibility measurements. In this inversion process, the unsampled spatial frequencies are typically set to zero power. The image that results is called the "dirty image."[30] By definition, the point-spread function (PSF) for the dirty image is the system response to an impulse sky brightness distribution ($I(l, m) = \delta(0, 0)$, where in this context δ represents the Dirac delta function). The PSF is also called the "dirty beam" because it typically has a substantial sidelobe pattern, responsible for the low fidelity of the dirty image. The PSF width and sidelobe pattern amplitude can be altered by adjusting the scheme by which nearby uv points are averaged or "gridded" (with trade-offs against the thermal noise level; Briggs 1995).

The PSF sidelobe response can be (effectively, albeit not perfectly) removed from the dirty image through image deconvolution. The most widely used deconvolution algorithm in the radio astronomy community is CLEAN (Högbom 1974), which we describe in Section 3. For more background on the CLEAN family of algorithms, see Chapter 11 of Thompson et al. (2017).

An alternative family of imaging algorithms are the "maximum entropy" techniques (Cornwell & Evans 1985; Narayan & Nityananda 1986; Cárcamo et al. 2018) and, more generally, the regularized maximum likelihood (RML) methods (Event Horizon Telescope Collaboration et al. 2019; Nakazato et al. 2019). Rather than focusing on image deconvolution, these algorithms forward-model visibility measurements using flexible, nonparametric models of the sky brightness distribution conditioned by well-motivated image priors. We will present the results of an RML technique (as implemented in the `MPoL` package; Czekala & Loomis 2020) applied to the MAPS LP in a forthcoming MAPS paper (I. Czekala et al. 2021, in preparation).

This article is arranged as follows. In Section 2 we describe in detail how the configuration of antennas within the interferometric array dictates the characteristics of the synthesized image PSF. We describe how we use the CLEAN algorithm to deconvolve the PSF sidelobe response from synthesized images in Section 3. In Section 4 we examine the implications of the so-called Jorsater & van Moorsel (1995; JvM) effect, an important flux scaling issue critical to correctly

[27] Though l and m are technically unitless, for small angular extent, they could also be considered to have units of radians, implying that u and v can also be interpreted in units of cycles per radian.

[28] Following the convention for l and m, u increases to the east and v increases to the north. Since u and v correspond to the baseline locations of the array, we plot u increasing to the right, as if the array were viewed on the surface of the Earth from above (e.g., Thompson et al. 2017, Figure 3.2).

[29] Often these weights are presented proportionally $w_k \propto \sigma_k^{-2}$ such that the constant of proportionality is determined via the rms noise on a single correlator channel corresponding to a visibility of unit weight (Section 3.2 of Briggs 1995). In our definition, we assume that this calibration has already been carried out and that the weights have a direct statistical interpretation.

[30] Both because of its typically lousy aesthetics and because it forms the starting point for the CLEAN family of deconvolution algorithms.

interpreting faint line emission. We discuss additional strategies of CLEAN masking and *uv* tapering in Sections 5 and 6, respectively. In Section 7 we discuss how the signal-to-noise ratio (S/N) might be interpreted in the image products, with reference to the nonimaging matched-filter approach used in Cataldi et al. (2021) and Ilee et al. (2021). We conclude in Section 8 and describe available data products from the MAPS LP in Section 9.

2. Synthesized Beams and the Impact of *uv* Coverage

The projected length and orientation of each baseline in the interferometric array directly correspond to the spatial frequency (*uv* point) of the visibility function that it measures (Chapter 2 of Thompson et al. 2017). A longer baseline will sample a higher spatial frequency, which corresponds to finer angular scales in the image plane. As the Earth rotates over the course of an observation, the projected length and orientation of a baseline relative to the astronomical source will change and the visibility function will be sampled over a range of *uv* values (Chapter 5 of Thompson et al. 2017).

To maximize image sensitivity and resolution, multiple sets of visibilities from short- and long-baseline array observations are often concatenated together into a single measurement set and imaged. The MAPS program utilized two separate configurations—nominally equivalent to the C43-4 ("short") and C43-7 ("long") ALMA configurations—with baselines ranging from 15 to 1400 m and from 40 to 3600 m, respectively. The full details of the array configurations for each observation are presented in Tables 9 and 10 of Öberg et al. (2021).

As introduced in Section 1, the dirty image is generated by taking the inverse Fourier transform of the visibility measurements while the unsampled spatial frequencies are assumed to carry zero power. Following Briggs (1995), the dirty image is formalized as

$$I_{\rm data}(l, m) = C\sum_{k=1}^{N^\dagger} T_k D_k w_k V_{{\rm data},k} \exp\{2\pi i(u_k l + v_k m)\} \quad (5)$$

where T_k is an optional visibility taper and D_k is a *uv* density weight. In this instance and in what follows, we assume that the set of visibilities have been augmented to include their complex conjugates, $N^\dagger = 2N$. When making images, it is necessary to sum over the Hermitian conjugates of the visibilities to ensure that the sky image is real. The normalization constant is

$$C = 1 \Big/ \sum_{k=1}^{N^\dagger} T_k D_k w_k. \quad (6)$$

For the immediate discussion of this section, we assume that $T_k = 1\,\forall k$ and the D_k values correspond to the default density weighting of non-tapered MAPS images, which is `robust = 0.5` (Briggs 1995). We will return to a discussion of both tapering and density weighting in more detail in Section 6. As discussed in Briggs (1995), the units of the dirty image are such that a point source with flux density S will have a peak numerical value of S in the dirty image—for discussion purposes one can reference the flux units of janskys per dirty beam.

The PSF or dirty beam of the synthesized image is calculated with Equation 5 by setting $V_{{\rm data},k} = 1\,{\rm Jy}\,\forall k$, which is the Fourier transform of an impulse sky brightness distribution, i.e., a 1 Jy point source located at the phase center. The PSF is then

$$B_{\rm PSF}(l, m) = C\sum_{k=1}^{N^\dagger} T_k D_k w_k \exp\{2\pi i(u_k l + v_k m)\}. \quad (7)$$

Phrased differently, if one considers the interferometric array transfer function W under a choice of density weighting and tapering parameters

$$W(u, v) = \sum_{k=1}^{N^\dagger} T_k D_k w_k \delta(u - u_k, v - v_k), \quad (8)$$

the PSF is its Fourier dual ($B_{\rm PSF} \rightleftharpoons W$). The units of janskys per dirty beam are technically undefined, since it is not guaranteed that the PSF integrates to a finite volume; however, the maximum is always $B_{\rm PSF}(0, 0) = 1$.

In the first row of Figure 1, we show the *uv*-plane sampling for a representative observation of a disk in the MAPS sample: MWC 480 HCN $J = 3 - 2$ in spectral setting B6-2 (for a full description of the MAPS spectral setup, see Tables 2–4 of Öberg et al. 2021). These samples are split into short and long baselines in the left and middle columns, respectively, and they are combined in the right column. For a sense of scale, the combined panel contains 408,697 individual visibility samples per spectral channel, which is typical of Band 6 MAPS observations.

The PSFs corresponding to the short-, long-, and combined-baseline subselections are shown in the second row of Figure 1. Each disk in the MAPS LP was also observed by ALMA at a different elevation, leading to a different set of projected baselines and thus to different PSF responses. For the same disk, all molecular transitions observed at the same time as part of a single spectral setup have the same baseline configuration, but different spectral setups were observed at different stages of array (re)configuration and therefore have slightly different baseline distributions and PSFs.

Given the way that ALMA array configurations were originally designed to sample the *uv*-plane,[31] individual ALMA configurations retain approximately Gaussian beams, with more extended configurations yielding narrower beams and better spatial resolution images (see Figure 1, Columns 1 and 2). Enhanced resolution comes with trade-offs, however.

The maximum recoverable scale (MRS) is a measure of the largest angular scale that can be usefully imaged from a set of visibility measurements. By definition, the more extended configurations of ALMA have fewer short baselines and thus are less sensitive to emission on large spatial scales. If a source has emission on spatial scales larger than the MRS of an array configuration, image flux carried at these low spatial frequencies will not be recovered in the synthesized image (for an analysis of the missing flux using analytic sky brightness distributions, see Appendix A of Wilner & Welch 1994).

The morphology of molecular line emission in protoplanetary disks is a function of velocity (frequency), with emission near the systemic velocity of the source (the transition rest-frame frequency) usually being the most spatially extended. The MRS[32] of the nominal C43-7 configuration is only $\approx 1\farcs1$,

[31] http://library.nrao.edu/public/memos/alma/main/memo400.pdf and http://library.nrao.edu/public/memos/alma/main/memo598.pdf

[32] Using the 5th percentile definition in the ALMA technical handbook (Remijan et al. 2019). See Section 2 of Huang et al. (2021) for a discussion of how the MRS of the combined MAPS configurations affects the interpretation of large-scale CO emission in GM Aur.

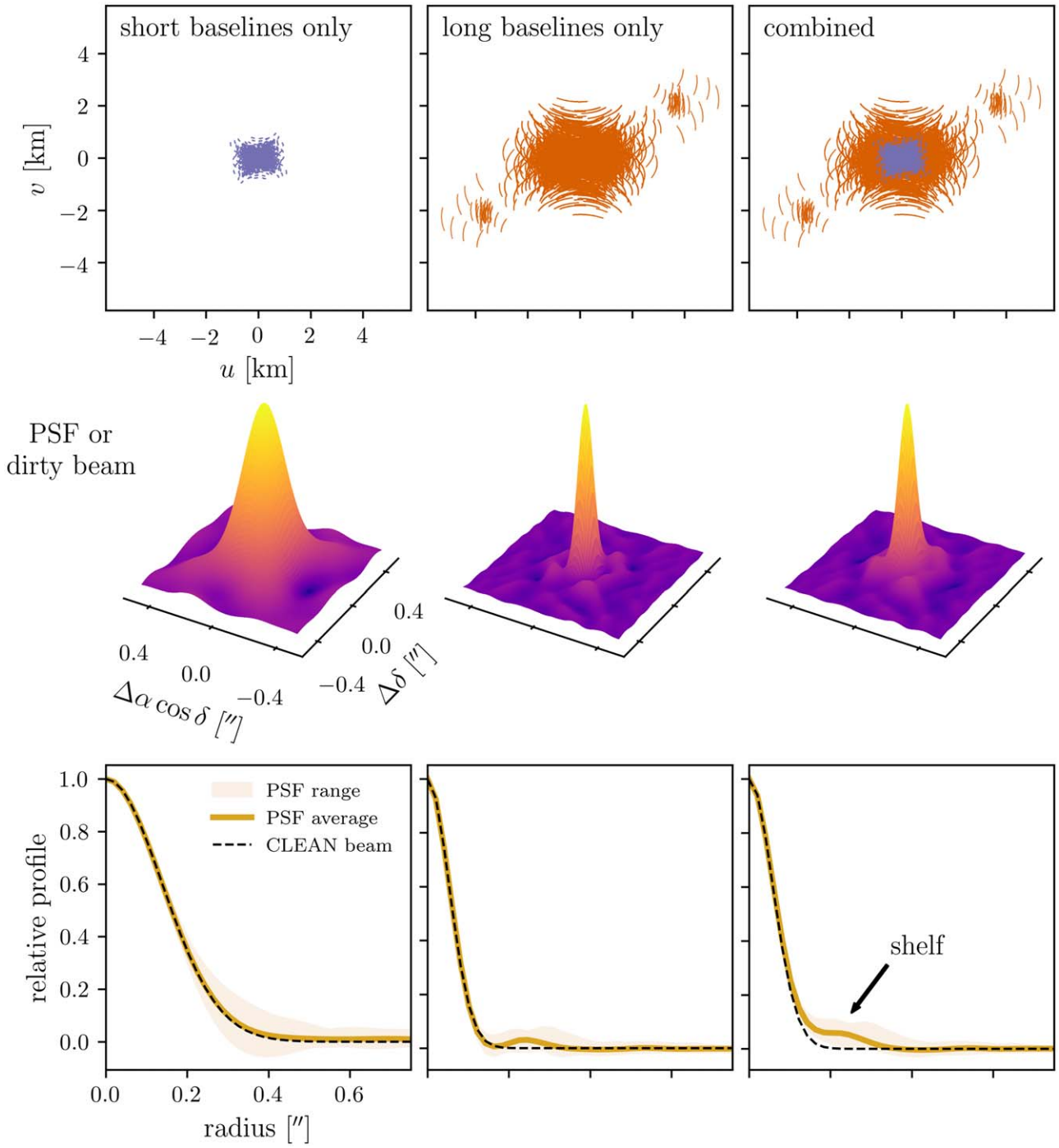

Figure 1. How *uv*-plane coverage impacts the synthesized PSF or dirty beam. Top row: The two sets of baselines (i.e., short and long baselines from the nominal C43-4 and C43-7 configurations, respectively) resulting from a representative observation of a representative disk in the MAPS LP. Middle row: The PSF that results from the default Briggs weighting of MAPS (`robust` = 0.5). The CLEAN (or restoring) beam is an elliptical Gaussian function fit to the main lobe of the PSF. Bottom row: The deprojected and azimuthally averaged radial profiles of the PSF and CLEAN beams, with the minimum-to-maximum range of the azimuthal variation of the PSF indicated by the shaded region. Considered individually, the ALMA nominal configurations yield dirty beams reasonably approximated (on average) by elliptical Gaussian CLEAN beams. However, the joint-baseline configuration results in a PSF with a substantial "shelf" to the main Gaussian core.

while all the disks targeted in the MAPS LP (IM Lup, GM Aur, AS 209, HD 163296, and MWC 480) were known to have emission on spatial scales larger than this (see Section 2.1 of Öberg et al. 2021 and references therein). This makes it necessary to observe with a combination of array configurations to properly sample the visibility function of each MAPS target, especially at the velocities (frequencies) where the emission is most extended.

The combination of observations from short- and long-baseline configurations—or at least those utilized by the MAPS LP—yields a dirty beam that has a substantial shelf at larger radii (Figure 1, right column). To highlight this shelf, we compare each dirty beam to its corresponding CLEAN beam in the third row of Figure 1.

The CLEAN beam is an elliptical Gaussian fit to the main lobe of the dirty beam. The beam is reported using the FWHM

along the major and minor axes (θ_a and θ_b, respectively) and the position angle of the major axis (ϕ; degrees east of north). The beam power pattern is then

$$B_{\rm CLEAN}(l, m) = \exp\left(-\frac{1}{2}\left[\left(\frac{l'}{\sigma_{l'}}\right)^2 + \left(\frac{m'}{\sigma_{m'}}\right)^2\right]\right) \quad (9)$$

where

$$l' = l\cos\phi - m\sin\phi \quad (10)$$

$$m' = l\sin\phi + m\cos\phi \quad (11)$$

and

$$\sigma_{l'} = \sigma_l = \theta_b/(2\sqrt{2\ln 2}) \quad (12)$$

$$\sigma_{m'} = \sigma_m = \theta_a/(2\sqrt{2\ln 2}). \quad (13)$$

To convert from units of janskys per CLEAN beam to janskys per square arcsecond, for example, one needs to divide by the effective solid angle (angular area) of the CLEAN beam

$$\Omega_{\rm CLEAN} = \iint B_{\rm CLEAN}(l, m)\, dl\, dm = \frac{\pi\theta_a\theta_b}{4\ln 2}. \quad (14)$$

This effective solid angle can be calculated by considering the beam response to a spatially uniform source (e.g., the cosmic microwave background). Alternatively, the size and shape of the PSF can be characterized by considering the beam as a 3D solid with its peak normalized to 1. The effective area is then the "volume" of this solid in units of 1 $\times$ arcsec^2, for example, and is graphically illustrated for the dirty beams in the middle row of Figure 1. The CLEAN beam sizes for all MAPS products are listed in Öberg et al. (2021, Table 5). To form the 1D beam profiles in the bottom row of Figure 1, we "deproject" both the dirty and CLEAN beams by the aspect ratio of the CLEAN beam and azimuthally average them. In reality, the dirty beams are not azimuthally symmetric, so this is only an approximation for the purposes of visualization (the full range of azimuthal variation is conveyed by the shaded region).

The dirty beam exhibits non-Gaussianity even for the shortest-baseline configurations (e.g., a slightly elevated "tail"). For the combined configurations, non-Gaussianity manifests as a shelf outside an approximately Gaussian core. Though the shelf may appear small, it is at the root of several issues that ramify throughout the image deconvolution process. We will now describe this process and how we mitigate these issues.

3. The Cleaning Process

We synthesized and deconvolved image cubes using the `tclean` task in the Common Astronomy Software Applications (CASA) package, version 6.1.0 (McMullin et al. 2007). The salient components of this process are illustrated in Figure 2 using a single channel containing significant but faint emission from an astrophysical source (AS 209 DCN $J = 3 - 2$). The first algorithmic decision is which functional basis set to use for CLEAN components. The simplest version of CLEAN uses Dirac δ-functions (Högbom 1974). We used the more advanced "multiscale" algorithm (`deconvolver=''multiscale''`), which is built on a set of variously sized axisymmetric tapered parabolic components similar to 2D Gaussian functions (Cornwell 2008). CLEAN components may also take on (small) negative amplitudes so that components placed in later iterations can refine a CLEAN model built from larger positive components placed in earlier iterations. We set the component scales to `scales=[0, 5, 15, 25]` pixels, where the pixel size was chosen to correspond to approximately one-seventh of the beam size. This pixel size adequately oversamples the beam FWHM without making the image size impractically large: the smallest image dimensions were for Band 3 or 0.″3 tapered images (1024 $\times$ 1024 pixels) and the largest image dimensions were for Band 6 CO (2048 $\times$ 2048 pixels).

The CLEAN algorithm starts by setting a "residual image" equal to the dirty image and a "CLEAN model" equal to a blank image. Each iteration of the CLEAN algorithm introduces new CLEAN components and two things happen. First, the CLEAN model is gradually built up when CLEAN components are placed at locations corresponding to the current peak flux in the residual map. It is possible to use a binary mask to restrict the placement of new CLEAN components. Second, the newly placed CLEAN components are deconvolved from the residual map. The deconvolution is carried out by subtracting the convolution of the CLEAN component with the dirty beam. The algorithm continues iterating until an exit criterion is triggered. The criterion may be as simple as a maximum number of iterations or it may correspond to a threshold on the noise properties of the residual map. We have chosen the latter and we discuss our choice of masks and thresholds in Section 5.

When the deconvolution procedure has finished, the CLEAN model is convolved with the CLEAN beam to form a convolved CLEAN model image. The final residual image is added to the CLEAN model to form the restored (or "CLEANed") image. For multichannel measurement sets like those of the MAPS LP, CLEAN images and deconvolves each channel with completely independent CLEAN components.

Though CLEAN is best known for deconvolving dirty beam sidelobes, an arguably more important function of the CLEANing process is creating an interpretable flux model. Because the flux units of the dirty image are technically undefined (janskys per dirty beam), any measurements made using a dirty image are highly sensitive to the *uv*-sampling distribution. For an extreme example, consider the ill-advised task of measuring the flux of a point source using the dirty image from a two-element interferometer, better known as a fringe pattern (e.g., Chapter 9 of Wilson et al. 2013).[33] Depending on how one drew the photometric aperture, one could just as easily include or exclude various nulls and maxima of the fringe pattern and measure fluxes ranging from a positive peak to a negative trough of the sine wave. However, if one were to first deconvolve the fringe PSF from the dirty map, reasonable inferences could be drawn on the source flux (if not on the location, in this contrived example).

Because the CLEAN model is built up with CLEAN components, it has real, physical interpretability: the units of the CLEAN model are janskys per pixel and the units of the convolved CLEAN model are janskys per CLEAN beam. Because the CLEAN beam has finite volume (Equation 14), CLEANed images are on much surer flux footing than dirty images, whose default dirty beam is not guaranteed to have finite volume.

CLEAN components like δ-functions or tapered parabolic components are rarely a perfect basis to represent a spatially

[33] A task more easily carried out by measuring the amplitude of the fringe pattern or (equivalently) using a model-fitting approach (Equation 3).

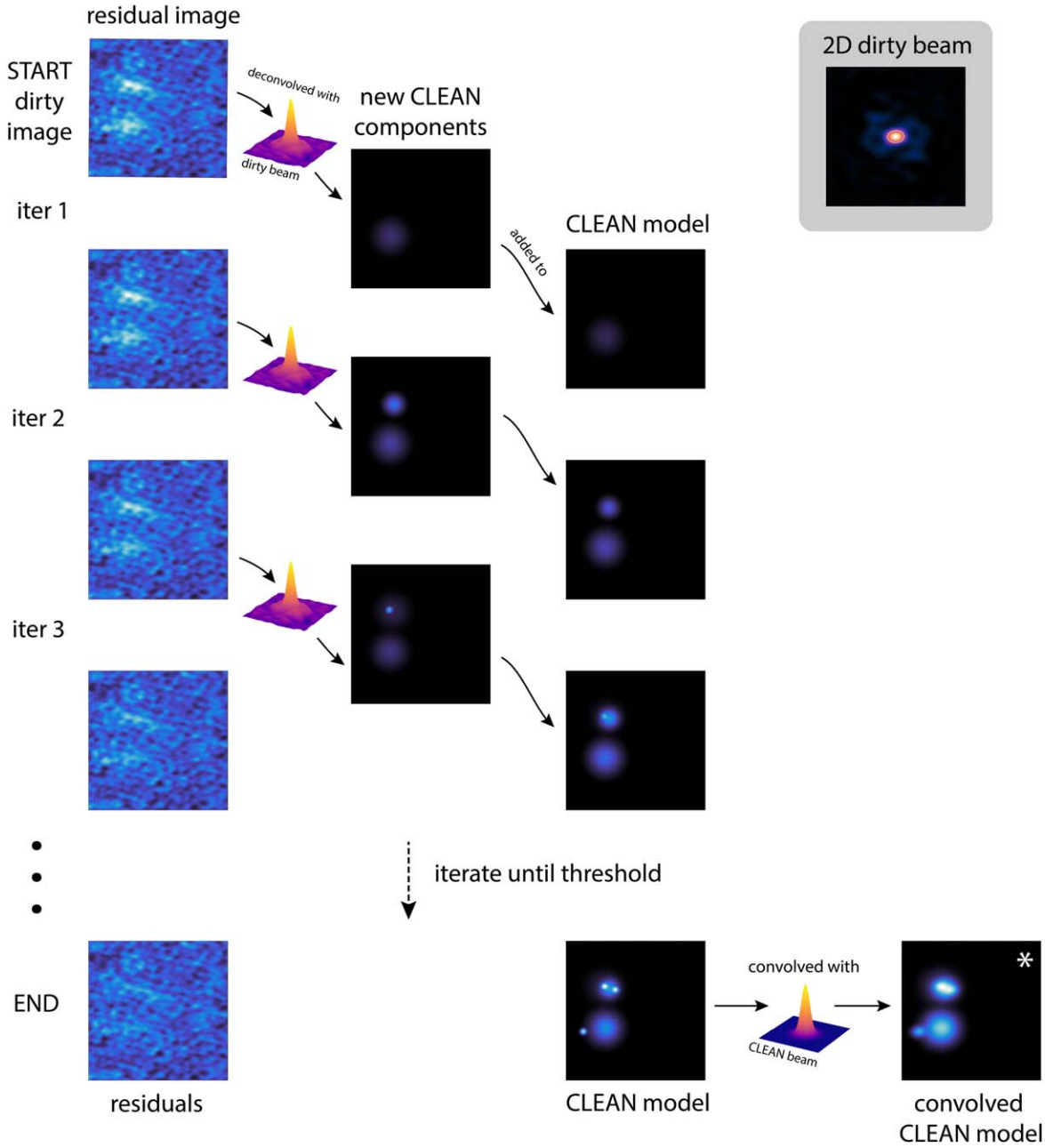

Figure 2. A schematic of the CLEANing procedure tracking the residual image as new CLEAN components are deconvolved and added to the CLEAN model. The residual image is initialized with the dirty image (which has units of janskys per dirty beam). With each iteration, the residual image is deconvolved by subtracting the convolution of the dirty beam with the new CLEAN component(s), removing the effects of beam sidelobes (a 2D representation of the dirty beam is shown in the upper right gray panel). At the end of the process, the CLEAN model is smoothed by convolution with the CLEAN beam (the FWHM of the CLEAN beam is marked by a white ellipse in the 2D dirty beam panel); the resulting image has units of janskys per CLEAN beam. The CLEAN models are shown with an aggressive color stretch (square root stretch) to better demonstrate the accumulation of CLEAN components. (*) The same convolved CLEAN model is visualized with a softer linear color stretch in Figure 3.

resolved source. However, in the limit of many low-amplitude components, reasonable models of source structure can be achieved. If the CLEAN components were a good match to source morphology, e.g., δ-functions for a field of quasars, then the CLEAN model itself would be a reasonable product to use for analysis. Like the CLEAN model in Figure 2, this is rarely achieved in practice; it is more often the case that the higher-resolution information that could be conveyed by an accurate, native-resolution CLEAN model is gladly traded for a more visually pleasing convolved CLEAN model, where errors stemming from a highly discretized but imperfect CLEAN basis set have been low-pass filtered out by CLEAN beam convolution (for a visual explanation, see the discussion of optimal resolution in Chael et al. 2016, Section 5.3). This shortcoming is one reason why flexible, nonparametric imaging techniques can often produce higher-resolution image products than CLEAN (see Event Horizon Telescope Collaboration et al. 2019 for a general discussion, and see Pérez et al. 2019, Disk Dynamics Collaboration et al. 2020, and Jennings et al. 2020 for discussions specific to protoplanetary disks).

While the CLEAN procedure may be familiar to many radio astronomers, it is still a nonlinear process subject to myriad

algorithmic choices like loop gain, threshold stopping criteria, and masking regions, among other parameters embodied in the `tclean` argument list.[34] When the basis sets are mismatched from the source morphology (e.g., when multiscale Gaussian components are used to deconvolve a source that is actually composed of concentric rings), it is very important to correctly tune these algorithm parameters to obtain faithfully restored images. It is not common practice within the protoplanetary disk community to publish representations of the CLEAN model in scientific analysis, but we argue that inspection of the CLEAN model (and its potential deficiencies) should be part of any radio astronomy workflow, especially where fine-featured source morphologies are concerned.

4. JvM Effect and Correction

4.1. The Problem: The CLEAN Model and Residual Image Have Different Units

Now that we have outlined the broad contours of the CLEANing process, we revisit the final step of the CLEAN algorithm, when the final CLEANed image is formed by summing the residual image and the convolved CLEAN model, shown graphically in the top panel of Figure 3 as the "standard" workflow. There are usually two reasons why radio astronomers carry out this final step, even though we have just discussed the reasons why the CLEAN model is a reasonable scientific product on its own terms. One, this gives the final CLEANed image some representation of the thermal noise, which is useful for interpreting the significance of features. Two, the residual map still contains some (ideally a small level of) real astrophysical flux that was not adequately deconvolved (e.g., see the "residuals" panel in the top row of Figure 3). Adding the residuals back to the CLEAN model provides some insurance that this real flux at least appears in the final image, though it will still exhibit the effects of the sidelobe response.

Because the residual map originated as the dirty image, it technically has units of janskys per dirty beam, while the convolved CLEAN model has units of janskys per CLEAN beam. This means that the final CLEANed image is created with mismatched units. If the CLEAN beam accurately approximates the dirty beam, the unit mismatch is inconsequential. However, if the dirty beam has even a small shelf, such as the ∼10% amplitude shelf on the `robust=0.5` beam shown in Figure 1 and replicated in Figure 4, there can be severe implications for accurate flux recovery. Because the shelf occurs at a large beam radius, it contributes to a large mismatch in differential volume, even though it is small in relative profile. This small shelf means that a CLEAN beam fit to this dirty beam main lobe will encompass only 60% of the full dirty beam volume (using the first null as a proxy for dirty beam extent, even though the response extends much further). Therefore, the units janskys per CLEAN beam and janskys per dirty beam differ substantially. The shelf of naturally weighted images is usually even more severe than that of robust or uniformly weighted images.[35]

To our knowledge, this issue and its implications for flux conservation were first described in Jorsater & van Moorsel (1995, Appendix A), and so we term this the JvM effect throughout the MAPS paper series. We calculate the ratio of the beam volumes as

$$\epsilon = \left. \frac{V_{\rm CLEAN}(r)}{V_{\rm dirty}(r)} \right|_{\rm first\ null}, \tag{15}$$

using the CLEAN beam and the `.psf` dirty beam files produced by `tclean`. The procedure originally described in Jorsater & van Moorsel (1995) calculates ϵ using the ratio of the CLEAN model to the difference between the dirty image and the residual map.[36] Our calculation using the ratio of the beam volumes yields a similar result via direct calculation. The bottom panel of Figure 4 shows a 1D representation of this calculation, though in practice we use the 2D beam profile since the dirty beam is not axisymmetric.

One might think that the unit mismatch can be avoided if only one fully CLEANed one's images, i.e., if the CLEAN model contained all of the real astrophysical flux and the residual map contained only noise. We concur with Jorsater & van Moorsel (1995) that this is unattainable in any real-world application of CLEAN. Consider an example where the CLEAN threshold is set at $2\times$ rms. At the end of the CLEANing process the residual map will contain some real astrophysical flux below this threshold while the CLEAN model will contain some components that were erroneously deconvolved from noise spikes above this threshold. Varying the threshold just changes the balance of flux in the residual map or CLEAN model—it is impossible to CLEAN to a zero-flux threshold without also adding a significant number of erroneous components to the CLEAN model. The JvM effect makes imaging faint molecular line emission challenging because (by definition) a significant fraction of the flux in each channel will exist at or below a typical CLEANing threshold level (e.g., a few $\times$ rms). When particularly faint data sets, such as the one chosen to illustrate Figures 2 and 3 (AS 209 DCN $J = 3-2$), are CLEANed to anything but the deepest thresholds (generally inadvisable for other reasons; see Section 5), the residuals may contain 50% or more of the total flux in the image. If one were to erroneously assume the products in the top row of Figure 3 were all in units of janskys per CLEAN beam (as is normal practice), the residuals, CLEAN model, and final image would contain 70 mJy, 33 mJy, and 103 mJy of flux, respectively.

The JvM effect is often less pronounced when synthesizing images of high-S/N continuum emission because in that application the CLEAN model will hold a higher proportion of astrophysical flux. However, it may still hamper the characterization of fainter regions within those images.

4.2. The Solution: Rescaling the Residual Image Using JvM Correction

While fully CLEANing images is not a practical solution to the unit mismatch issue, we can approximately convert the

[34] See the CASA `tclean` documentation for a full description (https://casa.nrao.edu/casadocs-devel/stable/global-task-list/task_tclean/about) and the NRAO pipeline scripts (https://github.com/AstroChem/MAPS/tree/master/NRAO_processing) for all parameter settings. Any nondefault parameter choices we made are discussed in this article.

[35] Because visibility-fitting techniques and forward-model imaging approaches eschew image plane deconvolution entirely, they have the potential to accurately recover flux even when the dirty beam has a nontrivial shelf.

[36] The ratio of the beam volumes need not always be $\epsilon < 1$; depending on the choice of CLEAN beam, which can technically be set to any arbitrary finite-volume beam, values $\epsilon > 1$ are possible. In practice, though, we fit elliptical Gaussian CLEAN beams to the main lobe of the dirty beam and thus $\epsilon \lesssim 1$ across MAPS image products.

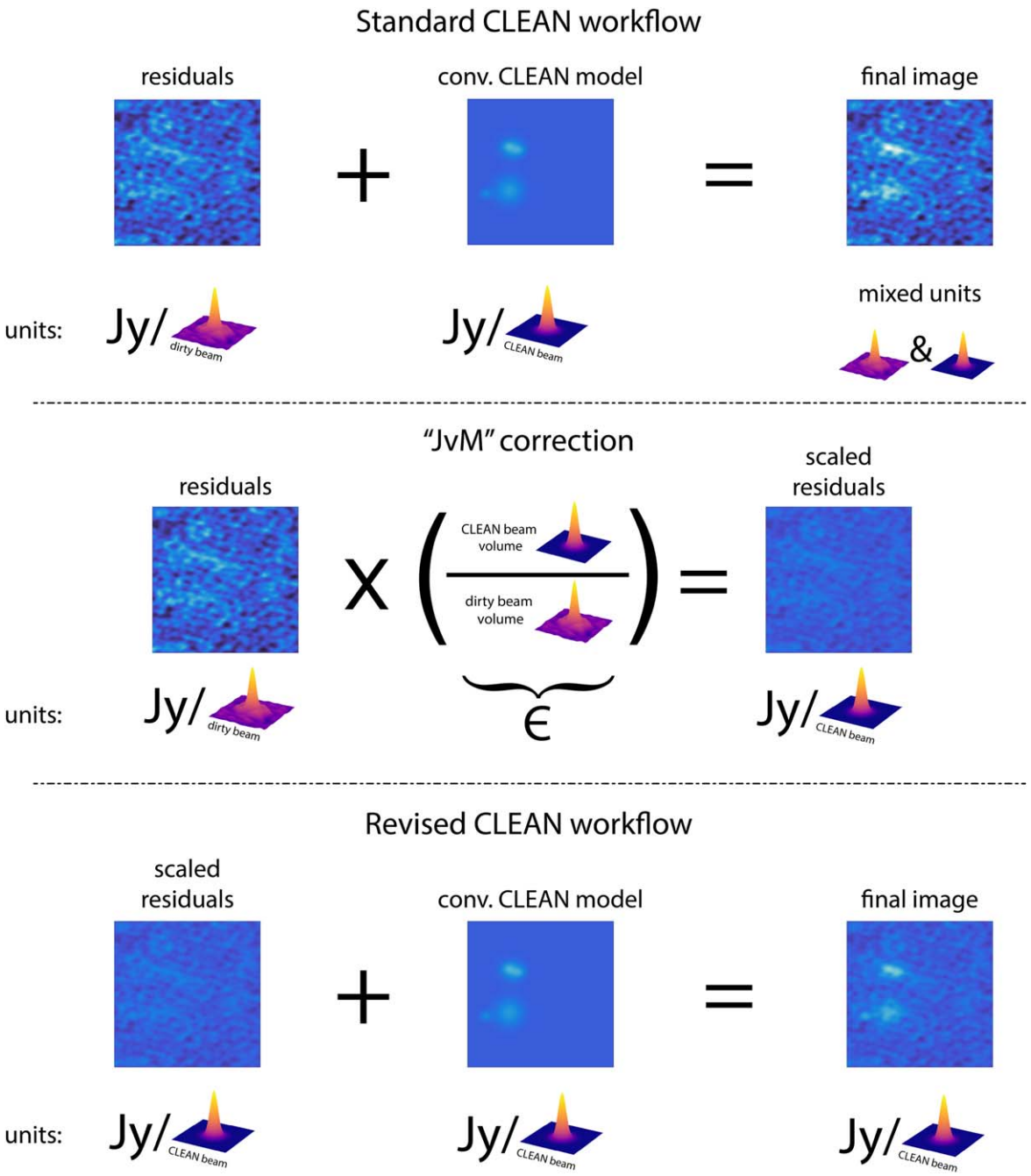

Figure 3. After the deconvolution iterations are complete, the residual map is added to the convolved CLEAN model to form the final image. Because the residual map is derived from the dirty image, it has units of janskys per dirty beam. In the standard CLEAN workflow (top panel), this creates a final CLEANed image with mismatched units, and therefore compromised interpretability. Following Jorsater & van Moorsel (1995), one solution is to scale the residual map by the ratio of the CLEAN beam volume to the dirty beam volume, a process that we term "JvM correction." In the bottom panel we show our revised CLEAN workflow that includes the JvM-corrected residual, which results in a final CLEANed image with the correct intensity units. All images in this figure appear with the same color scale stretch and limits.

residual map from units of janskys per dirty beam to units of janskys per CLEAN beam. This is achieved by rescaling the residual map by a factor of ϵ before it is added to the CLEAN model (see Figure 3, middle panel; for this data set, $\epsilon = 0.359$). We call this JvM correction, which we implement using the `immath` CASA task and the `.model` and `.residual` outputs from `tclean`. Under the revised workflow, the scaled residuals, CLEAN model, and final image in the bottom row of Figure 3 contain 25 mJy, 33 mJy, and 58 mJy of flux, respectively (resulting in a dramatic 50% change in total flux as compared to the standard workflow). As demonstrated by this faint data set, failure to apply JvM correction can have a profound effect on both the total flux contained within an image and the morphological characteristics of that flux (see Figure 3, final images).

The ϵ values for all MAPS image cubes are listed in Table 11 of Öberg et al. (2021). Band 3 image cubes typically have ϵ values in the range 0.7–1.0, while Band 6 image cubes typically have ϵ values in the range 0.2–0.7. The effect of JvM correction becomes more significant the more ϵ deviates from 1.

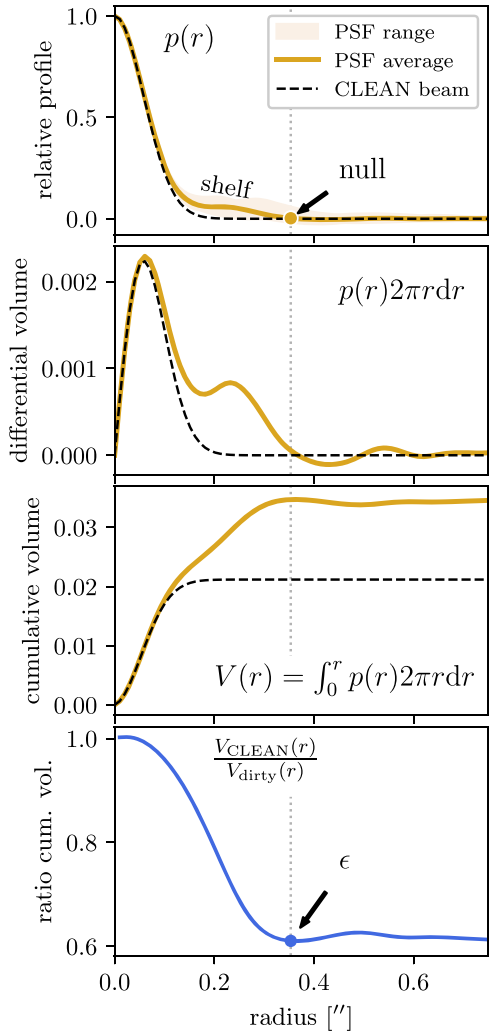

Figure 4. How a dirty beam shelf leads to a divergence of integrated beam volume. Top panel: The PSF profiles $p(r)$ from the joint-baseline configuration in Figure 1, with the first null (zero-crossing) of the dirty beam labeled as a proxy for the total "size" of the dirty beam. Second and third panels: Though the shelf of the dirty beam is small relative to the PSF peak, because it is at a large radius, it leads to a significant divergence in the total integrated volume of the dirty beam relative to the CLEAN beam. Last panel: The ratio of the CLEAN beam volume to the dirty beam volume, ϵ, can be used to convert the residuals from the CLEANing process into proper units of janskys per CLEAN beam.

JvM still matters for signal-free channels! Since JvM correction modifies the residual map, this creates ambiguity around an image plane rms measurement. Technically, even signal-free channels with zero CLEAN components need to be corrected for the JvM effect since they are formed from a residual map in units of janskys per dirty beam but are reported in units of janskys per CLEAN beam. Unless otherwise stated, throughout the MAPS paper series we specify rms values in janskys per CLEAN beam measured from emission-free images corrected for the JvM effect (but not yet corrected for the primary beam's sensitivity).[37]

A notable exception is when we specify CLEANing thresholds. Since the `tclean` procedure references threshold values in the dirty map itself, we report these thresholds with respect to rms values measured in the non-JvM-corrected dirty map.

[37] Thus, the rms values are equivalent to an rms value measured at the phase center of an image corrected for the primary beam's sensitivity.

5. Masking to Improve the CLEAN Model

Binary image plane masks are useful for guiding the placement of CLEAN components in locations believed to correspond to physical flux. Simple elliptical masks are often reasonable choices, especially for single-channel continuum images. Since the spatial distribution of molecular line emission in a protoplanetary disk changes considerably (but predictably) as a function of observing frequency (corresponding to velocity; e.g., Figure 3 of Öberg et al. 2021), using the same elliptical mask for all channels means that large areas of blank sky will be at risk of having erroneous CLEAN components placed within them.

CASA allows masks to be hand-drawn for complicated spatial emission. This is often the most accurate option, but mask drawing quickly becomes onerous for even a single large spectral cube. For the MAPS LP, the time investment is prohibitive. CASA's `auto-multithresh` algorithm (Kepley et al. 2020) is a promising solution; however, given its computational overhead, we instead exploited the Keplerian rotation pattern of the protoplanetary disk (e.g., Horne & Marsh 1986; Semenov et al. 2008) to create parametric CLEAN masks (e.g., Rosenfeld et al. 2013).

5.1. Keplerian Masking

We generated a Keplerian mask for each disk by the following procedure. Each image plane pixel (l, m) was first deprojected into disk-centric cylindrical coordinates (r, ϕ, z)$_{\rm disk}$, based on an assumed disk inclination i, position angle, and constant emission surface slope (z/r; e.g., Teague et al. 2019) representing the surface of an optically thick cone in 3D. The disk inclination and position angle values were drawn from the literature (see Öberg et al. 2021, Table 1), while the emission surface slope was refined by hand as described below. The disk-centric coordinates correspond to the location where a ray drawn from pixel (l, m) intersects the surface of the cone. Depending on the disk inclination and orientation, not all pixels necessarily intersect the cone. Using the disk-centric coordinates for each pixel, the projected Keplerian velocity is

$$v_{\rm Kep,proj}(r,\,z) = \sqrt{\frac{GM_\star r^2}{(r^2+z^2)^{3/2}}}\,\sin(i)\cos(\phi) \qquad (16)$$

where $M_\star$ is the central stellar mass. The total projected velocity component at each pixel is the sum of the Keplerian rotation and the systemic velocity, $v_0 = v_{\rm Kep,proj} + v_{\rm LSR}$.

Motivated by the fact that disk temperature declines with increasing r, the intrinsic line width is also assumed to narrow following a power-law profile:

$$\Delta V(r) = \Delta V_0 \times \left(\frac{r}{1''}\right)^{\Delta V_q} \qquad (17)$$

where $\Delta V_q \leqslant 0$. An initial mask was generated such that for a channel with central velocity $v_{\rm chan}$ and width $\Delta v_{\rm chan}$, masked regions satisfied $|v_{\rm chan} - v_0| \leqslant (\Delta v_{\rm chan} + \Delta V)$. For image cubes with low spectral resolution, the channel spacing $\Delta v_{\rm chan}$ dominates the mask. For image cubes with high spectral resolution, the line width ΔV is of greater importance in determining masked regions. After the initial mask was

Table 1
Keplerian Mask Parameters

Source	$r_{\rm out}$ ($''$)	ΔV_0 (m s^{-1})	ΔV_q	z/r	$\theta_{\rm conv}$ ($''$)
IM Lup	5.8	400	−1.0	0.3	0.5
AS 209	1.8	400	−0.5	0.1	0.5
GM Aur	3.0	500	−0.5	0.2	0.3
HD 163296	5.6	500	−1.0	0.3	0.3
MWC 480	3.0	300	−0.5	0.3	0.3

generated, it was convolved with a 2D Gaussian profile to smooth the mask edges (with the FWHM $\theta_{\rm conv}$ tuned to provide an adequate "buffer" region near otherwise sharp contours). The full script used to generate the masks is available in the `keplerian_mask` package (Teague 2020).

We tuned the mask parameters to match the ^{13}CO $J = 2 - 1$ emission for each disk; the final parameters are listed in Table 1. The mask parameters should not be considered true properties of the protoplanetary system; rather they are the parameters that best represent the spatial distribution of emission using the framework described above. The same mask parameters were used for all default image products for all observed transitions, with the exception of ^{12}CO $J = 2 - 1$, whose extended and complex emission morphology warranted bespoke, hand-drawn masks. For some specialized applications, it was advantageous to tune the Keplerian mask parameters from the values used here to produce the default image products. Where applicable, these alternative masks are noted in MAPS publications.

5.2. Stippling with Deeply CLEANed Models

In Figure 5 we compare the results of the CLEANing process for three channels of an example measurement set (AS 209 ^{13}CO $J = 2 - 1$) using an elliptical mask versus a Keplerian mask. To demonstrate the impact of the differently shaped masks, we CLEANed to two different depths: $4 \times$ rms and $1 \times$ rms. As noted previously, for CLEAN applications we used the rms measured from a signal-free region prior to the JvM correction.

If the measurement set is properly calibrated and the `tclean` parameters are set optimally—i.e., the basis set is an adequate match to the source morphology, loop gain is sufficiently conservative given the complexity of the deconvolution task, and the CLEANing threshold is appropriate—a well-designed mask will not significantly improve the end result (see elliptical versus Keplerian masks at a threshold $= 4 \times$ rms in Figure 5). If the algorithm is tuned correctly, CLEAN will eventually converge as it steadily but surely deconvolves real astrophysical flux (presumed to be responsible for the brightest features in the residual map at each iteration) down to the threshold. Through careful experimentation, we arrived at $4 \times$ rms as the ideal cleaning depth to balance the number of erroneous CLEAN components against flux retained in the residual map.

However, the optimal `tclean` parameters can be difficult to ascertain without some trial and error, and the threshold is an example of a parameter that is sometimes difficult to tune a priori. In these instances, properly fitting binary masks can guard against the incorrect placement of CLEAN components and mitigate (but not eliminate) some of the adverse effects of improperly set `tclean` parameters. In the `threshold`=1 $\times$ rms example in Figure 5, we see that numerous low-amplitude CLEAN components are added to the model, corresponding to noise spikes above the threshold level that were erroneously deconvolved from the residual map (top row: "conv. model"). Because the elliptical mask encompasses a larger area without plausible astrophysical flux, there are more opportunities for components to be incorrectly deconvolved.

Under the standard CLEAN workflow, the visual appearance of the final product (labeled "standard" in Figure 5) is not significantly affected at either threshold. However, we remind the reader that the flux units of the CLEANed image produced under the standard workflow are inconsistent (Section 4), and therefore any residual flux that was not deconvolved (particularly that outside the mask) is not represented with the appropriate strength. Moreover, as a matter of principle, it is not desirable to have a CLEAN model containing components known to correspond to pure noise, especially if it is to be used for further data reduction or analysis (e.g., for self-calibration; Brogan et al. 2018).

When the residual map and CLEAN model are properly combined under the JvM workflow (labeled "JvM" in Figure 5), the defects of the CLEAN model become apparent in the form of a stippled pattern. This stippling occurs because when $\epsilon < 1$, the flux distribution loses support under the CLEAN model: the CLEAN algorithm deconvolves the residual map with a dirty beam that is substantially larger than the CLEAN beam can restore. When the residual map is properly scaled to janskys per CLEAN beam using JvM correction, the previously artificially high residual map drops to a lower level and the stippling pattern appears. The stippling effect is most apparent in regions of low astrophysical flux and might lead to the misinterpretation of substructures in what would otherwise be interpreted as smooth regions (Disk Dynamics Collaboration et al. 2020). As mentioned, we used the carefully tuned $4 \times$ rms cleaning depth to minimize the appearance of stippling across the MAPS data products (see the panels with blue borders in Figure 5).

6. *uv* Tapering to Improve Beam Shape

The primary goal of this section is to describe how the tapering coefficients T_k of the dirty image equation (Equation 5) can be modified to standardize the PSF characteristics throughout the MAPS LP and aid comparisons between molecules and across disks. However, because PSF tapering is interrelated with the density weighting coefficients D_k, we first briefly review how uniform, natural, and robust weighting affect PSF shape through *uv* cell averaging and density weighting (see also Chapter 10.2.2 of Thompson et al. 2017). We also comment on how the images generated with these weighting schemes relate back to the data likelihood originally discussed in Section 1. Then, we discuss how the tapering and density weighting coefficients can be covaried to achieve a range of desired PSFs. In Figure 6 we demonstrate how the density weighting schemes and tapering coefficients affect the beam sidelobes, beam (non)axisymmetry, and JvM correction factor.

6.1. Cell Averaging, Density Weighting, and Maximum Likelihood Images

As discussed in the introduction (Section 1), Equation 5 is used to synthesize a dirty image from a set of discrete visibility samples $\{V_{{\rm data},k}\}_{k=1}^{N}$. In practice, visibilities and their Hermitian conjugates are first gridded, or interpolated, onto a regularly

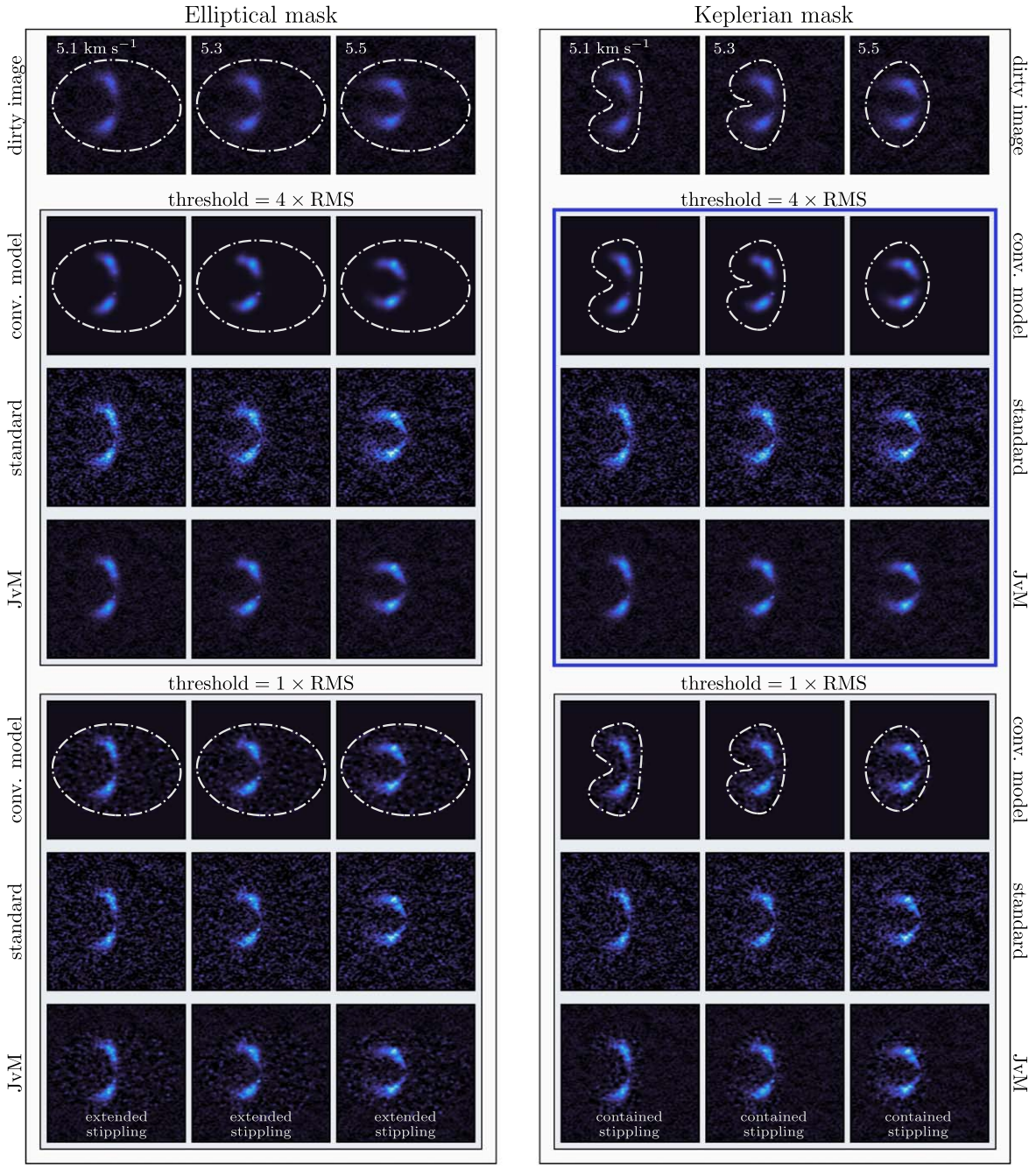

Figure 5. A comparison of elliptical (left column) vs. Keplerian (right column) binary masks for three channels of the AS 209 ^{13}CO $J = 2 - 1$ measurement set CLEANed to an ideal depth of 4 × rms (top grouping) and over-CLEANed to a greater yet still plausible depth of 1 × rms (bottom grouping). All panels are on the same color scale stretch. If the parameters of the `tclean` algorithm are set optimally, a well-designed mask will not significantly improve image quality over a generic mask, as demonstrated by the identical images generated using the 4 × rms threshold. However, if an image is CLEANed to a deep noise threshold, a poorly designed mask will more easily allow components to be erroneously added to the CLEAN model. These errant components are barely noticeable in the standard CLEAN workflow (which unfortunately has inconsistent flux units; Figure 3), but when the JvM correction scales the residuals by a factor of ϵ the stippled pattern becomes apparent. Well-designed Keplerian masks can mitigate but not eliminate this behavior. The default choices for MAPS data products are labeled with a thick blue border (Keplerian masks, `threshold`=4 × rms).

spaced array of u_i, v_j points[38] so that the fast Fourier transform may be used to accelerate the imaging calculation. This grid presents a convenient partition centered on the u_i, v_j pairs with cell sizes Δu and Δv that can be (a) used to average multiple visibilities contained within the cell and (b) used as a bounding box to calculate the local sample density. The cell sizes used for the averaging and density calculations need not be the same as those used in the gridded array, but for demonstration purposes we assume that they are.

For the sake of discussion, we consider an arbitrary cell i, j containing L visibilities, indexed such that each l index

[38] The method by which visibilities are interpolated to grid center points turns out to have implications for image fidelity, especially for high dynamic range applications. Nearest-neighbor or linear interpolation can result in significant numerical errors; convolutional regridding schemes using prolate spheroidal wave functions are good choices to minimize these errors (Schwab 1984; Czekala et al. 2015). For the discussion that follows, however, the main conclusions are unchanged if simple nearest-neighbor interpolation is assumed.

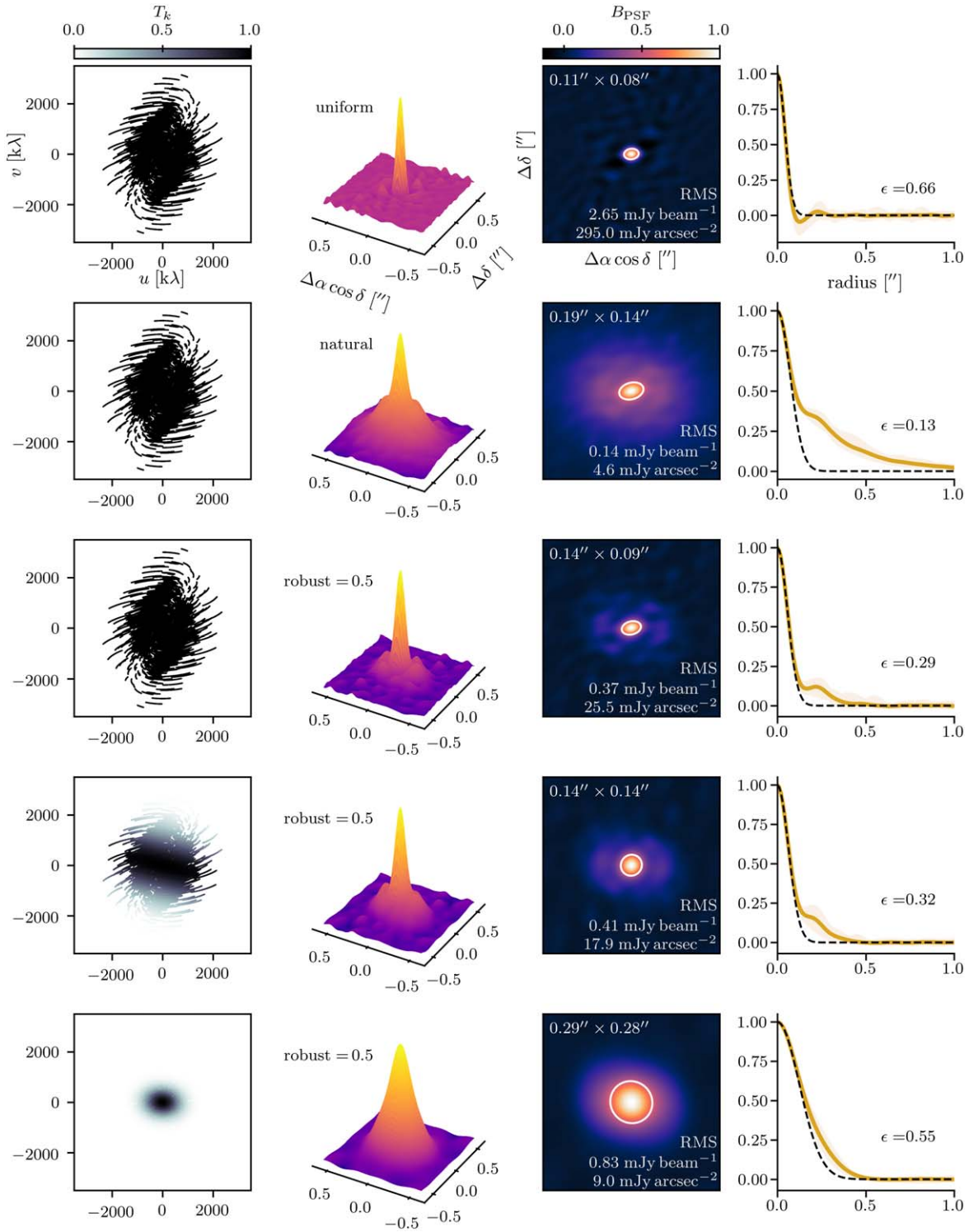

Figure 6. A representation of how the visibility tapering profile $T(u, v)$ (left column) and density weighting scheme influence the dirty beam resolution, point-source sensitivity, PSF profile, and JvM correction factor ϵ. The same dirty beam profile is shown in the second, third, and fourth columns: the 3D view best represents the beam structure, the 2D view best conveys the beam ellipticity (the white ellipse represents the FWHM), and the 1D view best represents the divergence from a Gaussian CLEAN beam. Each row represents a different combination of tapering profile and robust value. The first three rows are untapered, while the final two rows are tapered to circular FWHM resolutions of 0.″15 and 0.″30, respectively (the final $\leqslant$5% adjustment is carried out using the `imsmooth` task, which is not shown).

corresponds to some specific k index. For example, $l = \{1, 2, 3, \ldots, L\}$ might correspond to $k = \{304, 305, 307, \ldots, 320\}$, depending on how exactly the visibilities are indexed relative to the grid boundaries. When all L visibilities inside the cell are gridded, the "effective" visibility located at the cell center is

$$V_{\mathrm{grid},i,j} = \sum_{l=1}^{L} T_k D_k w_k V_{\mathrm{data},k}. \tag{18}$$

To simplify the comparison of density weights that follows, we will assume that there is no tapering ($T_k = 1$) and that all visibilities within a cell have a common expectation value $\langle V_{\mathrm{data},k}\rangle_{i,j} = \mathcal{V}(u_i, v_j)$, since the cell size is assumed to be small relative to the expected features in the visibility function (i.e., the chosen image size and resolution "Nyquist-sample" the visibility function).

Uniform weighting. Uniform weighting is arrived at by minimizing the sidelobe level of the main beam (for a derivation, see Section 3.3.3 of Briggs 1995), which yields constant density weights for all visibilities within the cell:

$$D_k = \frac{1}{w_{\mathrm{cell},i,j}} \tag{19}$$

where $w_{\mathrm{cell},i,j} = \sum_{l=1}^{L} w_l$. The effective value of the uniformly gridded visibility is

$$V_{\mathrm{grid},i,j} = \frac{\sum_{l=1}^{L} w_l V_{\mathrm{data},l}}{\sum_{l=1}^{L} w_l} = \langle V_{\mathrm{data},k}\rangle_{i,j}. \tag{20}$$

Uniform weighting delivers the minimum variance estimate of the visibility mean, since it is obtained by weighting each sample by its inverse variance ($w_k = 1/\sigma_k^2$). Moreover, the statistical uncertainty on $V_{\mathrm{grid},i,j}$ is encapsulated with $w_{\mathrm{cell},i,j}$ in the same way as that for ungridded visibilities $V_{\mathrm{data},k}$ with w_k. While uniform weighting usually produces the highest-resolution dirty beam (Figure 6, first row), the downside is that the dirty image usually has the worst sensitivity.

Natural weighting. The rms thermal noise in the dirty image is minimized when

$$D_k = 1, \tag{21}$$

since $w_k = 1/\sigma_k^2$ (for a derivation, see Section 3.3.1 of Briggs 1995). This is called natural weighting and results in an image that has the lowest rms and thus the greatest sensitivity. The effective value of the naturally gridded visibility is

$$V_{\mathrm{grid},i,j} = \sum_{l=1}^{L} w_l V_{\mathrm{data},l}. \tag{22}$$

Natural weighting diverges from estimating $\langle V_{\mathrm{data},k}\rangle_{i,j}$ and instead upweights those grid cells containing large quantities of high-S/N visibilities. Natural weighting results in a beam (Figure 6, second row) that is useful for detection (especially of point sources), but at the cost of substantial sidelobes negatively impacting resolution.

Robust weighting. Briggs (1995) developed the robust weighting scheme:

$$D_k = \frac{1}{1 + w_{\mathrm{cell},i,j} f^2} \tag{23}$$

where

$$f^2 = (5 \times 10^{-R})^2 \Bigg/ \left(\frac{\sum_i \sum_j w_{\mathrm{cell},i,j}^2}{\sum_{k=1}^{N^\dagger} w_k} \right). \tag{24}$$

Values of $R = 2$ approximate natural weighting (maximizing point-source sensitivity) while values of $R = -2$ approximate uniform weighting (maximizing resolution). Intermediate R values offer a trade-off between these two extremes. Because the performance curve is nonlinear (Figure 3.23 of Briggs 1995), significant resolution gains can be achieved at minimal loss in point-source sensitivity (e.g., in millijanskys), though loss in surface brightness sensitivity is more significant (see Figure 6, third column), especially when it is expressed in units of constant solid angles (e.g., mJy arcsec^{-2}) or brightness temperature (e.g., K).

The effective value of the robustly gridded visibility is

$$V_{\mathrm{grid},i,j} = \sum_{l=1}^{L} \frac{w_l V_{\mathrm{data},l}}{1 + w_{\mathrm{cell},i,j} f^2} \Bigg/ \sum_{k=1}^{N^\dagger} \frac{w_k}{1 + w_{\mathrm{cell},i,j} f^2}. \tag{25}$$

Maximum likelihood images. If the visibility function $\mathcal{V}$ were sampled at all u, v locations where it had significant power, then V_{grid} would have a direct relationship to the sky brightness via the inverse Fourier transform. Unfortunately, this is never achieved for actual submillimeter interferometric observations of astrophysical sources, and the impact of the transfer function ($W(u, v)$, Equation (8)) must be considered when producing images from visibility samples.

Nevertheless, because visibility model fitting (Section 1) is a common and highly useful procedure for parameter inference—especially for annular protoplanetary disk structures—it is worthwhile to consider how uniform, natural, and robust images relate back to $\mathcal{V}$. By referencing the effective visibility under each weighting scheme (Equations (20), (22), and (25)), we see that only uniform weighting converges to the expectation value of the visibility function in that cell (Briggs 1995), while natural and robust weighting instead result in an effective visibility that is tied to the number and quality of visibility measurements within that cell. Therefore if the dirty image were used as a sky plane model $I(l, m)$, only the uniform dirty image would maximize the data likelihood function (Equation 3).

Of course, there are legitimate reasons for using other weighting schemes, such as the need to maximize sensitivity with natural weighting for a detection experiment. And in practice, given its nonlinear trade-off curve, the most scientifically useful weighting scheme for the detection and characterization of faint protoplanetary disk features is often an intermediate robust value, like the `robust=0.5` we used as the default setting for untapered MAPS products.

The CLEANing process and maximum likelihood. The dirty image is constructed under the assumption that all unsampled spatial frequency components are set to zero power. As the CLEANing process deconvolves the dirty image and image plane components are added to the CLEAN model, the corresponding visibility plane model V_{model} is effectively interpolated from the gridded locations to otherwise unsampled (u, v) values. Because the image plane representation is directly equivalent to the inverse Fourier transform of the visibility plane model, the accuracy of the interpolated values relative to truth mirrors the quality of the deconvolution process. Using a basis set appropriate to the source morphology (e.g., choosing multiscale CLEAN over Högbom CLEAN for extended emission) can result in more accurate Fourier interpolations/image reconstructions, particularly at higher spatial frequencies/finer resolutions.

Each iteration of the CLEAN algorithm also brings V_{model} closer to maximizing the likelihood function (Schwarz 1978). However, V_{model} will only achieve the maximum possible likelihood value when $V_{\mathrm{model},k} = V_{\mathrm{data},k}\ \forall k \in 1, \ldots, N$, which

can happen in the absence of noise or in the case where the entire image is fully cleaned to a 0 Jy threshold (the Fourier equivalent of overfitting). Most practical applications will stop CLEANing when a noise threshold is reached in the residual map.

Because imaging is an ill-defined inverse process, there exist an infinite number of maximum (or approximately maximum) likelihood images (and therefore models) consistent with the data: these have $V_{\mathrm{model},k} \approx V_{\mathrm{data},k}\, \forall k \in 1, \ldots, N$ but may take on any value of V_{model} at unsampled spatial frequencies u, v. Because beam characteristics will affect the order in and scales at which flux is deconvolved from the dirty image, for any spatially resolved source the deconvolution algorithm will likely converge to different CLEAN models under different visibility weighting schemes.

6.2. Tapering

Visibility tapering profiles can be used to "force" beams of arbitrary dimensions (Briggs 1995). For MAPS, we produced fiducial data products using tapered visibilities to boost image plane sensitivity to low surface brightness features (at the expense of resolution) and provide a common circular beam that is useful for comparing molecular species across the Band 3 and Band 6 observations (for a listing of all fiducial beam sizes, see Table 5 of Öberg et al. 2021).

Beam tapers can be described by a multiplicative profile in the uv-plane $T(u, v)$ or equivalently by convolution with a kernel in the image plane $t(l, m)$. If we wish to taper our original dirty beam, whose main lobe is approximately an elliptical Gaussian, to a dirty beam whose main lobe is approximately a circular Gaussian, the appropriate functional form of the taper is an elliptical Gaussian. By analogy with the CLEAN beam (Section 2), this profile is described by a position angle (ϕ_T) and beam dimensions. In the uv-plane, these FWHMs are Θ_a, Θ_b in units of kilolambdas. In the image plane, these are θ_a, θ_b in units of arcseconds. The Gaussian FWHMs are related[39] such that

$$\Theta_a = 10^{-3}\left(\frac{60^2 \times 180''}{\pi \text{ radians}}\right)\frac{4\ln 2}{\pi\theta_a} \tag{27}$$

$$\Theta_b = 10^{-3}\left(\frac{60^2 \times 180''}{\pi \text{ radians}}\right)\frac{4\ln 2}{\pi\theta_b}. \tag{28}$$

We follow a convention such that $\Theta_a > \Theta_b$, which implies that $\theta_a < \theta_b$. The multiplicative uv-plane profile is

$$T(u, v) = \exp\left(-\frac{1}{2}\left[\left(\frac{u'}{\sigma_{u'}}\right)^2 + \left(\frac{v'}{\sigma_{v'}}\right)^2\right]\right) \tag{29}$$

with

$$u' = u\cos\phi_T - v\sin\phi_T \tag{30}$$

$$v' = u\sin\phi_T + v\cos\phi_T \tag{31}$$

and

$$\sigma_{u'} = \sigma_u = \Theta_b/(2\sqrt{2\ln 2}) \tag{32}$$

$$\sigma_{v'} = \sigma_v = \Theta_a/(2\sqrt{2\ln 2}). \tag{33}$$

In the image plane, the convolutional profile is

$$t(l, m) = \exp\left(-\frac{1}{2}\left[\left(\frac{l'}{\sigma_{l'}}\right)^2 + \left(\frac{m'}{\sigma_{m'}}\right)^2\right]\right) \tag{34}$$

with

$$l' = l\cos\phi_T - l\sin\phi_T \tag{35}$$

$$m' = m\sin\phi_T + m\cos\phi_T \tag{36}$$

and

$$\sigma_{l'} = \sigma_l = \theta_b/(2\sqrt{2\ln 2}) \tag{37}$$

$$\sigma_{m'} = \sigma_m = \theta_a/(2\sqrt{2\ln 2}). \tag{38}$$

Both the uv tapering and image convolutional profiles are normalized to peak values of 1. Although it is possible to analytically calculate tapering parameters given a desired beam size (Appendix C of Briggs 1995), we found that there is considerable "mechanical backlash" in CASA v6.1.0's beam fitting subroutines due to the pixelized representation of the dirty beam. This means that a smooth, monotonic relationship between the tapered dirty beam size and the fitted CLEAN beam size does not exist; in practice we found that direct forward modeling yielded a more consistent set of tapering parameters. We emulated CASA's dirty beam formulation routine[40] to compare the fitted CLEAN beam (with FWHMs θ'_a and θ'_b) to a CLEAN beam with the desired parameters (θ_a, θ_b, ϕ). We desired a circularized beam such that $\theta_a = \theta_b = \theta_T$, with sizes $\theta_T = \{0\farcs15, 0\farcs2, 0\farcs3\}$ for Band 6 observations and $\theta_T = \{0\farcs3, 0\farcs5\}$ for Band 3 observations.

Because density weighting also affects beam shape, there are many combinations of `robust` values and tapering profiles that deliver the same circularized CLEAN beam profile. To preserve point-source sensitivity without introducing large sidelobes, we calculated tapering profiles using a starting value of `robust=0.5`. We used `scipy.optimize.minimize` (Virtanen et al. 2020) to find the best-fitting set of beam tapering parameters that minimized the fit metric:

$$f(\Theta_a, \Theta_b, \phi_T) = (\theta_t - \theta'_a)^2 + (\theta_t - \theta'_b)^2. \tag{39}$$

If no solution could be found that delivered a beam within 95%–100% of the target resolution, we iterated by reducing the `robust` value by `0.25` (toward a uniformly weighted beam) and tried the fitting procedure again. For most of the Band 6 observations (including those shown in Figure 6), we were able to successfully find tapered beams using `robust=0.5` values. For several of the Band 3 observations, however, lower robust values were required to achieve pretapered beams with $\theta_a \leqslant 0\farcs3$. Examples of the visibility tapering profiles needed to achieve $\theta_T = \{0\farcs15, 0\farcs3\}$ are shown in the fourth and fifth rows of Figure 6, respectively. As a final step (not shown in Figure 6), the images were smoothed to the remaining <5% of

[39] We derived the FWHM relationships following the Fourier dual relationships for Gaussians, which are more simply phrased in terms of σ values: $\sigma_{u,v} = 1/(2\pi\sigma_{l,m})$ (e.g., Bracewell 2000). It appears there is an inconsistency in the CASA 6.1 implementation of `uvtaper` in `tclean` where the incorrect FWHM relationship

$$\Theta = 10^{-3}\left(\frac{60^2 \times 180''}{\pi \text{ radians}}\right)\frac{1}{\theta} \tag{26}$$

is used to calculate the `uvtaper` profile, which equates to a scale factor difference of ≈1.13. We emulated this behavior such that the final images were tapered to the correct target resolution.

[40] https://github.com/ryanaloomis/beams_and_weighting

their target resolution using the CASA `imsmooth` task. By carrying out the bulk of the tapering in the *uv*-plane (instead of entirely on the final image), the CLEAN algorithm is able to build a more accurate CLEAN model during the deconvolution process and thus improve final-image fidelity.

While our tapering profiles were successful in standardizing beam sizes across the MAPS data products, it is important to realize that even tapered dirty beams whose main lobes deliver circular Gaussian CLEAN beams have an extended asymmetric shelf. For example, consider the 2D PSF representation of the $0\farcs15$ tapered beam in the fourth row, third column, of Figure 6. This means that any astrophysical flux remaining in the residual map will retain the asymmetric features of the dirty beam.

7. S/N in MAPS Products

The MAPS spectral setup (Öberg et al. 2021, Table 4) covered more than 40 molecular transitions of interest, many of which are at or near the detection threshold in the five protoplanetary disks we targeted. Given the diverse scientific goals of the MAPS LP, we employed several algorithms to detect and characterize the emission. In this section we discuss the general principles behind the quantification of S/N and how these apply to several core data products from the LP.

Colloquially, S/N is usually quoted as a 1D quantity in multiples of a σ value corresponding to a fractional probability of the Gaussian distribution, e.g., 2σ or 95% (regardless of whether the posterior distribution itself is actually Gaussian). Model assumptions are key to contextualizing any S/N statistic, since it is these (often hidden) assumptions that define the likelihood function and prior distributions and thus the posterior distribution of the parameter(s) of interest.

S/N in detection experiments. Detection experiments are usually discussed in terms of a 1D posterior probability distribution of the amplitude a of the source $p(a \mid \text{data})$. The detection S/N is the probability that the inferred amplitude of the source is greater than 0, $p(a > 0 \mid \text{data})$. In truth, the full posterior distribution is likely to be multivariate because the model usually has hidden parameters—b, c, etc. In the best situations, the uncertainty pertaining to other model parameters is marginalized out:

$$p(a \mid \text{data}) = \int p(a, b, c \mid \text{data}) db\, dc. \tag{40}$$

However, for computational or implementation reasons, it may be difficult to explore the probability distributions of these other parameters, and the S/N is estimated using the conditional posterior distribution $p(a \mid b, c, \text{data})$ with all other model parameters held fixed. This may overestimate the detection S/N when the model varies significantly under reasonable choices for the other parameters.

Consider the scenario of detecting a point source against a blank background with Gaussian noise. We assume a simple model: a δ-function with a known position but an unknown amplitude. Assuming the beam is well characterized, the amplitude posterior is defined by a Gaussian centered on the value of flux measured at the location of the point source. The width of the posterior Gaussian corresponds to the thermal rms of the image. The fraction of the posterior corresponding to fluxes greater than 0 defines the S/N detection probability (e.g., 2σ or 95%). Most realistic scenarios quickly diverge from this idealized scenario to involve more complex models (for example, if the location of the δ-function is not known, then the l and m positions must be marginalized over and the significance of any particular candidate is diminished).

7.1. Spatially Resolved Emission

When considering spatially resolved emission, it is important to consider the model assumptions that undergird the interpretation of S/N. If the model is misspecified (e.g., a point source when the emission is in fact diffuse) then the S/N calculation, conditional on those assumptions, may not reflect the S/N calculation one actually desires. The more closely a model matches reality, generally speaking, the higher the detection significance that can be achieved. If the model is overly flexible, however, then unknown parameters of the search space (e.g., spatial location or rest frequency) must be marginalized over and the S/N of the detection will suffer. One way to gain intuition for model sensitivity is to attempt to fit the data with a range of different model assumptions and effectively explore some of the hidden parameters, albeit in a limited manner.

Matched filter. The application of a matched filter to detect spatially resolved but weak line emission from a protoplanetary disk in Keplerian rotation nicely demonstrates some of the model complexity trade-offs (e.g., Loomis et al. 2018). In such a framework, a template (such as uniform surface brightness inside some Keplerian pixel mask; e.g., Table 1) is assumed, Fourier-transformed, and cross-correlated against all available frequency channels. The template amplitude is the sole free parameter of the model. Since the matched filter is linear, the interpretation of the filter response relative to a signal-free region (σ) is directly equivalent to the point-source detection scenario discussed above, and carries many of the same caveats (Ruffio et al. 2017; Loomis et al. 2018).

Detection significance will be maximized when the template is an accurate representation of reality; in most real-world applications an imperfect template will reduce the significance of a detection. This is demonstrated to some degree in Figure 7 with the matched filter applied to the HC_3N $J = 11 - 10$ transition in MWC 480, using templates generated from Keplerian masks with different outer radii (100, 200, and 400 au). The line is detected with all three templates; however, the significance is maximized when using the template with the smallest radial extent (100 au), which best matches the emission morphology of this weak, compact line. Since even the 100 au template is misspecified to some degree (real disk emission does not appear to be uniform within some Keplerian mask, but rather radially varies in intensity), the matched-filter detections most likely represent lower limits to the maximal S/N detection that could be achieved with an ideal template. More details on the matched-filter procedure applied to MAPS data are provided in Ilee et al. (2021).

Shifting and stacking. To aid the detection and characterization of weak lines, we used the shift-and-stack technique (Teague et al. 2016; Yen et al. 2016; Matrà et al. 2017) on several image cubes. In this process, the Doppler shift corresponding to the Keplerian rotation of the protoplanetary disk is removed from each position–position–velocity pixel in the data cube and each channel in the cube is summed across all pixels. This coherently sums emission across the data cube and results in a higher-S/N detection compared to a traditional spatially integrated spectrum. For a full description of the shift-and-stack technique applied to MAPS data, as well as the

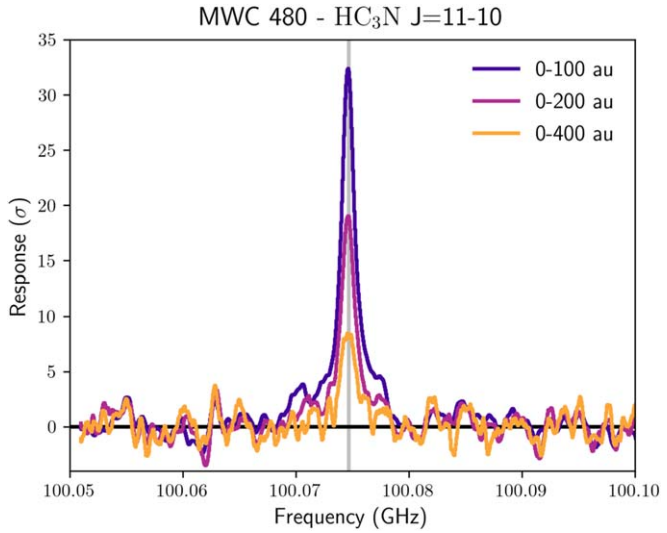

Figure 7. The response of Keplerian matched-filter templates with outer radii of 100, 200, and 400 au applied to the MWC 480 HC_3N $J = 11 - 10$ transition measurement set. The detection significance is higher when the template more accurately matches the true spatial distribution of the (inherently compact) emission.

calculation of the uncertainties that result from the deprojection and aggregation of the data, see Ilee et al. (2021) and Cataldi et al. (2021).

An important benefit to using multiple techniques to detect and characterize faint molecular emission is that confidence builds when independent techniques return consistent results. In Figure 8 we compare the results of the matched filter and spectral shift-and-stack technique applied to the $H^{13}CO^+$ $J = 1 - 0$ transition in the HD 163296 and MWC 480 disks (for more information on this and the HCO^+ $J = 1 - 0$ transition, see Aikawa et al. 2021). The $H^{13}CO^+$ emission is faint—it cannot be recovered by visual inspection of the channel maps for either disk. Encouragingly, the emission is detected in HD 163296 using both the matched filter (with a Keplerian mask template $r_{out} = 400$ au) and the spectral shift-and-stack technique, while the line is not detected in MWC 480 using either technique.

The spectral shifting and stacking technique is useful for measuring the disk-integrated flux and radial intensity profile (especially for transitions with hyperfine structures; e.g., Bergner et al. 2021; Cataldi et al. 2021; Guzmán et al. 2021), while the matched filter is a more efficient way to search the full MAPS LP for weak transitions (instead of imaging the entire data set). We applied the matched filter to all of the spectral windows to search for any additional lines that were not primary science targets and serendipitously detected satellite lines of c-C_3H_2 and C_2H (Guzmán et al. 2021; Ilee et al. 2021). For a full list of the molecules detected by the MAPS LP, see Öberg et al. (2021, Tables 2 and 3).

8. Summary

The MAPS LP represented a significant effort to calibrate and image a large volume of molecular line emission for five protoplanetary disks. In this work, we have described the non-Gaussian dirty beam that can arise from multiconfiguration ALMA observations and, after reviewing the CLEANing process, the challenges it presented for accurate flux recovery of faint, extended features. We chose to remedy this issue by implementing the JvM correction originally proposed by Jorsater & van Moorsel (1995), which properly scales the CLEAN residual map into consistent units of janskys per CLEAN beam. We have also described how we generated custom Keplerian CLEAN masks to guide the deconvolution process, and discussed some imaging artifacts that can result from nonideal `tclean` parameters, such as stippling. To aid the comparison of different molecular transitions, we also produced image products using beam profiles tapered to common resolutions of FWHM 0.″15, 0.″2, 0.″3, and 0.″5. Finally, we have briefly discussed the interpretation of S/N and detection significance across the MAPS data products.

We centralized our data processing pipeline on the North American ALMA Science Center (NAASC) computing cluster, located in Charlottesville, VA, USA. Python scripts documenting the reduction and imaging procedures are available.[41] During times of heavy development, we availed ourselves of multiple 16-core machines for days at a time. Including development and incremental reprocessing efforts, we estimate that we utilized 1 yr worth of core hours to produce the MAPS LP image products (i.e., two 16-core machines fully utilized for two weeks). Though still small compared to the computational demands of protoplanetary disk hydrodynamical simulations, for example, this represents a considerably larger computational demand compared to that of most other ALMA observations.

9. Imaging Data Products

Here we enumerate the molecular line imaging data products of the MAPS LP, which are accessible through a portal to the ALMA archive available after peer review. For a detailed description of the naming conventions for all MAPS LP image products, see Section 3.5 of Öberg et al. (2021). For a description of the continuum-only MAPS image products, see Sierra et al. (2021).

For each combination of disk (e.g., MWC 480; see Table 1 of Öberg et al. 2021) and transition (e.g., DCN $J = 3 - 2$; see Tables 2 and 3 of Öberg et al. 2021), the archive contains two minimal measurement sets produced with the `cvel2` and `split` tasks with visibilities pertaining to that disk and transition pair. The first contains the visibilities including continuum emission, while the second contains the line visibilities with the continuum subtracted. During the invocation of `cvel2`, we coarsened the spectral channels slightly from their native spacings (see Table 4 of Öberg et al. 2021) to a uniform set of channels spaced 0.5 km s^{-1} apart in Band 3 and 0.2 km s^{-1} apart in Band 6.

The archive also contains a set of image products generated from each measurement set using various beams. For the Band 3 transitions, these beams are untapered `robust=0.5`, tapered 0.″30, and tapered 0.″50. For the Band 6 transitions, these beams are untapered `robust=0.5`, tapered 0.″15, tapered 0.″20, and tapered 0.″30. For a full description of the beam sizes available in each band, see Table 5 of Öberg et al. (2021). For each beam setting, the following image products are available:

1. Keplerian binary CLEAN mask cube
2. (Unconvolved) CLEAN model cube

[41] www.almamaps.info

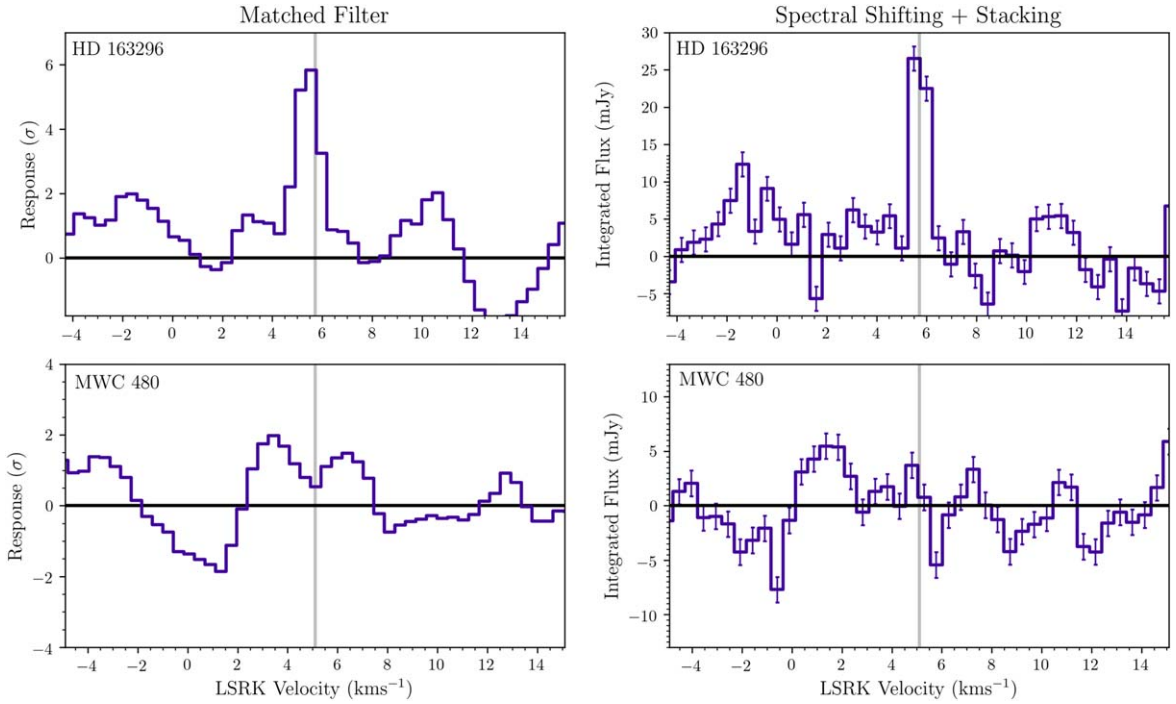

Figure 8. A comparison of the matched-filter (left column) and spectral shift-and-stack (right column) detection techniques applied to the $H^{13}CO^+$ $J = 1 - 0$ transition in HD 163296 (top row) and MWC 480 (bottom row). In HD 163296, the transition is detected by both techniques at similar significance (relative to the off-source velocities), but in MWC 480 it is not detected by either technique, demonstrating consistency between the two methods.

3. Residual cube from CLEANing process (before JvM correction)
4. CLEANed image cube under standard workflow (Figure 3), with and without primary-beam correction
5. CLEANed image cube under JvM correction (Figure 3), with and without primary-beam correction. The primary-beam-corrected and JvM-corrected cube is the recommended data cube for most scientific use cases.
6. A Python script to reproduce the data products from the minimal continuum-subtracted measurement set

The following is a typical workflow to produce a set of image products. We used the `tclean` task with multiscale CLEAN and `scales=[0, 5, 15, 25]` pixels, where the pixel size was chosen to correspond to approximately one-seventh of the beam FWHM. For all lines except CO, we used Keplerian CLEAN masks matched to the ^{13}CO $J = 2 - 1$ emission. We iterated the CLEAN algorithm such that the peak residual emission was below a threshold of $4 \times$ rms. For untapered beams we used Briggs weighting of `robust=0.5`. For tapered beams we forward-modeled the CASA beam fitting process to calculate the value of the `uvtaper` argument that achieves the target resolution using the largest (most natural) `robust` value still $\leqslant 0.5$. Finally, we calculated the JvM factor ϵ via the ratio of the CLEAN beam volume to the dirty beam volume, scaled the residual map by ϵ, and summed it with the convolved CLEAN model to produce the JvM-corrected image cube. We recommend consulting the Python script accompanying each set of image products for the specific `tclean` parameters used to generate a particular image product.

For more information on the additional value-added data products like moment maps, radial profiles, and emission surfaces provided for most transitions across a range of beam sizes see Law et al. (2021a, 2021b).

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2018.1.01055.L. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. We thank the NAASC for allowing the use of computational resources on the NAASC computing cluster.

C.J.L. acknowledges funding from the NSF Graduate Research Fellowship Program under grant DGE1745303. R.T. acknowledges support from the Smithsonian Institution as a Submillimeter Array Fellow. K.I.Ö. acknowledges support from the Simons Foundation (SCOL No. 321183) and an NSF AAG grant (No. 1907653). J.H. and S.M.A. acknowledge funding support from the National Aeronautics and Space Administration under grant No. 17-XRP172-0012 issued through the Exoplanets Research Program. J.H. acknowledges support for this work provided by NASA through NASA Hubble Fellowship grant No. HST-HF2-51460.001-A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. I.C. was supported by NASA through NASA Hubble Fellowship grant No. HST-HF2-51405.001-A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Y.A. acknowledges support by NAOJ ALMA Scientific Research Grant code 2019-13B and Grant-in-Aid for Scientific Research Nos. 18H05222 and 20H05847. R.L.G. acknowledges support from a CNES fellowship grant.

K.Z. acknowledges the support of the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation, and the support of NASA through Hubble Fellowship grant No. HST-HF2-51401.001, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. C.W. acknowledges financial support from the University of Leeds, STFC, and UKRI (grant Nos. ST/R000549/1, ST/T000287/1, MR/T040726/1) J.D.I. acknowledges support from STFC under ST/T000287/1. K.R.S. acknowledges the support of NASA through Hubble Fellowship Program grant No. HST-HF2-51419.001, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. G.C. is supported by NAOJ ALMA Scientific Research grant code 2019-13B. J.B.B. acknowledges support from NASA through NASA Hubble Fellowship grant No. HST-HF2-51429.001-A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. J.B. acknowledges support by NASA through NASA Hubble Fellowship grant No. HST-HF2-51427.001-A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Y.Y. is supported by IGPEES, WINGS Program, the University of Tokyo. H.N. acknowledges support by NAOJ ALMA Scientific Research grant code 2018-10B and Grant-in-Aid for Scientific Research No. 18H05441. T.T. is supported by JSPS KAKENHI grant Nos. JP17K14244 and JP20K04017. F.M. acknowledges support from ANR of France under contracts ANR-16-CE31-0013 (Planet-Forming-Disks) and ANR-15-IDEX-02 (through CDP "Origins of Life"). L.M.P. acknowledges support from ANID project Basal AFB-170002 and from ANID FONDECYT Iniciación project No. 11181068. A.S.B acknowledges the studentship funded by the Science and Technology Facilities Council of the United Kingdom (STFC). A.D.B. and E.A.B. acknowledge support from NSF AAG grant No. 1907653. L.I.C. gratefully acknowledges support from the David and Lucille Packard Foundation and Johnson & Johnson's WiSTEM2D Program. V.V.G. acknowledges support from FONDECYT Iniciación 11180904 and ANID project Basal AFB-170002. A.R.W. acknowledges support from the Virginia Space Grant Consortium and the National Science Foundation Graduate Research Fellowship Program under grant No. DGE1842490. N.T.K. acknowledges support provided by the Alexander von Humboldt Foundation in the framework of the Sofja Kovalevskaja Award endowed by the Federal Ministry of Education and Research.

Facility: ALMA.

Software: Astropy (Astropy Collaboration et al. 2013, 2018), CASA (v6.1.0; McMullin et al. 2007), GoFish (Teague 2019), NumPy (van der Walt et al. 2011), SciPy (Virtanen et al. 2020), `keplerian_mask` (Teague 2020).

ORCID iDs

Ian Czekala https://orcid.org/0000-0002-1483-8811
Ryan A. Loomis https://orcid.org/0000-0002-8932-1219
Richard Teague https://orcid.org/0000-0003-1534-5186
Alice S. Booth https://orcid.org/0000-0003-2014-2121
Jane Huang https://orcid.org/0000-0001-6947-6072
Gianni Cataldi https://orcid.org/0000-0002-2700-9676
John D. Ilee https://orcid.org/0000-0003-1008-1142
Charles J. Law https://orcid.org/0000-0003-1413-1776
Catherine Walsh https://orcid.org/0000-0001-6078-786X
Arthur D. Bosman https://orcid.org/0000-0003-4001-3589
Viviana V. Guzmán https://orcid.org/0000-0003-4784-3040
Romane Le Gal https://orcid.org/0000-0003-1837-3772
Karin I. Öberg https://orcid.org/0000-0001-8798-1347
Yoshihide Yamato https://orcid.org/0000-0003-4099-6941
Yuri Aikawa https://orcid.org/0000-0003-3283-6884
Sean M. Andrews https://orcid.org/0000-0003-2253-2270
Jaehan Bae https://orcid.org/0000-0001-7258-770X
Edwin A. Bergin https://orcid.org/0000-0003-4179-6394
Jennifer B. Bergner https://orcid.org/0000-0002-8716-0482
L. Ilsedore Cleeves https://orcid.org/0000-0003-2076-8001
Nicolas T. Kurtovic https://orcid.org/0000-0002-2358-4796
François Ménard https://orcid.org/0000-0002-1637-7393
Hideko Nomura https://orcid.org/0000-0002-7058-7682
Laura M. Pérez https://orcid.org/0000-0002-1199-9564
Chunhua Qi https://orcid.org/0000-0001-8642-1786
Kamber R. Schwarz https://orcid.org/0000-0002-6429-9457
Takashi Tsukagoshi https://orcid.org/0000-0002-6034-2892
Abygail R. Waggoner https://orcid.org/0000-0002-1566-389X
David J. Wilner https://orcid.org/0000-0003-1526-7587
Ke Zhang https://orcid.org/0000-0002-0661-7517

References

Aikawa, Y., Cataldi, G., Yamato, Y., et al. 2021, ApJS, 257, 13
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bergner, J., Öberg, K. I., Guzmán, V. V., et al. 2021, ApJS, 257, 11
Bracewell, R. N. 2000, The Fourier Transform and Its Applications (Boston, MA: McGraw Hill)
Briggs, D. S. 1995, PhD thesis, New Mexico Institute of Mining and Technology
Brogan, C. L., Hunter, T. R., & Fomalont, E. B. 2018, arXiv:1805.05266
Cárcamo, M., Román, P. E., Casassus, S., Moral, V., & Rannou, F. R. 2018, A&C, 22, 16
Cataldi, G., Yamato, Y., Aikawa, Y., et al. 2021, ApJS, 257, 10
Chael, A. A., Johnson, M. D., Narayan, R., et al. 2016, ApJ, 829, 11
Cornwell, T. J. 2008, ISTSP, 2, 793
Cornwell, T. J., & Evans, K. F. 1985, A&A, 143, 77
Czekala, I., Andrews, S. M., Jensen, E. L. N., et al. 2015, ApJ, 806, 154
Czekala, I., Chiang, E., Andrews, S. M., et al. 2019, ApJ, 883, 22
Czekala, I., & Loomis, R. 2020, MPoL-dev/MPoL: v0.1.1 Release, Zenodo, doi:10.5281/zenodo.4939048
Disk Dynamics Collaboration, Armitage, P. J., Bae, J., et al. 2020, arXiv:2009.04345
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019, ApJL, 875, L4
Guzmán, V. V., Bergner, J. B., Law, C. J., et al. 2021, ApJS, 257, 6
Guzmán, V. V., Huang, J., Andrews, S. M., et al. 2018, ApJL, 869, L48
Högbom, J. A. 1974, A&AS, 15, 417
Hogg, D. W., & Foreman-Mackey, D. 2018, ApJS, 236, 11
Horne, K., & Marsh, T. R. 1986, MNRAS, 218, 761
Huang, J., Bergin, E. A., Öberg, K. I., et al. 2021, ApJS, 257, 19
Ilee, J., Walsh, C., Booth, A. S., et al. 2021, ApJS, 257, 9
Jennings, J., Booth, R. A., Tazzari, M., Rosotti, G. P., & Clarke, C. J. 2020, MNRAS, 495, 3209
Jorsater, S., & van Moorsel, G. A. 1995, AJ, 110, 2037
Kepley, A. A., Tsutsumi, T., Brogan, C. L., et al. 2020, PASP, 132, 024505
Law, C. J., Loomis, R. A., Teague, R., et al. 2021a, ApJS, 257, 3
Law, C. J., Teague, R., Loomis, R. A., et al. 2021b, ApJS, 257, 4
Loomis, R. A., Öberg, K. I., Andrews, S. M., et al. 2018, AJ, 155, 182

Matrà, L., MacGregor, M. A., Kalas, P., et al. 2017, ApJ, 842, 9
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
Nakazato, T., Ikeda, S., Akiyama, K., et al. 2019, in ASP Conf. Ser. 523, Astronomical Data Analysis Software and Systems XXVII, ed. P. J. Teuben et al. (San Francisco, CA: ASP), 143
Narayan, R., & Nityananda, R. 1986, ARA&A, 24, 127
Öberg, K. I., Guzmán, V. V., Walsh, C., et al. 2021, ApJS, 257, 1
Pérez, S., Casassus, S., Baruteau, C., et al. 2019, AJ, 158, 15
Remijan, A., Biggs, A., Cortes, P., et al. 2019, ALMA Technical Handbook, Doc. 7.3 v1.1, https://almascience.nrao.edu/documents-and-tools/cycle7/alma-technical-handbook/view
Rosenfeld, K. A., Andrews, S. M., Hughes, A. M., Wilner, D. J., & Qi, C. 2013, ApJ, 774, 16
Ruffio, J.-B., Macintosh, B., Wang, J. J., et al. 2017, ApJ, 842, 14
Schwab, F. R. 1984, in Indirect Imaging. Measurement and Processing for Indirect Imaging, ed. J. A. Roberts (Cambridge: Cambridge Univ. Press), 333
Schwarz, U. J. 1978, A&A, 65, 345
Semenov, D., Pavlyuchenkov, Y., Henning, T., Wolf, S., & Launhardt, R. 2008, ApJL, 673, L195
Sierra, A., Pérez, L. M., Zhang, K., et al. 2021, ApJS, 257, 14
Speagle, J. S. 2019, arXiv:1909.12313
Tazzari, M., Beaujean, F., & Testi, L. 2018, MNRAS, 476, 4527
Teague, R. 2019, JOSS, 4, 1632
Teague, R. 2020, richteague/keplerian_mask: Initial Release, v1.0, Zenodo, doi:10.5281/zenodo.4321137
Teague, R., Bae, J., & Bergin, E. A. 2019, Natur, 574, 378
Teague, R., Guilloteau, S., Semenov, D., et al. 2016, A&A, 592, A49
Thompson, A. R., Moran, J. M., Swenson, & George, W. 2017, Interferometry and Synthesis in Radio Astronomy (3rd ed.; Berlin: Springer)
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wilner, D. J., & Welch, W. J. 1994, ApJ, 427, 898
Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2013, Tools of Radio Astronomy (Berlin: Springer)
Yen, H.-W., Koch, P. M., Liu, H. B., et al. 2016, ApJ, 832, 204
Zhang, K., Bergin, E. A., Blake, G. A., et al. 2016, ApJL, 818, L16