



SPECTRAL LINE ANALYSIS/MODELING (SLAM) I: PVANALYSIS

YUSUKE ASO ^{1†} & JINSHI SAI (INSA CHOI) ²

¹Korea Astronomy and Space Science Institute, 776 Daedeok-daero, Yuseong-gu, Daejeon 34055, Republic of Korea
²Academia Sinica Institute of Astronomy and Astrophysics, 11F of Astronomy-Mathematics Building, AS/NTU, No.1,
Sec. 4, Roosevelt Rd, Taipei 10617, Taiwan

E-mail: yaso@kasi.re.kr

(Received July 07, 2023; Revised October 12, 2023; Accepted October 29, 2023)

ABSTRACT

Line observations of young stellar objects (YSOs) at (sub)millimeter wavelengths provide essential information of gas kinematics in star and planet forming environments. For Class 0 and I YSOs, identification of Keplerian rotation is of particular interest, because it reveals presence of rotationally-supported disks that are still being embedded in infalling envelopes and enables us to dynamically measure the protostellar mass. We have developed a python library SLAM (Spectral Line Analysis/Modeling) with a primary focus on analyses of emission line data at (sub)millimeter wavelengths. Here, we present an overview of the `pvanalysis` tool from SLAM, which is designed to identify Keplerian rotation of a disk and measure the dynamical mass of a central object using a position-velocity (PV) diagram of emission line data. The advantage of this tool is that it analyzes observational features of given data and thus requires few computational time and parameter assumptions, in contrast to detailed radiative transfer modelings. In this article, we introduce the basic concept and usage of this tool, present an application to observational data, and discuss remaining caveats.

Key words: Circumstellar disks (235) — Circumstellar envelopes (237)

1. INTRODUCTION

The circumstellar disk is a key structure in low mass star formation because planet formation is expected to occur in the disks where material is supported by rotation against the gravity of the central star in a young stellar object (YSO). The rotation balancing with the gravity is called Keplerian rotation and is expressed by the rotational velocity of

$$v_{\text{rot}}(r) = \sqrt{\frac{GM_*}{r}}, \quad (1)$$

where r is the radius, G is the gravitational constant, and M_* is the central stellar mass. Disks around YSOs are expected to exhibit Keplerian rotation when the gravity of the central star is a dominant force determining gas motion. Keplerian disks were identified first around YSOs in the Class II phase, called protoplanetary disks (??). In this evolutionary phase, a YSO simply consists of a central star and a disk, and thus Keplerian rotation of the disks can be identified relatively easily through molecular line observations at millimeter and submillimeter wavelengths.

Disks have been also identified around YSOs in earlier phases, Class 0/I or protostellar phase (e.g., ???). In this evolutionary phase, a YSO, i.e., a protostellar system consists of a central protostar, a disk, an envelope surrounding the disk, and an outflow driven by the rotation in the disk. Because envelopes and outflows have

different kinematics from Keplerian rotation, identification of a disk requires to distinguish the disk from those structures in terms of kinematics (e.g., ?). The outflow can be distinguished relatively easily because its morphology is elongated in the direction perpendicular to the associated disk inclined from the line of sight (e.g., ?). The velocity gradient of an outflow is also expected in a direction perpendicular to that produced by disk rotation (?). In contrast to outflows, envelopes are hard to distinguish from disks because envelopes often show a flattened, disk-like morphology and are rotating like disks (??). The rotation in an envelope is, however, not fast enough to support the material against the gravity of the central star. The material is then infalling in an envelope toward the associated disk. The infalling material is thought to conserve its specific angular momentum (j), producing the rotational velocity inversely proportional to the distance from the central protostar (?):

$$v_{\text{rot}}(r) = \frac{j}{r}. \quad (2)$$

This rotational velocity has a steeper radial profile ($v_{\text{rot}} \propto r^{-1}$) than Keplerian disk ($v_{\text{rot}} \propto r^{-0.5}$) and thus cannot support the material against the gravity. It is, therefore, necessary to verify whether the rotation around a protostar is Keplerian rotation or not through observations in order to identify a protostellar disk. Protostellar disks have recently gained a spotlight as the place for planet formation because more evolved, Class II disks may not be massive enough to form Jupiter-mass

† CORRESPONDING AUTHOR

planets (?). Recent observational works also reported several Class 0/I disks with substructures (e.g., ?????), which are often thought to be a sign of planet formation. The identification of disks is essential to the study of planet formation during the protostellar phase.

Previous works have suggested two directions for identifying protostellar disks. One direction is to construct models, including a disk, and compare them with observational data. Recent sophisticated models of a protostellar system incorporate various calculations, which provide quantities such as a temperature distribution based on a given stellar luminosity and a given density field, non-LTE population of each quantum energy level of a given molecule, the intensity at each position on the plane of the sky through radiative transfer (e.g., ?). Finding a model that reproduces a given observational data set the best can cost more computational time as more realistic conditions and more free parameters are considered. It is no longer very rare for observational studies to include parameter search that takes a month or longer time scale. The other direction is to analyze observational data in detail focusing on observable features expected in specific physical structures, such as a disk. Such analytic estimates have advantages of the lower computational cost and simpler processes than model fitting including radiative transfer with a number of free parameters. Analytic methods also enable us to focus on a specific quantity, such as the central stellar mass, without complicated parameter degeneracy. The python package of `pvanalysis` implemented in the `SLAM` (Spectral Line Analysis/Modeling) code (?)¹ has been developed to identify protostellar disks embedded in a protostellar envelope. In this article, we summarize its methodologies and a basic usage, and introduce an application to observational data of this tool.

2. METHODOLOGIES

The `pvanalysis` tool performs a two-steps analysis using a given position-velocity (PV) diagram made by cutting a 3D cube (2D space projected on the plane of the sky plus 1D line-of-sight velocity) obtained from emission line observations along the major axis of the expected disk. Since rotation causes a velocity gradient along the major axis, PV diagrams along the major axis show the emission distribution as a function of the rotational velocity v and the offset or radius from the central protostellar position r . The first step of this tool is to extract data points from the PV diagram that represent the relation between the rotational velocity and the radius. The second step is to fit these data points with a power-law function. These processes are visualized in Figure 1. We describe details of each step in the following subsections.

2.1. EDGE AND RIDGE

The first step, extracting data points (r_i, v_i) from a PV diagram, is performed through two independent meth-

ods. One method uses the emission ridge in 1D profiles along the positional axis (the velocity axis) at a given velocity (position). This ridge method is demonstrated in previous observational studies (????) with discussion about its potential problems (???), as well as in a study using synthetic observations of a magnetohydrostatic (MHD) simulation of protostellar evolution (?). The other method uses the emission outer edge in the 1D profiles. This edge method was developed using an MHD simulation of protostellar evolution (?) and is used in previous observational studies as well (e.g., ?). The advantage and disadvantage of each method are controversial. For example, ? reported that, depending on the spatial resolution, the ridge method can underestimate the central stellar mass by $\sim 30\%$, while the edge method can overestimate it by a factor of ~ 2 .

The edge and ridge radii are obtained in the 1D profile along the positional axis at the given velocity ($xcut$); those radii are obtained at every channel (Figure 1). The edge radius is defined as the outermost position with a given threshold level (e.g., 5σ) of emission. The uncertainty of the edge radius is calculated as the absolute value of the intensity gradient at the edge radius multiplied by the noise level of the PV diagram. The intensity gradient is calculated by interpolating the 1D profile. The ridge radius is defined as the intensity-weighted mean position in the 1D profile, or the center derived by Gaussian fitting. The uncertainty of the ridge radius is calculated from the noise level of the PV diagram through the error propagation in the case of the mean position, while it is provided by the Gaussian fitting in the case of the Gaussian center. Similarly the edge and ridge velocities are obtained in the 1D profile along the velocity axis at a given position ($vcut$; Figure 1). In the $vcut$ case, those velocities are obtained with a positional separation of half of the beam major-axis, rather than one pixel, to avoid sampling more than the Nyquist rate. The oversampling causes to put unnecessarily high weights on the $vcut$ points, relative to the weights on $xcut$ points, in the χ^2 fitting with a power-law function.

Then, obtained ridge and edge points are passed to filters so that data points better representing the rotation curve feature can be fitted with a power-law function in a latter fitting stage. The filtering conditions are illustrated in Figure 2. In each of edge and ridge method, $xcut$ points at velocities higher than a “middle” velocity and $vcut$ points at radii outer than a “middle” radius are combined as the final pairs of (r_i, v_i) (top of Figure 2). The middle velocity and radius are determined from the closest pair of points between the $xcut$ and $vcut$, as explained in ? in more detail. In addition, in $xcut$ ($vcut$), the edge/ridge radii (velocities) can be removed if the radii (velocities) are at lower velocities (smaller radii) than the velocity (radius) with the maximum edge/ridge radius (velocity) (middle of Figure 2). Furthermore, the edge/ridge radii or velocities can be removed if those are located in the opposite quadrant (bottom of Figure 2). Removing points is done separately between the redshifted and blueshifted velocities

¹This tool is public on GitHub: <https://github.com/jinshisai/SLAM>.

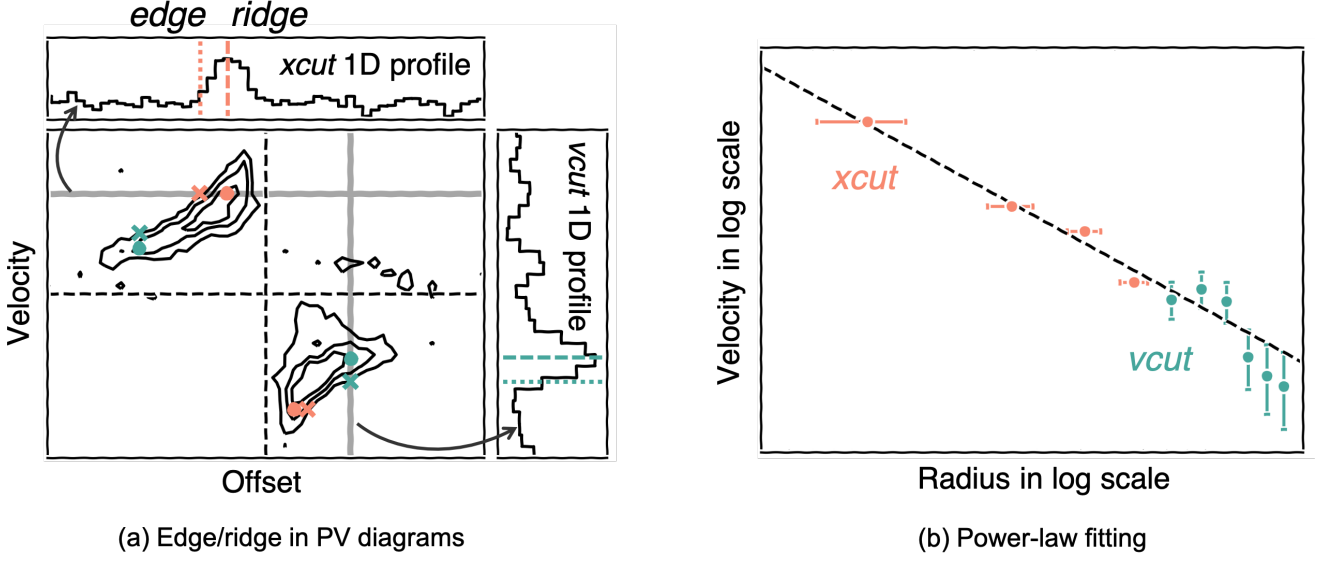


Figure 1. Two steps of pvanalysis. (a) Edge/ridge radii and velocities are extracted from the PV diagram, respectively, and then (b) the obtained edge/ridge points are fitted with a power-law function. The left panel illustrates a PV diagram and edge/ridge points derived from *xcut* and *vcut* 1D profiles of the PV diagram. Pink and green markers indicate *xcut* and *vcut* data points, respectively. Crosses and circles in the PV diagram denote edge and ridge points, respectively. Dotted and dashed lines in the *xcut* and *vcut* 1D profiles also indicate edge and ridge points, respectively. The right panel shows a figure of the fitting of a power-law function to edge/ridge points.

(or positive and negative offsets).

2.2. POWER-LAW FITTING

The pairs of (r_i, v_i) obtained in the edge and ridge methods are separately fitted with a single- or double-power function as follows:

$$v = \text{sgn}(r) v_b \left(\frac{|r|}{r_b} \right)^{-p} + v_{\text{sys}},$$

$$p = p_{\text{in}} \quad \text{if } |r| < r_b \quad \text{else } p_{\text{in}} + dp, \quad (3)$$

where v_b is the velocity at the radius of r_b , p_{in} is the power-law index, dp is the difference of the index between the inner and outer parts, and v_{sys} is the systemic velocity. The sign of the first term on the left hand side of Equation (3) means that emission tracing rotation appears in the first and third quadrants in the PV diagram. When the rotation appears in the second and fourth quadrants, this sign is inverted. dp is limited to be ≥ 0 to produce a steeper or the same index outside r_b , such as the inner Keplerian rotation versus the outer rotation where the angular momentum is conserved. $dp = 0$ means the single power-law fitting. Equation (3) provides a function of radius, defined here as $V_{\text{fit}}(r)$. When $dp \geq 0$, the inverse function of velocity $R_{\text{fit}}(v)$ can also be defined. Using these functions, χ^2 to be minimized is defined as

$$\chi^2 = \sum_{i \text{ for } x\text{cut}} (r_{i,\text{obs}} - R_{\text{fit}}(v_{i,\text{obs}})) / \sigma_{i,\text{obs}}^2$$

$$+ \sum_{j \text{ for } v\text{cut}} (v_{j,\text{obs}} - V_{\text{fit}}(r_{j,\text{obs}}))^2 / \sigma_{j,\text{obs}}^2, \quad (4)$$

where σ is the uncertainty of each radius/velocity defined in Section 2.1: The edge uncertainty is calculated through the emission gradient, the mean-ridge uncertainty is calculated through the error propagation, and the Gaussian-ridge uncertainty is calculated through the 1D Gaussian fitting. If the *vcut* points are sampled on every pixel, the sum of *vcut* becomes more dominant in Equation (4) with smaller pixels by summing multiple points within the half-beam, which have effectively the same information. Then, the fitting process will search for a solution that only minimizes the sum of *vcut*, ignoring the sum of *xcut*. To avoid such imbalance, the *vcut*-points are sampled at the half-beam interval (Section 2.1).

The χ^2 is minimized by the Markov chain Monte Carlo method using the open python package `emcee` (?) widely used in modelings of YSOs. We adopt this method because it searches for parameters efficiently around the best parameters, and it shows acceptable parameter ranges and degeneracy among parameters through the posterior distribution. By default, the numbers of walkers per parameter, burn-in steps, and steps after the burn-in are 16, 2000, and 2000, respectively. The 50 percentile of each parameter is adopted as the best-fit value. The uncertainty of each parameter is calculated as the 16 and 84 percentiles. These numbers will be justified in the actual application to observational data in Section 4. The search ranges for r_b and v_b are limited between the maximum and minimum of the edge/ridge radii and velocities, respectively. The ranges for p_{in} and dp are limited to $[0.01, 10]$ and $[0, 10]$, respectively. These ranges are wide enough because the indices expected from physics are around 0.5 (Keplerian

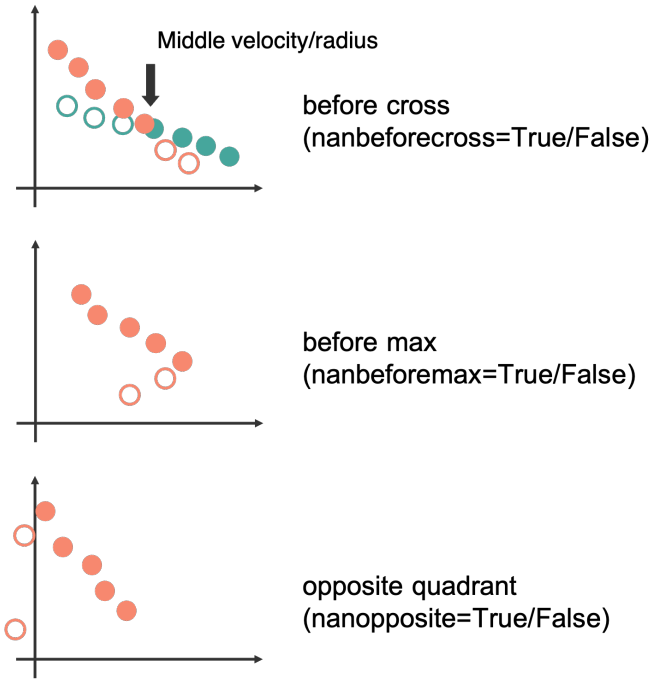


Figure 2. Filtering conditions so that edge/ridge points better represent the rotation curve. Filled and open circles indicate data points that will be used for and removed from the rotation curve fitting, respectively. The `nanbeforecross` parameter removes the `vcut`-points at smaller radii and the `xcut`-points at lower velocities than the closest pair to obtain a one-to-one relation between the radius and the velocity. The `nanbeforemax` parameter removes the `vcut`(`xcut`)-points at the radii smaller (velocities lower) than the radius (velocity) of the maximum velocity (radius) to achieve the spin-up relation between the radius and the velocity. The `nanopposite` parameter removes the points on the unexpected quadrant. These three filters are described in Section 2.1 in more detail.

rotation) or 1.0 (conservation of angular momentum). The range for v_{sys} is limited to $[-1 \text{ km s}^{-1}, 1 \text{ km s}^{-1}]$ relative to a given systemic velocity. This range is much wider than typical velocity resolutions for YSO studies, and thus v_{sys} can be estimated within this range before using our tool. No other prior is set for the MCMC fitting. In the case of single power-law fitting, v_b is fixed at the logarithmic center of the search range of v_b : When the range is $[v_{\text{min}}, v_{\text{max}}]$, the logarithmic center is $\sqrt{v_{\text{min}}v_{\text{max}}}$.

3. USAGE AND PARAMETERS

In this section, we describe actual usage and parameters of the `pvanalysis` tool. In addition to the detailed descriptions in this section, an example script can be found in our GitHub repository².

3.1. INPUTS

The each analysis step described in Section 2 is provided as a method for a python class, `PVAnalysis`, in this tool.

The instantiation requires five parameters, including the input data:

```
impv = PVAnalysis(fitsfile, rms, vsys, dist, incl)
```

`fitsfile` — Path to the input fits file of a PV diagram.

`rms` — The root-mean-square (rms) noise level of the PV diagram.

`vsys` — The systemic velocity of the object in the unit of km s^{-1} . When the systemic velocity v_{sys} is a free parameter, this value is the center of the range over which the free parameter is searched.

`dist` — The distance to the object in the unit of pc to convert the unit of radius from arcsecond to au.

`incl` — The inclination angle of the disk in the unit of degree, which is used to calculate the central stellar mass. If not given, an edge-on configuration (i.e., an inclination angle of 90°) is assumed.

The input of observational data must be a fits file of a PV diagram. The input fits file is supposed to be two or three dimensions: it should have offset and velocity/frequency axes, or have a one-dimensional stokes axis (i.e., with only stokes I and no polarization) in addition to the offset and velocity/frequency axes. It is assumed that the offset axis is the first axis, and the stokes axis is the third axis if it is present. The second axis can be either velocity or frequency. If the second axis is given as velocity, it must be in the unit of m s^{-1} , and then it is converted in the unit of km s^{-1} . If the second axis is given as frequency, it is automatically converted to velocity based on the line rest frequency provided in the fits header.

The first phase of the analysis where edge and ridge points are determined (with a method of `get_edgeridge`) requires inputs to limit which data of a PV diagram are used. For the simplest use, one can only give the output file name and the `thr` parameter:

```
impv.get_edgeridge(outname='filename', thr=5)
```

`outname` — Header name of output files. The output file will be, for example, `outname.ridge.dat`.

`thr` — The pixels where the intensity is higher than this threshold are used for determining the ridge points. The outermost radius (the highest velocity) with this value is the edge radius (velocity). The unit is the given rms noise level. Defaults to 5.

Then, edge/ridge radii and velocities will be calculated using the whole radial and velocity ranges of a given PV diagram following the method and criteria described in Section 2. One might want to calculate edge/ridge points only within limited radial and/or velocity ranges. The following parameters can be used for such a case in the method of `get_edgeridge`:

`xlim` — When `xlim = (a, b, c, d)`, the edge and ridge radii are removed unless they are between a and b or c and d .

`vlim` — Same as `xlim` but for the velocities.

`Mlim` — When `Mlim = (a, b)`, the edge and ridge points (r_i, v_i) are removed unless $r_i V_i^2 / G$ is between a and b .

²<https://github.com/jinshisai/SLAM>

use_position — Whether this process uses the edge and ridge radii, i.e., `xcut`.

use_velocity — Whether this process uses the edge and ridge velocity, i.e., `vcut`.

These inputs allow to remove points that appear suffering from any effect that this tool does not suppose, such as the resolving-out effect in interferometric observations and a low signal-to-noise ratio (S/N). In addition, some of the filtering processes could be turned-on and off through the following parameters in the method of `get_edgeridge`. All of which are True by default.

nanbeforecross — This switches on the removal of lower velocities and smaller radii to combine the `xcut` and `vcut` points.

nanbeforemax — This switches on the removal of the edge and ridge radii (velocities) at velocities (radii) lower (smaller) than the velocity (radius) with the maximum radius (velocity).

nanopposite — This switches on the removal of the points on the unexpected quadrants.

To avoid that only few points that has S/N much higher than others control the power-law function fitting, the inputs of `minabserr` and `minrelerr` are introduced in the method of `get_edgeridge`:

minabserr — This sets the minimum absolute error for each edge or ridge point. The unit is the major-axis of the observational beam for the radius, while it is the channel width for the velocity. When the absolute error bar for an edge or ridge point is smaller than this value, it is replaced with this value. Defaults to 0.1.

minrelerr — This sets the minimum relative error for each edge or ridge point. When the relative error bar for an edge or ridge point is smaller than this value, it is replaced with this value. Defaults to 0.01.

Ridge points can be derived through not only the intensity-weighted mean but also the Gaussian fitting. They can be switched using the following parameter in the method of `get_edgeridge`:

ridgemode — ‘mean’ or ‘gauss’. The method for determining the ridge points: the intensity-weighted mean radius (velocity) or the Gaussian-center radius (velocity). Defaults to ‘mean’.

The power-law fitting part (`fit_edgeridge`) requires inputs to define the model function.

include_vsyst — Whether the systemic velocity is a free parameter.

include_dp — Whether the difference between the inner and outer power-law indices is a free parameter.

include_pin — Whether the inner power-law index is a free parameter.

fixed_pin — The inner power-law index is fixed to this value when it is not a free parameter.

fixed_dp — The difference between the inner and outer power-law indices is fixed to this value when it is not a free parameter.

For example, the following command means that the

model function is the Keplerian rotation, which has only one free parameter:

```
impv.fit_edgeridge(include_vsyst=False,
                  include_dp=False,
                  include_pin=False,
                  fixed_pin=0.5,
                  fixed_dp=0,
                  calc_evidence=False).
```

The last input, `calc_evidence`, is an optional parameter and allows you to calculate a model evidence using the `dynesty` package (?) when it is set to `True`. A model evidence is the integral of a likelihood function over the parameter space. One can calculate the Bayes factor by taking a evidence ratio between two models, which would be useful for discussing the goodness of fits among different models (e.g., `include_pin=True` versus `False` in Section 4).

In addition, the `PVAnalysis` class provides methods for writing the obtained edge and ridge points to a text file (`write_edgeridge(outname='filename')`), outputting the fitting results in the terminal window (`output_fitresult()`), and plotting PV diagrams with the ridge/edge points and the best-fit curves (`plot_fitresult()`).

3.2. OUTPUTS

The main output is the best-fit parameters, r_b , v_b , p_{in} , dp , and v_{syst} with their error bars; the error bar is zero when the parameter is not free. In addition, ranges of radius and velocity are calculated using the best-fit model function, as summarized in Figure 3. With only `xcut`, the highest velocity, v_{high} , is the highest velocity among the velocities where edge/ridge radii are obtained. Then, the innermost radius r_{in} is calculated from Equation (3) as the radius at $v = v_{high}$, i.e., $r_{in} = R_{fit}(v_{high})$. The pair of (r_{out}, v_{low}) is similarly defined in the case with only `xcut`. With only `vcut`, the innermost radius, r_{in} , is the smallest radius among the radii where edge/ridge velocities are obtained. Then, the highest velocity v_{high} is calculated from Equation (3) as the velocity at $r = r_{in}$, i.e., $v_{high} = V_{fit}(r_{in})$. The pair of (r_{out}, v_{low}) is similarly defined. With the combination of `xcut` and `vcut`, (r_{in}, v_{high}) is defined as in the case with only `xcut`, while (r_{out}, v_{low}) is defined as in the case with only `vcut`. This is inner, higher-velocity points in the `xcut` set and outer, lower-velocity points in the `vcut` set are supposed to be combined to obtain the final edge/ridge points (Section 2.1) in the case with both `xcut` and `vcut`. These definitions of (r_{in}, v_{high}) and (r_{out}, v_{low}) for different cases are illustrated in Figure 3.

From the pair of (r_b, v_b) , the central stellar mass can be calculated as $M_* = v_b^2 r_b / G / \sin^2 i$, where G is the gravitational constant and i is the inclination angle. Estimation of the inclination angle is out of the scope of this tool; the inclination angle is often estimated from the aspect ratio of the associated (sub)mm continuum emission or independent modeling of the associated outflow. If the power-law index p_{in} is significantly different from the Keplerian index 0.5, the estimated M_*

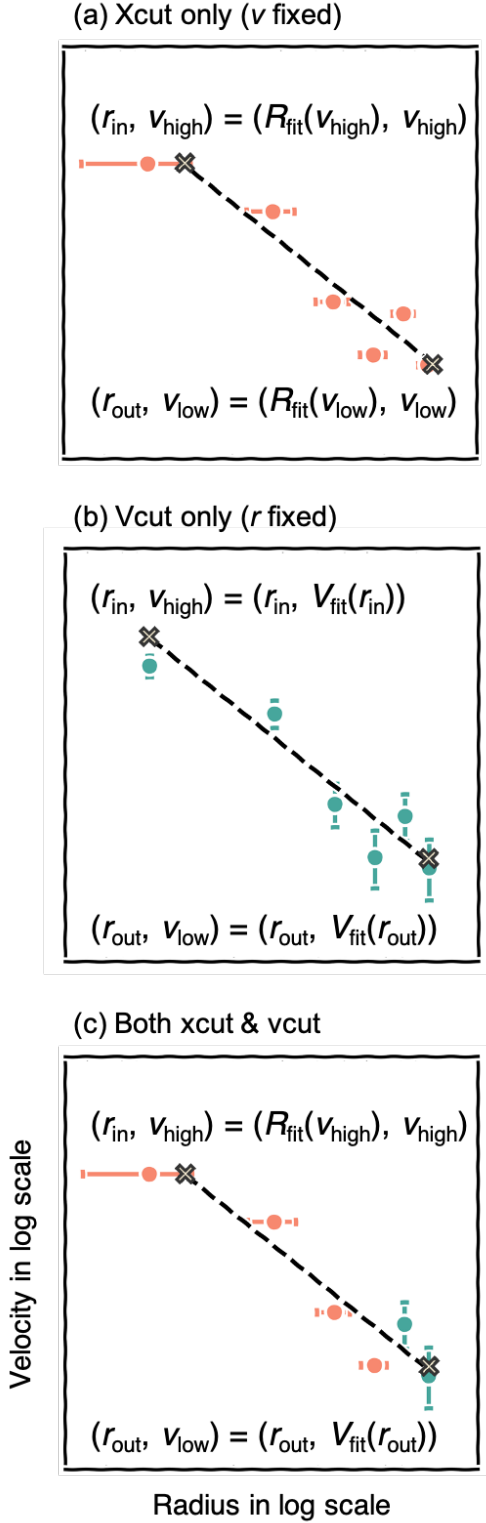


Figure 3. Illustration of definitions of $(r_{\text{in}}, v_{\text{high}})$ and $(r_{\text{out}}, v_{\text{low}})$ in cases of (a) xcut-only, (b) vcut-only, and (c) the combination of xcut and vcut. Dashed lines indicate the best-fitted power-law function and crosses denote points where $(r_{\text{in}}, v_{\text{high}})$ and $(r_{\text{out}}, v_{\text{low}})$ are defined.

does not have a physical meaning of the central stellar mass. This tool also provides two more masses:

$$M_{\text{in}} = v_{\text{high}}^2 r_{\text{in}} / G / \sin^2 i \text{ and } M_{\text{out}} = v_{\text{low}}^2 r_{\text{out}} / G / \sin^2 i.$$

This tool also outputs two text files and four figure files. Two text files are produced by the method of `write_edgeridge()`. One of the two text files shows the pairs of ridge (r_i, v_i) with their error bars. The error bar of r_i (v_i) is zero for the vcut (xcut) points. The other text file similarly shows the edge pairs. Two figures show the corner plots produced by the MCMC fitting, which are produced by the method of `fit_edgeridge()`. One figure is for the ridge fitting, and the other is for the edge fitting. The remaining two figure are produced by the method of `plot_fitresult()` and shows a PV diagram with the edge and ridge points and the best-fit curves; one figure for the linear scale and one figure for the logarithmic scale. The logarithmic PV diagram is produced by calculating the mean between the first- and third-quadrants (or second- and forth-quadrants) of the linear PV diagram. When the systemic velocity is a free parameter, the mean of the edge- and ridge-systemic velocities is used to calculate the logarithmic PV diagram.

4. EXAMPLES WITH OBSERVATIONAL DATA

Figure 4 shows the result of our tool using an actual observational data set³ of a molecular line in a protostellar system that consists of a disk and an envelope. Figure 4(a) shows a PV diagram along the major axis of the disk. The abscissa is the distance from the central protostellar position, i.e., radius in the unit of au (distance is 140 pc), while the ordinate is the velocity relative to the systemic velocity (6.4 km s^{-1}). This diagram shows a tendency that the emission is closer to the center at higher velocities. Because PV diagrams along the disk major axis are used to investigate the relation between the rotational velocity and the radius, this tendency can be interpreted as a spin-up rotation i.e., the rotational velocity is higher at a position closer to the central protostellar position. The points in Figure 4 are the edge/ridge xcut/vcut points; the triangles and circles are the edge and ridge points, respectively, while the red/blue and pink/cyan points denote the xcut and vcut points, respectively. These points successfully trace the spin-up tendency in the PV diagram.

Figures 5 and 6 show the edge and ridge points plotted in 1D intensity profiles with which the edge or ridge points are obtained. The xcut and vcut points in these figures are before combined. The panels without the edge or ridge point (blue or red dashed line) have S/N ratios not enough to obtain these points. The offsets of the vcut profiles are sampled at the Nyquist rate. The ridge points trace the peak offset and velocity of each 1D intensity profile, and the edge points trace an outer position and velocity of the profile, which demonstrate that the edge and ridge points are well-defined.

To describe the radial profile of rotational velocity, a broken power-law function is adopted that fixes $p_{\text{in}} =$

³The project codes of the data used in this demonstration are 2013.1.01086.S and 2015.1.01415.S. The targeted protostar has a name of TMC-1A. The angular resolution is $0''.12 \times 0''.08$ (P.A.= 35°). The velocity resolution is 0.35 km s^{-1} . The PV diagram is produced with a cut at P.A.= 75° .

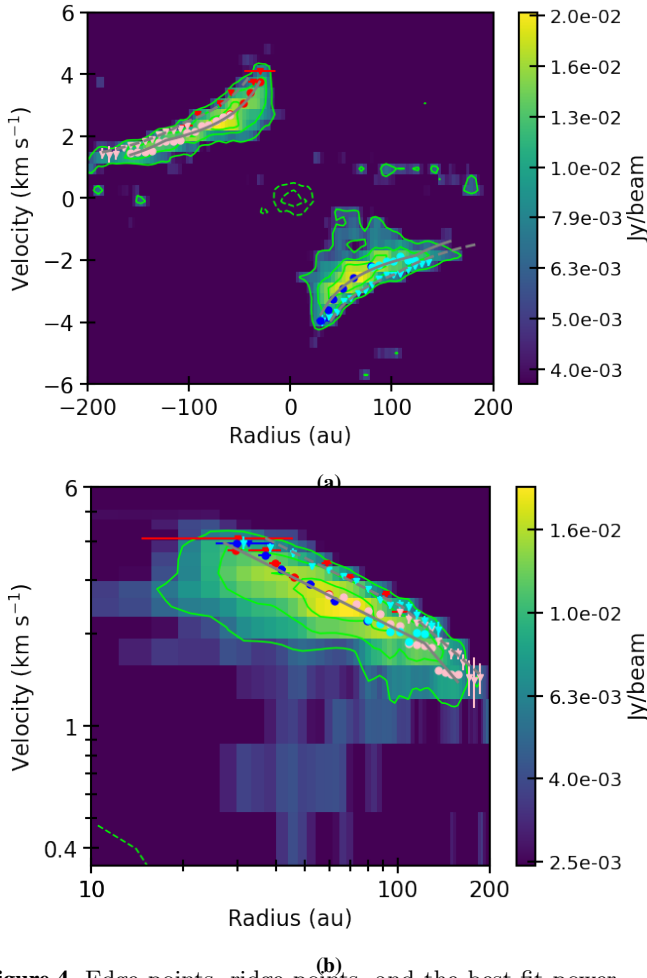


Figure 4. Edge points, ridge points, and the best-fit power-law function overlaid in a position-velocity diagram on the (a) linear and (b) logarithmic scales. The velocity is relative to the systemic velocity 6.4 km s^{-1} . The triangles and circles are the edge and ridge points, respectively. The red and blue points are the xc points, while the pink and cyan points are vc points. The xc points have error bars in the positional direction. The vc points have error bars in the velocity direction. The dashed and solid curves show the best-fit power-law functions for the edge and ridge points, respectively. The contour levels are ± 3 , ± 6 , and $\pm 9\sigma$, where 1σ is $1.7 \text{ mJy beam}^{-1}$.

0.5 (Keplerian rotation) and $v_{\text{sys}} = 6.4 \text{ km s}^{-1}$, while varying v_b , r_b , and dp in the MCMC fitting. The gray curves in Figure 4 show the best-fit power-law functions; the dashed and solid curves show those for the edge and ridge points, respectively.

Figure 7 shows the mean (panel a) and the autocorrelation (panel b) of the three parameters sampled in the MCMC fitting as a function of the step. This figure shows the results for the ridge points; the results for the edge points similarly show the following results. In Figure 7(a), the denser blue lines show the mean values among the 16 chains, while the cyan lines show the mean \pm the standard deviation among the chains. Different lines with the same color show three independent results to inspect the effect of any random number in

the MCMC fitting. In Figure 7(b), the denser blue lines show the autocorrelation of the parameter offset from the mean among the steps. This autocorrelation is calculated using neighboring 5000 steps in Figure 7(a) and normalized by the maximum value at the no delay point. The cyan lines show the mean \pm the standard deviation of the autocorrelation among the chains. The meaning of different lines with the same color is the same as in Figure 7(a). These figures show that the sampled parameters are converged after 2000 steps, with the adopted number of walkers per parameter 16. In other words, the parameters sampled after 2000 steps are not correlated with the initial values. This supports the numbers of walkers and the burn-in steps (Section 2.2).

Figure 8 shows corner plots of the posterior distribution of the three free parameters obtained in the MCMC fitting, using 2000 steps after the burn-in. The posterior distribution is well sampled with the 2000 steps. The blue lines and points denote the parameter set that provides the smallest χ^2 . This parameter set is close to the 50 percentile parameter set, indicating that the 50 percentile can be adopted as the best-fit parameter set. The best-fit three parameters are listed in Table 1. The evidence of the posterior distribution was calculated through *dynesty* to be $\sim 2.3 \times 10^{-15}$ for the edge points and $\sim 1.1 \times 10^{-41}$ for the ridge points. The parameter dp means that the outer power-law index is ~ 0.9 in the edge method and ~ 1.2 in the ridge method, which may suggest a rotation with a constant specific angular momentum ($v_{\text{rot}} \propto r^{-1}$). Because the inner power-law index is assumed to be the Keplerian-rotation index, r_b is interpreted as the boundary radius between a Keplerian disk and a surrounding envelope: $r_b \sim 100\text{--}120 \text{ au}$. This radius is consistent with the disk radius reported by ? with multiple analyses and modeling to the same target. The difference of r_b between the edge and ridge methods is > 2 times larger than the error bar of each r_b . Similarly, the difference of v_b between the edge and ridge methods is much larger than the error bar of each v_b .

Using the best-fit power-law function, the two masses M_{in} and M_{out} are calculated; their error bars are also calculated through error propagation (Table 2). Previous works estimated the inclination angle of this system to be $i = 51^\circ$ (?) and 53° (?). Hence, we adopt $i = 50^\circ$ for this example calculation. M_b is the same as M_{in} in this fitting because the inner power-law index p_{in} is fixed at 0.5 (Keplerian rotation). Hence, $M_{\text{in}} (= M_b)$ is the central stellar mass estimated from this fitting. The edge mass and ridge mass are thought to provide an upper limit and a lower limit of the central stellar mass, respectively, i.e., $M_* \sim 0.8\text{--}1.2 M_\odot$. The mass before the inclination correction, $M_* \sin^2 i = 0.5\text{--}0.7$, is consistent with the one reported for the same target by ?, $M_* \sin^2 i = 0.56 \pm 0.05 M_\odot$. The derived error bars for M_{in} and M_{out} are a few times smaller than the difference between the edge and ridge masses. This suggests that the uncertainty of the stellar mass estimation is dominated by the difference between the edge and ridge

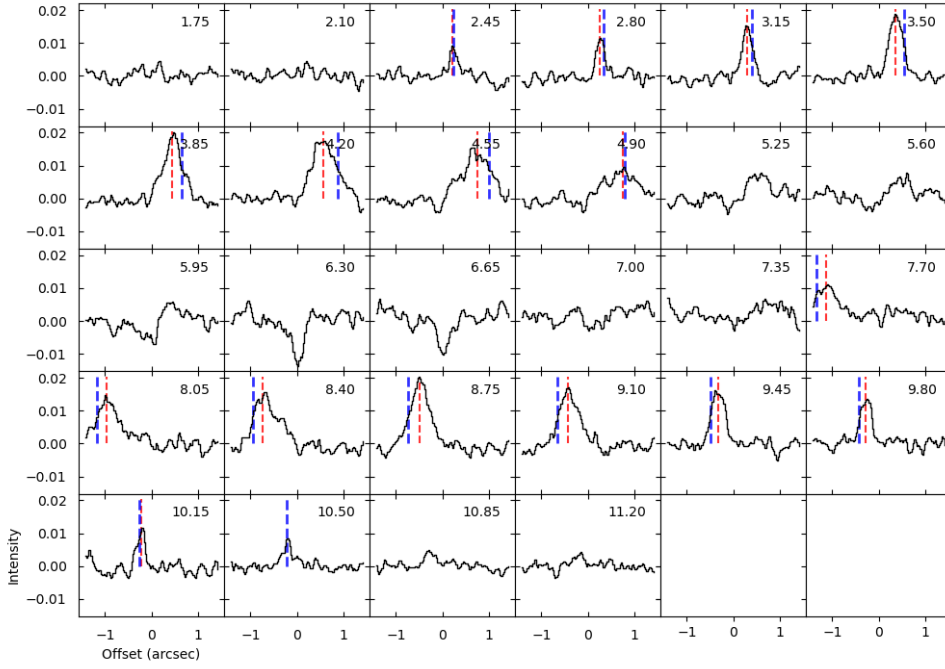


Figure 5. The x_{cut} points plotted in 1D intensity profiles. The blue and red points are the edge and ridge points, respectively. Each panel shows a 1D intensity profile where the abscissa is the offset from the central position. The velocity (not relative to the systemic velocity but absolute) of each profile is denoted on the upper right corner of each panel in the unit of km s^{-1} .

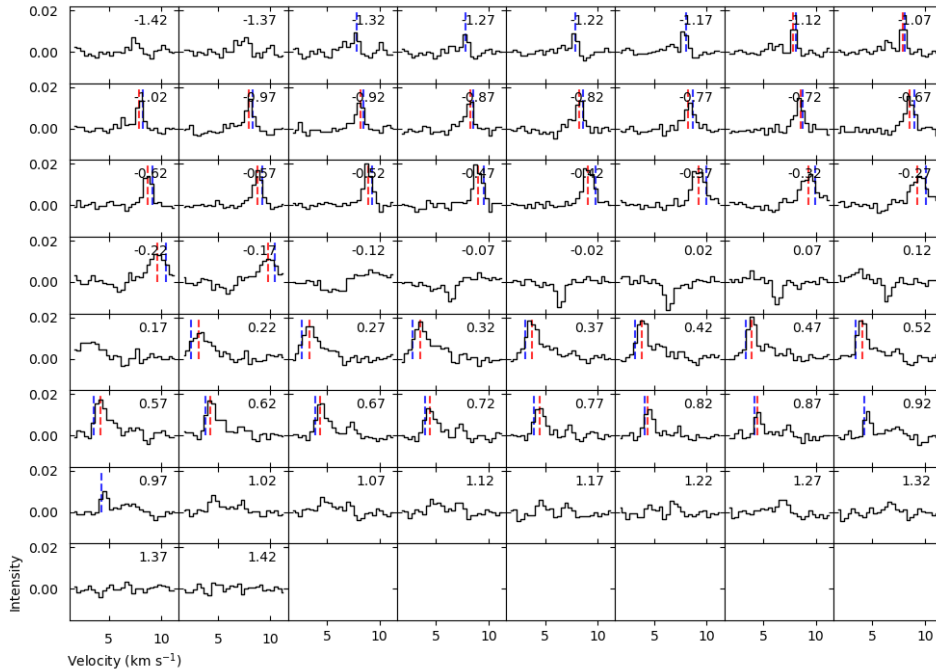


Figure 6. The v_{cut} points plotted in 1D intensity profiles. The blue and red points are the edge and ridge points, respectively. Each panel shows a 1D intensity profile where the abscissa is the velocity (not relative to the systemic velocity but absolute). The positional offset of each profile is denoted on the upper left corner of each panel in the unit of arcsecond.

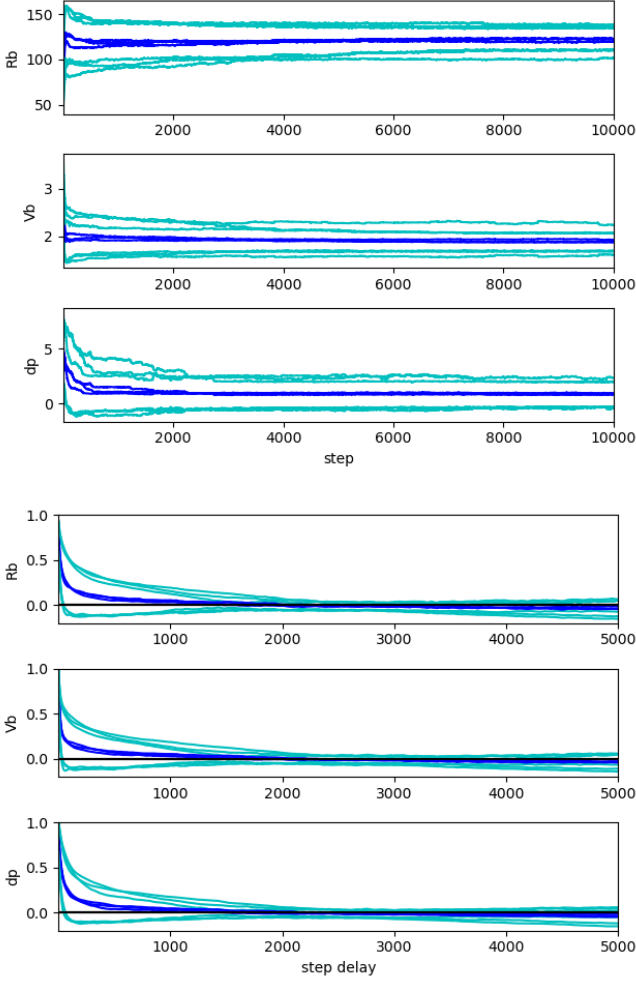


Figure 7. The three parameter (panel a) and their autocorrelation (panel b) obtained in the MCMC fitting. The denser blue lines show the mean, while the cyan lines show the mean \pm the standard deviation among 16 chains. Different lines with the same color show three independent results to inspect the impact of any random number in the MCMC fitting. The autocorrelation is calculated using neighboring 5000 steps and normalized by the maximum value at the no delay point.

Table 1. The best-fit parameters in Section 4

	r_b (au)	v_b (km s $^{-1}$)	dp
Edge	103.4 ± 8.9	2.45 ± 0.12	0.37 ± 0.11
Ridge	123.0 ± 2.3	1.86 ± 0.02	0.67 ± 0.11

v_b is not corrected by the inclination factor of $\sin i$.

methods. Because dp is significantly different from zero, M_{out} no longer has a meaning of the central stellar mass. The best-fit power-law function is also used to calculate the r and v ranges of the edge and ridge points (Table 2). Using this range, M_{out} can be converted to a specific angular momentum: $j = \sqrt{GM_{\text{out}}r_{\text{out}}} \sim 360$ and $290 \text{ km s}^{-1} \text{ au}$ in the edge and ridge methods, respectively.

As another example, a case with `include_pin=True`

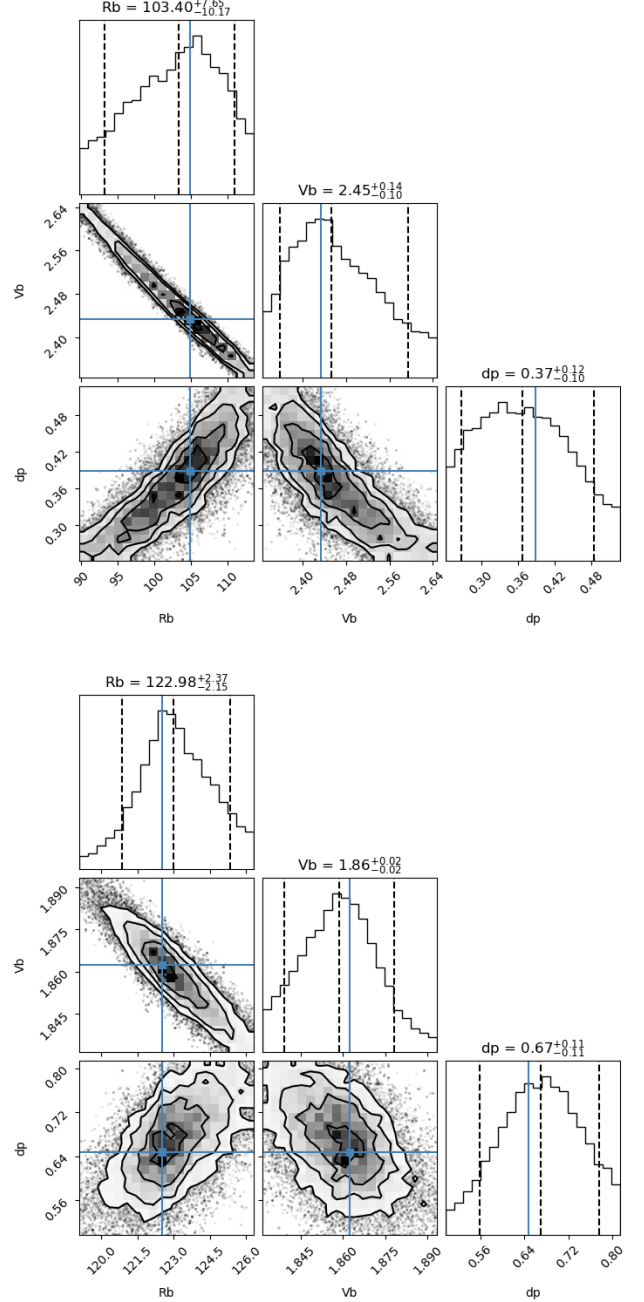


Figure 8. Corner plots of the three parameters, r_b in the unit of au, v_b in the unit of km s^{-1} , and dp , searched in the MCMC fitting with `emcee` to the (a) edge and (b) ridge points. The blue lines and points show the parameter set that provides the smallest χ^2 . The dashed lines show the 16, 50, and 84 percentiles of the posterior distribution of each parameter.

is performed, i.e., the inner power-law index is not fixed to the Keplerian law. The best-fit parameters of the edge method are $(r_b, v_b, p_{\text{in}}, dp) = (83.3 \pm 9.8 \text{ au}, 2.82 \pm 0.18 \text{ km s}^{-1}, 0.41 \pm 0.04, 0.37 \pm 0.06)$. Those of the ridge method are $(123.9 \pm 2.7 \text{ au}, 1.83 \pm 0.03 \text{ km s}^{-1}, 0.53 \pm 0.01, 0.58 \pm 0.12)$. The masses are calculated to be $(M_{\text{in}}, M_b, M_{\text{out}}) = (1.08 \pm 0.24 M_{\odot}, 1.27 \pm$

Table 2. Values calculated from the best-fit power-law function

	r range (au)	v range (km s ⁻¹)	M_{in} (= M_b ; M_{\odot})	M_{out} (M_{\odot})
Edge	37.0 – 185.5	1.48 – 4.10	1.20±0.15	0.78±0.17
Ridge	27.2 – 157.5	1.39 – 3.95	0.82±0.02	0.59±0.04

$M_{\text{in}} = M_b$ because p_{in} is fixed at 0.5 (Keplerian rotation). M is inclination-corrected, while the v range is not inclination-corrected. Note M_{out} no longer has a meaning of the central stellar mass because $dp > 0$.

0.22 M_{\odot} , 0.81±0.21 M_{\odot}) in the edge method, while they are (0.87±0.04 M_{\odot} , 0.79±0.03 M_{\odot} , 0.59±0.05 M_{\odot}) in the ridge method. When the error bars are included, the estimated r_b , v_b , and dp are overall close to those estimated in the example with the fixed p_{in} . The estimated inner index p_{in} is close to the Keplerian law, 0.5, in both methods, supporting $p_{\text{in}} = 0.5$ fixed in the previous example. In addition, the evidence is $\sim 0.8 \times 10^{-15}$ for the edge points and $\sim 0.4 \times 10^{-41}$ for the ridge points. These are ~ 3 times lower than those for the p_{in} -fixed case, implying that the p_{in} -fixed model could be better than the p_{in} -varied model. When p_{in} is not exact 0.5, M_{in} and M_b are not exactly the same. In this case, a middle value between these masses will provide a reasonable estimate of the stellar mass ($M_* \sim 1.3$ and 0.8 M_{\odot} in the edge and ridge methods, respectively), which is consistent with the example with the fixed p_{in} .

The fitting in the examples above used fixed values for the systemic velocity, the distance to the target, and the inclination angle. Among them, the systemic velocity can be incorporated in the fitting, if necessary, by setting `include_vsyst=True`, while the distance and inclination angle need to be estimated outside our tool. When the uncertainties of the distance d and the inclination angle i (in the unit of radian) are Δd and Δi , respectively, these cause additional uncertainties for estimated radius r , velocity v , mass M , and specific angular momentum j as $\Delta r = r\Delta d/d$, $\Delta v = v\Delta i/\tan i$, $\Delta M = M\sqrt{4(\Delta i)^2/\tan^2 i + (\Delta d)^2/d^2}$, and $\Delta j = j\sqrt{(\Delta i)^2/\tan^2 i + (\Delta d)^2/d^2}$. For example, when $\Delta d/d = 0.05$, $\Delta i = 3^\circ$, and $i = 50^\circ$, $\Delta r/r = 0.05$, $\Delta v/v = 0.04$, $\Delta M/M = 0.09$, and $\Delta j/j = 0.06$.

5. DISCUSSION

The example in the previous section demonstrates that our tool can provide a disk radius and a central stellar mass separately with the edge and ridge methods. The best-fit function including the outer power-law index also enables to calculate a specific angular momentum. The MCMC fitting provides a statistical error of each free parameter and the central stellar mass. In addition to the statistical error, the obtained disk radius and central stellar mass have uncertainties due to the difference between the edge and ridge methods. In the example presented in Section 4, the difference between the estimated stellar masses by the two methods is larger than the statistical errors in the edge and ridge methods by a factor of ~ 20 and ~ 3 , respectively. Thus, our tool adopts the edge and ridge values as upper and lower limits. Although previous works suggest how much the two methods over- or underestimates the central stellar

mass (e.g., ??), this is beyond the scope of our tool at this moment.

5.1. CAVEATS

Lastly, we note caveats in the usage of our tool. First of all, how accurately the method in our tool estimates physical quantities (M_* , R_{disk} , p , j , etc.) is still under heated debate, with other questions such as which of edge and ridge methods is better, whether these analytical methods are sufficient for a certain purpose or making a model with radiative transfer is necessary. The references in this report help to understand the current situation. In other words, it is currently difficult to evaluate any systematic error of the physical quantities. Our tool thus provides only a statistical error that originates in the given observational noise level. The statistical error is often smaller than differences due to choice of methods and parameters, such as edge versus ridge and M_{in} versus M_b . Hence, we emphasize that each free parameter could have uncertainties larger than the statistical error. The free parameters are selected on a case-by-case basis and could contribute to the uncertainty. For example, when p_{in} is not 0.5 but close to 0.5, the fitting provides different M_{in} and M_b . In comparison, the two masses are the same if p_{in} is fixed to be 0.5, which may be justified by theoretical prediction (an inner rotation must be in the Keplerian rotation). Hence, we recommend trying multiple parameter settings (single-power, double-power, fixing and varying p_{in} and p_{out} , etc.) to check how robust a result is. For example, when a double-power fitting provides r_b close to the innermost or outermost radius, this result suggests that the single-power model is better than the double-power model in the given PV diagram. In such a double-power case, a power-law index tends to be strangely high or low simply because there are too few points on the inner ($r < r_b$) or outer ($r > r_b$) side. Similarly, when a double-power fitting provides p_{in} close to the Keplerian index, 0.5, such as the second example in Section 4, it may be better to fix $p_{\text{in}} = 0.5$ because the Keplerian rotation is supported by physics, and fixing p_{in} may provide a better model in the sense of the evidence ratio, i.e., the Bayes factor.

The contribution of infall motion is also controversial. Although PV diagrams along the disk major axis are supposed to trace rotation in the system, the observed velocity is the line-of-sight component of a combination of rotation and infall motions. The outer power-law index is introduced based on a picture where gas is infalling and thus following a lower power-law index in an outer part than in an inner part. Because the infall velocity is expected to be proportional to $r^{-0.5}$, as is the

case with free fall, and is more dominant than rotation in outer radii, the infall motion could cause p_{out} to be closer to 0.5. Whether an infall motion is dominant can be verified by using a 2D distribution of the mean velocity (moment 1) and checking the quadrants in a PV diagram where emission is not expected in the case only with rotation. When the observed system includes infalling gas, the PV diagram along the disk major axis has emission in all the four quadrants, which could shift the edge and ridge points from the radius/velocity in the case only with rotation. In such a case, it is worthwhile to try limiting the velocity range to high velocities where emission is seen only in two quadrants to focus on the rotation-dominant part.

6. SUMMARY

We have developed the `pvanalysis` tool, which is implemented in a python library `SLAM`. The main objective of this tool is to kinematically identify disks around YSOs in the protostellar phase, which are embedded in infalling envelopes, by analyzing PV diagrams of emission line data. As presented in Section 2 and 3, the straightforward concept and usage of this tool make it highly accessible to examine Keplerian rotation of the disks. In Section 4, we demonstrated the application of this tool

to real observational data, which highlights its ability to extract key observational features and to distinguish between the rotational motions of a disk and an infalling envelope. It is important to acknowledge that certain uncertainties still exist, which are currently not taken into account in the tool. These include systemic errors inherent in the method itself and potential contamination from infalling velocities, as discussed in Section 5. Nevertheless, the analytic approach of the `pvanalysis` tool with few assumptions and its simple usage would be a great asset for identifying disks in the protostellar phase and measuring the dynamical mass of central objects.

ACKNOWLEDGMENTS

This paper makes use of the following ALMA data: ADS/JAO.ALMA#2013.1.01086.S and ADS/JAO.ALMA#2015.1.01415.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.