Revista Mexicana de Astronomía y Astrofísica, **43**, 271–281 (2007)

STATISTICAL ANALYSIS OF MOLECULAR LINE EMISSION FROM T TAURI DISK MODELS

Itziar de Gregorio-Monsalvo,[1] Paola D'Alessio,[2] and José F. Gómez[3]

Received 2006 December 4; accepted 2007 March 12

RESUMEN

En este trabajo modelamos la emisión esperada de la línea molecular $C^{17}O(J=3\rightarrow2)$ en discos protoplanetarios, modificando diferentes parámetros físicos para obtener distintas características observacionales. Nuestra meta es determinar la clase de observaciones que nos permitirán extraer información sobre los parámetros físicos de los discos. Con este propósito realizamos un análisis estadístico de componentes principales y una correlación lineal múltiple en el conjunto de resultados obtenidos a partir de los modelos. Además, presentamos un estudio sobre futuras observaciones de línea molecular en discos protoplanetarios usando SMA y ALMA.

ABSTRACT

In this work we model the expected emission from the molecular line $C^{17}O(J=3\rightarrow2)$ in protoplanetary disks, modifying different physical parameters to obtain distinctive observational signatures. Our aim is to determine the kind of observations that will allow us to extract information about the physical parameters of disks. With this purpose we perform a statistical analysis of principal components and a multiple linear correlation on our set of results from the models. We also present prospects for future molecular line observations of protoplanetary disks using SMA and ALMA.

Key Words: **ISM: MOLECULES — METHODS: STATISTICAL — PLANETARY SYSTEMS: PROTOPLANETARY DISKS — RADIATIVE TRANSFER**

1. INTRODUCTION

The mass of disks around young stars is $\sim 99\%$ gas and only 1% dust. However, since the dust opacity is large compared to the gas opacity in a wider range of wavelengths, the dust component dominates the absorption and reprocessing of stellar radiation and the emergent spectral energy distributions (SEDs) of disks around classical T Tauri (CTTS) and Herbig Ae stars (HAe). Most of the models constructed to explain observed SEDs have taken into account only the dust component in the calculation of the disk temperature (Chiang & Goldreich 1997; D'Alessio et al. 1998, 1999; D'Alessio, Calvet, & Hartmann 2001). These models have been successful to explain the observed SEDs. However, SEDs are not sensitive to kinematical information, details of the radial and vertical temperature distribution, the chemistry of the gaseous component (Aikawa et al. 1996, 1997; Aikawa & Herbst 1999; van Zadelhoff et al. 2001; Aikawa & Nomura 2006), the possibility that the gas in the upper layers is hotter than the dust in lower layers by absorption of UV and X-rays (i.e., Glassgold, Najita, & Igea 2004; Jonkheid et al. 2004; Kamp & Dullemond 2004) and/or photoelectrical effect (i.e., Nomura & Millar 2005), etc.

The study of molecular line emission from disks around young stars is an important tool to infer physical characteristics of disks (e.g., Dartois, Dutrey, & Guilloteau 2003; Carr, Tokunaga, & Najita 2004; Piétu, Guilloteau, & Dutrey 2005; Qi et al. 2006; Raman et al. 2006; Dutrey, Guilloteau, & Ho 2007). An advantage of a spectral line is that emission at different frequencies/velocities might be probing different disk regions, making lines an important test for disk models. However, this is also a

[1]European Southern Observatory, Chile.
[2]Centro de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, México.
[3]Instituto de Astrofísica de Andalucía (CSIC), Granada, Spain.

disadvantage in some sense, since one observes intensities convolved with the beam of the telescope and with a finite spectral resolution, and therefore the information of the different regions is mixed up in a complex way. Thus, the analysis and relationship between observations and model properties might become very complicated and difficult to disentangle. There are some previous works that compare molecular line emission from protoplanetary or circumbinary disks with specific models of such an emission (e.g., Koerner, Sargent, & Beckwith 1993; Guilloteau & Dutrey 1998; Qi et al. 2003, 2004) and, in general, agreement between model and observations is fairly good, at least in the general appearance of the maps. However, given the great deal of physical parameters involved in the resulting molecular line emission, it is not straightforward to determine those parameters from a particular observation just by fitting an emission model.

Our main aim in this paper is to identify a set of observational characteristics that give the most information on the physical properties of the disk. Such observational characteristics should then be given the heaviest weight in a fit between observations and models aiming to determine physical parameters in a disk.

This paper is structured as follows: in § 2 we provide an outline of this work. In § 3 we describe the assumptions to calculate the disk structure models and the initial input parameters. In § 4 we explain the radiative transfer calculation and we discuss the selection of the $C^{17}O$ (J=3 $\rightarrow$ 2) as the molecular line to make our study. In § 5 we outline the network of models and the general trends of the line emission maps. In § 6 we describe the statistical study that identifies the set of observational characteristics that give the most information on the physical properties of the disk, and we comment the results derived in § 7. § 8 contains comments and prospects for these studies. Finally, in § 9 we perform a study of the detectability of our modeled disks with SMA and ALMA and we summarize the conclusions in § 10.

2. GENERAL CONSIDERATIONS

In order to achieve the goal described in the introduction, we have developed a set of molecular line emission models calculated for various mass accretion rates, radii, viscosities and maximum dust grain radius distributions. The small scale size of a protoplanetary disk (~100 AU) and their low temperatures (~ 100 K; Beckwith et al. 1990; Miyake & Nakagawa 1995; Beckwith, Henning, & Nakagawa 2000) require observations with high sensitivity and subarsecond angular resolution, since 100 AU subtends $0\farcs7$ at 140 pc (the distance to the Taurus cloud). This makes interferometric observations necessary. With the intention of reproducing a real interferometric observation of a protoplanetary disk with different physical parameters, we integrated the radiative transfer equation and convolved each model with a beam of $0\farcs4$, as a compromise between resolution and sensitivity. For each resultant map, we have measured different observational signatures, as if they were data from a real interferometric observation. Finally, in order to obtain the best combinations of such observational parameters that yield more information about the physical characteristics of disks, we have undertaken a novel statistical approach to link observational properties of the expected molecular line emission with the underlying physical properties of the disk, by means of a principal component and multiple linear correlation analysis. We show that this is a promising type of analysis to prepare the observations to be made with the new generation of millimeter and submillimeter interferometers.

In this study it is important to choose an appropriate molecular transition sensitive to the relevant physical parameters. To carry out our study we have selected the $C^{17}O$(J=3 $\rightarrow$ 2) transition at 337 GHz. This line is a high excitation transition of a CO isotope with very low abundance, which makes it less susceptible to be affected by absorption and/or the emission from the surrounding cloud material. This transition is also a suitable candidate to be observed in protoplanetary disks using the Submillimeter Array (SMA) and the Atacama Large Millimeter Array (ALMA), as shown by Gómez & D'Alessio (2000).

3. DISK STRUCTURE MODELS

3.1. *Assumptions*

We base our calculations of molecular line emission on structure models of accretion disks irradiated by the central star, which have been previously used to explain different observations of classical T Tauri stars. The assumptions and calculation method of such models are described in D'Alessio et al. (1998, 1999, 2001). In summary, the disk is assumed to be in steady state, with a constant mass accretion rate $\dot{M}$ and an $\alpha-$viscosity (Shakura & Sunyaev 1973), with a constant value of the viscosity parameter α. The disk is in vertical hydrostatic equilibrium in the gravitational potential well of the star, and we neglect the disk self-gravity. We assume that gas and dust are thermally coupled, having the same temperature everywhere. This dust/gas temperature enters in the calculation of the disk volumetric density distribution through the integration of the hydrostatic equilibrium equation. The main heating mechanisms considered are viscous dissipation and stellar irradiation. The viscous dissipation is important for heating the inner regions (close to the star and close to the midplane), the direct stellar irradiation heats the disk atmosphere, and the stellar radiation scattered and reprocessed by the disk upper layers heats the whole vertical structure. The transfer of radiation through the disk is calculated taking into account that the dust scatters and absorbs stellar and disk radiation, implying that the temperature structure depends on the dust properties. The viscous irradiated disk models used in the present study show the temperature inversion previously found by Calvet et al. (1991, 1992), i.e., at the outer disk, $R \gtrsim 10$ AU, the upper layers are hotter than at the disk midplane, because they are heated by direct stellar irradiation (see also Chiang & Goldreich 1997 and D'Alessio et al. 1998).

The dust opacity is calculated using the Mie theory for compact spherical grains. We consider a distribution of sizes given by $n(a) = n_0 a^{-p}$, where a is the radius of the grains, n_0 is a normalization constant, and p is a free parameter. In this work, we have adopted $p = 3.5$ from Draine & Lee (1984) and the model of dust composition proposed by Pollack et al. (1994) with the variations introduced by D'Alessio et al. (2001). To account for the possibility of dust growth, we adopt different values of maximum grain sizes. Dust grains of different sizes have different continuum opacity at mm wavelengths, affecting the molecular line emission in different ways. The existence of larger grains in disks than in the interstellar medium was proposed to explain the observed slope of the continuum SED at millimeter wavelengths (Beckwith & Sargent 1991; Miyake & Nakagawa 1995). Large grains could be depleted from higher layers of the disk, but could be well mixed with gas below a few gas scale heights. For simplicity, the disk models adopted here (from D'Alessio et al. 2001) assume that gas and dust are well mixed. This seems a reasonable assumption if the emission from the molecular line arises mainly from areas closer to the midplane than to the upper layers. This point will be discussed later (subsection 4.1).

3.2. *Input parameters*

For the present study we have considered the following input parameters: maximum disk radius (R_d), maximum radius of dust grains (a_{max}), disk mass accretion rate ($\dot{M}$), and viscosity parameter (α) (see values in Table 1).

TABLE 1

INITIAL PHYSICAL PARAMETERS

R_d[a] (AU)	a_{max}[b] (μm)	$\dot{M}$[c] ($M_\odot$/year)	α[d]
50	1	10^{-9}	0.001
100	10	3×10^{-8}	0.005
150	10^2	10^{-7}	0.01
	10^3		0.02
	10^4		0.05
	10^5		

[a]Disk radius.

[b]Maximum radius of dust grains.

[c]Mass accretion rate.

[d]Viscosity parameter.

We have adopted typical parameters of a T Tauri star from Gullbring et al. (1998), i.e., M_*=0.5$M_\odot$, R_*=2$R_\odot$, and T_*=4000 K for all the models. The disks are assumed to be at 140 pc, the distance of the Taurus molecular cloud (Kenyon, Dobrzycka, & Hartmann 1994), with a typical inclination angle i=60°. It is important to mention that each disk structure is self-consistently calculated given these input parameters. This means that the whole disk structure is affected by all the parameters, consequently affecting the line properties. This might complicate the analysis of the resulting line properties, but we think this gives a more realistic description of the interplay between the different variables.

4. RADIATIVE TRANSFER AND MOLECULAR LINE EMISSION

The model of the disk structure provides a detailed density and temperature distribution through the disk as a function of the height and the distance to the disk center. To derive the line intensity for a given molecular transition, we must solve the transfer equation. We have used the same assumptions and formalism that Gómez & D'Alessio (2000) use, and we summarize them briefly here. We assume local thermal equilibrium for the population of the molecular energy levels and we consider thermal line profiles. We divide the disk in a grid of cells considering isovelocity lines and their perpendicular lines (see Appendix in Gómez & D'Alessio 2000). We integrate the transfer equation $dI_\nu/ds = \kappa_\nu \rho (S_\nu - I_\nu)$ through the line of sight at the center of each cell, where I_ν is the intensity, s is the length along the line of sight, κ_ν is the absorption coefficient that considers the contributions from the line and continuum, i.e., $\kappa_\nu = \kappa_l + \kappa_c$, ρ is the mass density of the gas, and S_ν the source function. The coefficient κ_c is dominated by dust and we consider pure absorption opacity of each kind of dust grain size.

In order to derive the flux density, we convolve all our models with a beam of $0\farcs4$ of HPBW, as a compromise between resolution and sensitivity. Finally we subtract the continuum emission to isolate the flux density of the molecular line, obtaining a set of model results that reproduce a real observation of a protoplanetary disk with different physical parameters.

4.1. *Selection of the emission line transition*

It is very common that young stellar objects like T Tauri stars are still embedded in the material of the parental cloud. The envelope that surrounds the disk-star system is composed of cold gas and dust that could hide the disk emission.

On the one hand, the surrounding material could absorb the emission from the hotter, innermost part of the disk structure. This problem can be solved by selecting a molecular transition whose energy is high enough to trace the hotter gas from the disk, while few molecules in the colder envelope are in the states involved in this transition, thus avoiding line absorption at the envelope. However, the frequency of the transition must not be too high, because otherwise the dust in the envelope would become optically thick. On the other hand, the molecular line emission from the whole cloud could hide the deeper emission from the disk, as is usually the case for the more abundant CO isotopes. To make sure that the observed emission comes only from the disk structure, we must select a molecular species of low abundance for which the envelope is optically thin and therefore its emission is negligible.

In order to fulfill these conditions, we have chosen the (J=3 $\rightarrow$ 2) transition of the $C^{17}O$ molecule at 337 GHz. Its molecular abundance relative to H_2 in molecular cores is low, 5.0×10^{-8} (White & Sandell 1995) and the frequency transition at 337 GHz is still low enough to avoid being significantly absorbed by the dusty envelope.

In a recent work, Dartois et al. (2003) have studied the disk vertical temperature structure using different isotopes of CO. Being characterized by different opacities, different lines trace the gas at distinct depths. Dartois et al. (2003) find a good agreement between the inferred temperature for each transition/isotope and the temperature where $\tau \sim 1$ in irradiated (dusty) disk models. Following Dartois et al. (2003), we have studied the formation regions of molecular lines from different isotopes of CO (CO, ^{13}CO, $C^{18}O$, and $C^{17}O$) in our set of disk models, and shown that the $C^{17}O$ is formed closer to the midplane than the rest of the isotopes. This fact makes the LTE assumption acceptable, since the region that mainly contributes to the line emission shows higher densities than the critical density of the selected transition ($\simeq 5 \times 10^4$ cm^{-3}).

In addition, given the high gas densities in the disk, we expect that the abundance of this isotope, would be less affected by photodisociation produced by the incident radiation, since it would be shielded against it, specially in the deepest layers. Moreover, theoretical models of the evolution of molecular abundances in protoplanetary disks, predict depletion of CO from the gas phase for temperatures below 20 K (Aikawa et al. 1996) and therefore, depletion is probably not significant for the chosen molecule and for the disk radii we are considering here ($<$ 150 AU). Thus, for simplicity, we adopt a constant abundance for $C^{17}O$ relative to H_2, given the molecular core value (White & Sandell 1995).

5. RESULTS

We have calculated the expected emission in the $C^{17}O$(J=3 $\rightarrow$ 2) transition for a typical T Tauri star with a disk inclination angle of 60°, and for all possible combinations of the physical parameters shown on Table 1. We have 241 different models for which we have integrated the radiative transfer equation at 12 different velocities, from -2.5 to 2.5 km s^{-1} at steps of 0.5 km s^{-1}. The results are maps like those shown in Figure 1, in which we only show the positive velocities, since the maps are similar and almost symmetrical for the negative velocities (there are slight differences though, due to the hyperfine structure of $C^{17}O$ transition). For each map we have measured the following observational signatures as if they were data from a real interferometric observation: intensity of the principal (the more distant from the observer, to the north in our maps) and secondary (closer to the observer, to the south in our maps) peak at each velocity, distance from the disk center to the principal peaks, half power size of the emission distribution, and velocity at which the maximum intensity is present. These represent a total of 43 different observational parameters for each input model.

5.1. *Line emission maps*

In all maps obtained from our set of models we observe the same tendencies as Gómez & D'Alessio

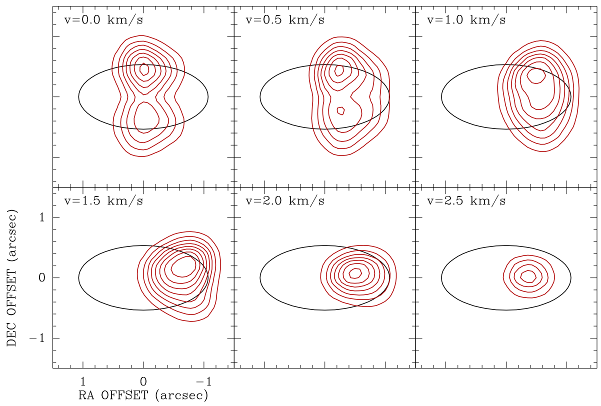

Fig. 1. Emission maps at different velocities for a disk with radius = 150 AU, a_{max}=10 μm, i=60^o, α=0.01 and mass accretion rate = $10^{-7} M_\odot$ year^{-1}. Maps have been convolved with a 0$''$.4 beam. The lowest contour and the increment step are 20 mJy beam^{-1}. The ellipse traces the outer edge of the disk.

(2000). Summarizing, we observe an asymmetry on both sides of the major axis, as expected from optically thick emission. The areas further away from the observer (positive declination in Figure 1) show line emission of higher intensity because the line of sight intercepts areas of the disk closer to the central star, where the gas is warmer. This result is confirmed in DM Tau spectral line observations by Dartois et al. (2003), where an inclination angle $i \simeq -37^o$ was assumed. We notice another asymmetry on both sides of the minor axis, that it is more pronounced when approaching the velocity of the cloud ($v = 0$ km s^{-1}). It is caused by the asymmetry of the hyperfine structure in the $C^{17}O$ molecular transitions.

We find that the maximum intensity of the line at the systemic velocity ($v = 0$ km s^{-1}) traces the outer edge of the disk. As previously discussed by Sargent & Beckwith (1991) and Gómez & D'Alessio (2000), this is a consequence of the fact that the effective area emitting at a given velocity within the beam increases with the distance to the central star more steeply than the decrease of brightness temperature with distance.

Moreover, at the center of the disk the emission intensity diminishes, and we can even see absorption lines in some of our maps. The high dust continuum opacities typical for the central parts of the disk reduce the contrast between the emission lines and the continuum, and, if they are high enough, lines could show up in absorption.

6. STATISTICAL TOOLS

As we mentioned in Section 5, we have chosen a set of 43 different observational parameters to characterize the maps resulting from our model calculations. These parameters are, in principle, somewhat arbitrary. If we want to simulate a real observation, in which we would like to extract information about the underlying physical characteristics of disks, it is obvious that some of these 43 parameters will have more informative power, while some may turn out to be irrelevant. Moreover, it is likely that not all the chosen observational parameters will be independent.

Here, we have undertaken a statistical analysis to try to identify a set or a combination of observational parameters that could render more information about the disks properties. First, we used a principal component analysis to reduce the number of observational parameters to a small, informative set. Later, we investigated whether we can obtain quantitative values of physical magnitudes from observational parameters by means of multiple linear regression.

6.1. *Principal Components*

The principal component analysis is a statistical technique that provides a dimensional reduction of a set of variables. In our case, each of the initial 43 observational parameters would be an axis in a system of coordinates in a multidimensional space. The method consists of finding a set of orthogonal axes in which the variance (heterogeneity) of our data is maximum (see Thurstone 1947; Kaiser 1958). This is solved through a linear and orthogonal transformation that corresponds to a rigid rotation of the original data into a new set of coordinates. The eigenvalues obtained provide information about the variances of the data in the new space, and the eigenvectors represent the direction of the axes in the new space of representation of our data. In our study, this analysis reduces the number of observational signatures necessary to derive information about the physical parameters to a smaller set of linearly independent parameters: the principal components.

6.2. *Multiple linear correlation*

In order to quantitatively estimate each physical parameter from the set of observational variables, we have also carried out a multiple regression analysis (Pearson 1908). In our case, the obtained principal components will be considered the independent variables, and the physical properties of the disk will be the dependent ones. The multiple linear correlation analysis will then try to make the best fit to derive a linear function of the form:

$$y_i = a_1x_1 + a_2x_2 + ... + a_jx_j \ ,$$

where y_i is the dependent variable, and x_j are the independent ones. In an ideal case (correlation coefficient $\simeq 1$), we would obtain a function with a good predictive power for the dependent variable. For our particular problem, we would like to obtain a formula with which, from a set of observational parameters, we could calculate the physical characteristics of the disk.

Moreover, to test the signification of the multiple regression we applied the F-Snedecor test (Snedecor 1934) as hypothesis testing. The ratio of two chi-squares divided by their respective degrees of freedom follows an F-distribution. The test consist of comparing the relation between the variance of the predicted values for the dependent valuable and the error variance with the value of the F-distribution.

7. OBSERVATIONAL SIGNATURES VS. PHYSICAL PARAMETERS

7.1. *Eigenvalues and Eigenvectors of Principal Components*

To reduce the number of relevant components in the principal component analysis, we adopted the Kaiser criterion (Kaiser 1960), which only retains factors with eigenvalues greater than 1. In our case, the criterion selects four factors, which account for 91% of the total variance of the system. In fact, the first and the second factors alone represent 82% of the total variance. Therefore, most of the following analysis is based on these two first principal components.

We show the eigenvectors for each principal component in Table 2. The numerical entries in this table show the linear combination coefficients for each observational parameter used to build the corresponding principal component. Therefore, they indicate the relative weight of each observational parameter on the components.

For the first principal component (PC1) the observational parameters with a larger weight in its definition are (in order of decreasing relative weights) the velocity of the peak emission, the half power sizes for principal peaks at intermediates velocities, and the distance from principal peaks to center. The parameters that define the second principal component (PC2) are the half power sizes of principal peaks at 1.5 km s^{-1}, the velocity of the peak emission, and the half power sizes of secondary peaks at 0.0 and 1.5 km s^{-1}. The third principal component (PC3)

TABLE 2

EIGENVECTORS

Observational Parameters	PC1	PC2	PC3	PC4
$I(v_0)$[a]	−0.0279	0.0020	0.0030	0.0033
$I(v_{0.5})$[a]	−0.0276	0.0019	0.0029	0.0034
$I(v_1)$[a]	−0.0281	0.0007	0.0022	0.0035
$I(v_{1.5})$[a]	−0.0340	−0.0023	0.0024	0.0040
$I(v_2)$[a]	−0.0276	−0.0086	0.0016	0.0040
$I(v_{2.5})$[a]	−0.0118	−0.0082	0.0028	0.0040
$r(v_0)$[b]	−0.1360	0.0388	−0.0072	−0.0137
$r(v_{0.5})$[b]	−0.1462	0.0325	−0.0041	−0.0169
$r(v_1)$[b]	**−0.1581**[*]	0.0448	−0.0015	−0.0086
$r(v_{1.5})$[b]	**−0.1930**	0.0100	0.0266	−0.0262
$r(v_2)$[b]	−0.1077	0.0075	−0.0030	−0.0150
$r(v_{2.5})$[b]	−0.0483	0.0310	−0.0094	−0.0062
$I_{sec}(v_0)$[c]	−0.0289	−0.0047	0.0004	0.0019
$I_{sec}(v_{0.5})$[c]	−0.0297	−0.0042	0.0006	0.0022
$I_{sec}(v_1)$[c]	−0.0179	−0.0091	−0.0024	−0.0018
$I_{sec}(v_{1.5})$[c]	−0.0003	0.0007	−0.0009	−0.0004
$I_{sec}(v_2)$[c]	0.0000	0.0000	0.0000	0.0000
$I_{sec}(v_{2.5})$[c]	0.0000	0.0000	0.0000	0.0000
$a(v_0)$[d]	−0.0595	0.0426	0.0316	0.0034
$b(v_0)$[d]	−0.0790	0.0334	0.0272	−0.0184
$a(v_{0.5})$[d]	**−0.1615**	−0.0309	0.0259	0.0043
$b(v_{0.5})$[d]	**−0.1947**	−0.0788	**0.0523**	−0.0076
$a(v_1)$[d]	−0.1138	−0.0250	0.0281	−0.0024
$b(v_1)$[d]	−0.0388	−0.0024	0.0041	0.0017
$a(v_{1.5})$[d]	**−0.1777**	**−0.1926**	−0.0350	**−0.0615**
$b(v_{1.5})$[d]	**−0.2271**	**−0.1499**	0.0086	**0.0743**
$a(v_2)$[d]	**−0.2420**	−0.0171	**0.0543**	−0.0023
$b(v_2)$[d]	−0.1181	0.0056	0.0133	−0.0084
$a(v_{2.5})$[d]	−0.0286	0.0168	−0.0054	0.0011
$b(v_{2.5})$[d]	−0.0028	0.0114	−0.0067	0.0010
$a_{sec}(v_0)$[e]	−0.0746	**0.1023**	**0.0703**	**0.0434**
$b_{sec}(v_0)$[e]	−0.0436	**0.1026**	0.0406	**−0.0607**
$a_{sec}(v_{0.5})$[e]	−0.0026	0.0161	−0.0127	−0.0076
$b_{sec}(v_{0.5})$[e]	−0.0006	0.0047	−0.0057	−0.0011
$a_{sec}(v_1)$[e]	0.0000	0.0000	0.0000	0.0000
$b_{sec}(v_1)$[e]	0.0000	0.0000	0.0000	0.0000
$a_{sec}(v_{1.5})$[e]	−0.0733	**0.0985**	**0.0683**	**0.0525**
$b_{sec}(v_{1.5})$[e]	−0.0398	**0.0951**	0.0398	**−0.0549**
$a_{sec}(v_2)$[e]	−0.0040	0.0185	−0.0141	−0.0097
$b_{sec}(v_2)$[e]	−0.0008	0.0043	−0.0060	−0.0006
$a_{sec}(v_{2.5})$[e]	0.0000	0.0000	0.0000	0.0000
$b_{sec}(v_{2.5})$[e]	0.0000	0.0000	0.0000	0.0000
$v_{I_{max}}$[f]	**0.3176**	**−0.1431**	**0.1527**	−0.0241

[a]Intensity of the principal peak at each velocity (0, 0.5, 1, 1.5, 2, 2.5 km s^{-1}) in Jy beam^{-1}.

[b]Distance from disk center to principal peaks in arcsec.

[c]Intensity of the secondary peak at each velocity.

[d]Minor (a) and major (b) axis of the half power sizes of emission for the principal peak of intensity at each different velocity, in arcsec.

[e]Minor (a) and mayor (b) axis of the half power sizes of emission for the secondary peak of intensity at each different velocity.

[f]Velocity in km s^{-1} at which the maximum intensity is present.

[*]Values in boldface represent the observational parameters with largest weights in the definition of each PC.

is defined by the velocity of the peak emission and the half power sizes for principal peaks at 0.5 and 2.0 km s^{-1}, and for secondary peaks at 0.0 and 1.5 km s^{-1}. Finally, the physical parameters that define the fourth principal component (PC4) are the half power sizes for principal peaks at 1.5 km s^{-1} and for secondary peaks at 0.0 and 1.5 km s^{-1}. In Table 2 we have marked in boldface the most representative observational parameters.

7.2. *PC1(kinematical component)-PC2(spatial component) diagrams*

Due to the fact that the first and the second principal component represent the 82% of the total variance of the system, we have represented all our models in a PC1-PC2 diagram, to check whether such diagrams can be used to discriminate disks with particular physical characteristics (if we see some clustering related to the physical properties). Considering the parameters with the larger weights in each of these principal components, we have named them as "kinematical component" and "spatial component" in the case of PC1 and PC2, respectively.

The most evident result is shown when we represent disks with different radii in the PC1(kinematical)-PC2(spatial) diagram (see Figure 2). The disks with radius 50 AU are distributed on the right part of the diagram, the ones with radius 100 AU are located in the central part and the disks with radius 150 AU are located on the left of the plot. This means that the kinematic principal component (x axis) is qualitatively good to discriminate among disks with different radii. This result indicates that we can get information about the radius of a protoplanetary disk (physical parameter) from the velocity of the peak emission, the half power sizes for principal peaks at intermediates velocities and from the distance from principal peaks to the center (observational signatures). That the disk radii can be discriminated relatively well by maps of line emission may seem a relatively obvious result, but it illustrates the power of this kind of statistical analysis.

Another interesting trend is seen from the representation in the diagram of disks with different mass accretion rate (see Figure 3). Disks with higher mass accretion rates tend to populate the upper parts of the PC1(kinematical)-PC2(spatial) diagram. Therefore, in this case it is the second principal component (y axis) that better discriminates among disks with different mass accretion rates. Considering the parameters that give rise to this component, we can say that the half power sizes of principal peaks at 1.5 km s^{-1}, the velocity of the peak emission and the half power sizes of secondary peaks at 0.0 and 1.5 km s^{-1} velocities provide information about the mass accretion rate of the disks.

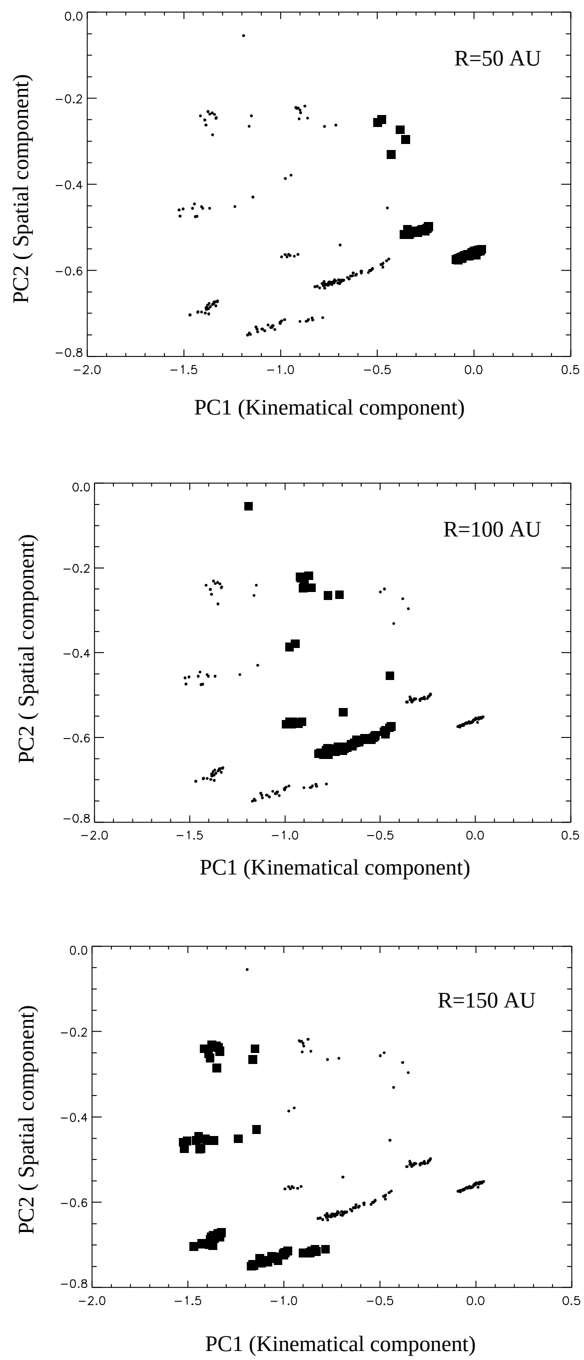

Fig. 2. Disk radii in the PC1-PC2 diagrams. Dots represent all the modeled disks with different physical parameters. Squares represent disks with radius 50 AU (upper), 100 AU (central) and 150 AU (bottom).

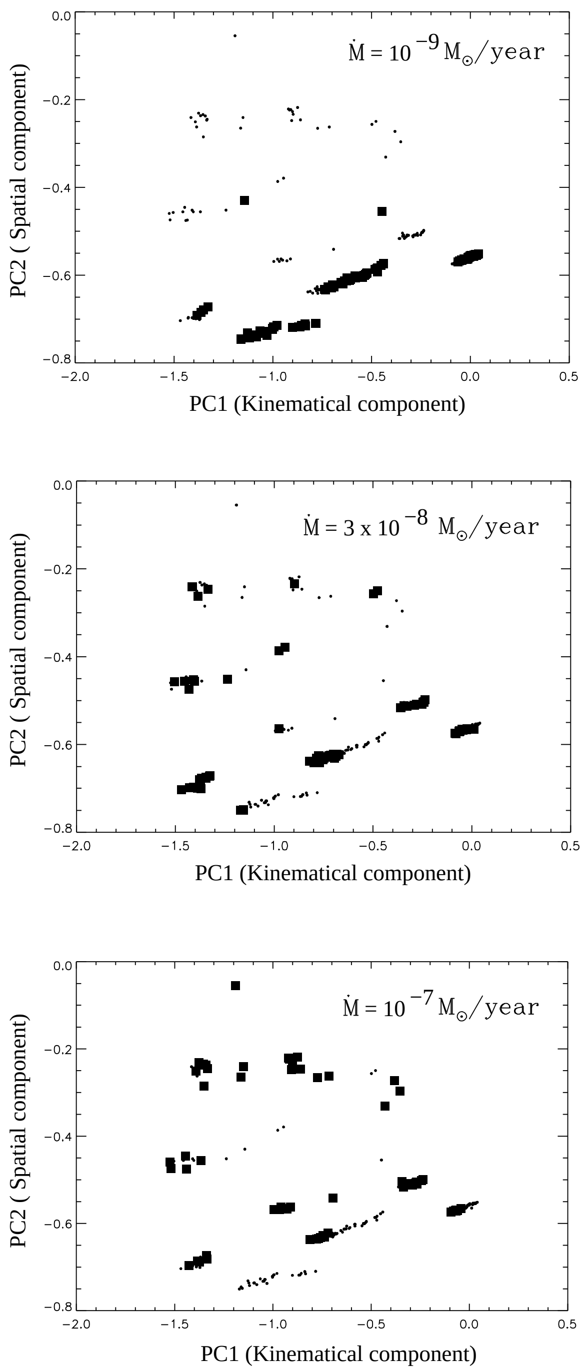

Fig. 3. Mass accretion rates in the PC1-PC2 diagrams. Dots represent all the modeled disks with different physical parameters. Squares represent disk with mass accretion rate of 10^{-9} $M_\odot$/year (upper), 3×10^{-8} $M_\odot$/year (central) and 10^{-7} $M_\odot$/year (bottom).

Other trends relating principal components and the rest of the physical parameters are also present, but qualitatively they are not as clear as the two we have mentioned. In the case of other physical parameters, it is the second principal component that provides more information about the α parameter and the maximum radius of dust grains. Further analysis including models of more molecular transitions will certainly be useful to obtain refined principal components that can better discriminate among these physical parameters.

7.3. *Multiple correlation*

We have seen in the previous section that diagrams of principal components can be useful to discriminate among disks with different physical characteristics. It would be interesting, however, to have a mathematical tool to easily obtain numerical values for these physical characteristics based on the observed maps. As a first approach, we have here performed a multiple linear correlation with the observational parameters as independent variables, and the physical characteristics as dependent ones. The resulting coefficients for such a fit are shown in Table 3.

As was deduced in the previous subsection, the linear correlation coefficients of each principal component (see r(PC$_i$) values in Table 3) show that PC1(kinematical component) is the best one to derive information about the disk radii and PC2(spatial component) provides the most information in the determination of the rest of the physical parameters, $\dot{M}$, α, and maximum radius of dust grains.

The most important piece of information that can be retrieved with this kind of study is the linear combination of observational signatures that provide quantitative information about physical characteristics. This linear combination can be derived from λ_i coefficients in Table 3 as follows:

$$P = \lambda_0 + \lambda_1 PC1 + \lambda_2 PC2 + \lambda_3 PC3 + \lambda_4 PC4$$

where P is the physical parameter. As an example, in the case of the radius of the protoplanetary disk, this combination would be the following:

$$R_d = -55.3 - (79.4 \times PC1) - (54.2 \times PC2) +$$

$$+ (133.3 \times PC3) - (137.7 \times PC4)$$

With these sets of linear combinations, we could estimate physical parameters from observations, provided that the fit is good enough.

The strongest correlation (see R values in Table 3) is obtained for radius (correlation coefficient R = 0.97), followed by mass accretion rate

TABLE 3

CORRELATION COEFFICIENTS AND F-SNEDECOR TEST VALUES

	r(PC1)[a]	r(PC2)[a]	r(PC3)[a]	r(PC4)[a]	R[b]	F
Radius	−0.9	−0.19	0.07	−0.001	0.97	951.62
$\dot{M}$	−0.18	0.5	−0.3	−0.16	0.57	26.93
α	0.16	−0.21	0.13	0.07	0.31	5.94
Max. Grain Radius	−0.08	−0.009	−0.012	0.08	0.19	2.07
	λ_0[c]	λ_1[c]	λ_2[c]	λ_3[c]	λ_4[c]	
Radius	−55.3	−79.4	−54.2	133.3	−137.7	
$\dot{M}$	1.8×10^{-7}	-0.16×10^{-7}	1.4×10^{-7}	-1.7×10^{-7}	-0.15×10^{-7}	
α	−0.0195	0.0024	−0.0349	+0.0465	−0.0091	
Max. Grain Radius	−16943.4	−100.4	−40213.2	+48703.1	+123751.8	

[a]r(PCi) are the linear correlation coefficients of each principal component.
[b]Multiple correlation coefficient.
[c]λ_i are the coefficients of the linear combination of principal components obtained from the multiple regression study, to derive the values of the physical parameters in the first column.

(R = 0.57), viscosity parameter (R = 0.31) and maximum radius of dust grains (R = 0.19). The multiple correlation coefficients alone are not good statistical indicators of the goodness of linear fits. To assess the validity of our fitted functions we carried out an F test to derive the significance of the linear regression. We assume a confidence limit of 95%, which means that we could admit as good fits with values $F < 1.44$. However, in all our results the values exceeded this critical value (see Table 3). This result suggests that our variables are far from the linear regime, which is certainly reasonable.

7.4. *Effects of measurements errors*

In our method, the derived value of a physical parameter depends linearly on the principal components, which are linear combinations of the observational data. Therefore we must consider on the one hand the error in the calculation of the principal components, and on the other the error derived from the calculation of the multiple correlation.

In a real observation, the error of the intensity will depend on the rms noise of the observations and, in particular, the achieved signal-to-noise ratio will determine the positional accuracies (1σ error in position $\simeq HPBW/(2\,\mathrm{SNR})$, assuming a compact emitting source). In our method for the determination of the most significant principal components, the most important sources of error are related to the observational parameters that contribute with a large weight, i.e., the half power sizes of emission for the principal and secondary peaks of intensity (whose error can be roughly estimated as $\simeq HPBW/\mathrm{SNR}$), and the velocity at which the maximum intensity is present (with 1σ error $\sim \Delta v/(2\,\mathrm{SNR})$, where Δv is the linewidth). Since principal components are linear combinations of observational parameters, the weight of the errors associated with each measured observational parameter will depend on theirs particular weight in the definition of each principal component (see Table 2). Nevertheless, we must point out that, in the case of calculations of the physical parameters of disks, the error is dominated by the uncertainties in the fits of the multiple regression. Since our derived variables are far from a linear regime (see F test above), measurement errors are not significant in comparison.

8. COMMENTS AND PROSPECTS FOR THESE STATISTICAL STUDIES

The results derived from the statistical study presented in this Paper show that this method could be a powerful tool to obtain from observational parameters information related to the physical characteristics of protoplanetary disks.

The most important information of our study is related to the radius of the disks. This preliminary study provides a way to obtain a first approximation of disk radii that depends only on the observations, avoiding the application of χ^2 fitting techniques.

This is a promising method of study, and considering the future set of observations in protoplanetary disks that will be carried out with the development of ALMA, we plan to complete and improve our models and analysis.

Nowadays it is not possible to test observationally our method because there are no observations of protoplanetary disks in our selected molecular transition, $C^{17}O(J=3\rightarrow2)$. But considering the excellent characteristics of this transition/isotope to observe protoplanetary disks (see subsection 4.1), one should be able to test our study in the near future with ALMA observations. In the meantime we will improve or study as follows.

First, to better determine the tendencies in the principal component diagram, we plan to increase the sample, building a more extended set of models with a wider variety of initial physical and observational parameters. Calculations in other molecular transitions or isotopes would further constrain the information on temperature and density. An update in the calculation of the disk structure models will be done, considering in detail the photodisociation or depletion effects in the line of study, along with dust grain growth and settling. We also plan to study the application of our method to non-axisymmetric disks, for which the development of a new disk structure that takes into account the lack of symmetry is needed. Finally, a more sophisticated method to derive the physical parameters of disks from the principal components is desirable, since the application of multiple correlation provides variables far from the linear regime and introduces the greatest source of error in our study.

9. DETECTABILITY OF MOLECULAR LINES WITH NEW INTERFEROMETERS

It would obviously be important to be able to test our results observationally in the future. With this mind, we have also studied the detectability of the calculated models, when observed with interferometers that will be able to reach subarcsecond resolution at the frequency of the $C^{17}O(J=3\rightarrow 2)$ transition, such as SMA and ALMA.

We have calculated the line intensity expected for all our line models. The lowest peak intensity ($\sim$7 mJy beam^{-1}) corresponds to disks with R_d=50 AU, $\dot{M}$=10^{-9} $M\odot$ yr^{-1}, α=0.05 and 1 μm maximum grain radius. On the other hand, the highest peak intensity ($\sim$180 mJy beam^{-1}) corresponds to disks with 150 AU radius, $\dot{M}$=10^{-7} $M\odot$ yr^{-1}, α=0.01 and 10^5 μm maximum grain radius.

With the SMA, a sensitivity of $\sim$85 mJy beam^{-1} is expected[4], considering the 8 antennas of the array working at 337 GHz, with 1 km s^{-1} velocity resolution and 10 hours of integration time, under standard values of precipitable water vapor ($\sim$2.0 mm for 300−355 GHz).

[4]http://sma1.sma.hawaii.edu/call.html

On the other hand, in the case of ALMA, considering 64 antennas, 1 km s^{-1} velocity resolution, only 1 hour of integration time and 1.5 mm of precipitable water vapor (median at the site over all hours and seasons), the sensitivity expected is 1.2 mJy beam^{-1} (Butler & Wootten 1999).

Considering 5σ emission as detections, we conclude that all our modeled disks, even the faintest, would be detected with a one-hour integration time with ALMA. Nevertheless, it will be extremely difficult of observe any of our modeled disks with the SMA, at least the ones with the physical characteristics showed in this work (see Table 1), since the SNR obtained for the modeled disks with the highest intensities ($\sim$180 mJy beam^{-1}) in 10 hours of integration time, would only be 2σ.

Therefore, ALMA will be a crucial instrument to observationally test the predictions and assumptions of our models.

10. CONCLUSIONS

In this paper we have presented a statistical method to derive physical parameters from observational characteristics in a protoplanetary disk. We modeled the expected emission of the $C^{17}O(J=3\rightarrow 2)$ transition from protoplanetary disks with different physical properties. We then applied a principal component and a multiple correlation analysis, to obtain a set of linear combinations of observational parameters that may provide relevant information of the physics of the disks. The main conclusions are the following:

- The most significant results of our analysis are related to the disk sizes. We can discriminate among disks with different radii using a principal component composed mainly of the velocity of the peak emission, the half power sizes for principal peaks at intermediates velocities and the distance from principal peaks to center.
- Moreover, some information about the mass accretion rate could be obtained from a principal component made of the half power sizes of principal peaks at 1.5 km s^{-1}, the velocity of the peak emission and the half power sizes of secondary peaks at 0.0 and 1.5 km s^{-1} velocities, although the results are much less significant than in the case of the radii.
- Our preliminary results shows that the statistical method presented here seems to be promising and useful, and it will be improved and completed in the future with studies of other transitions.

- We have performed a study of detectability with ALMA and SMA. All our modeled disks could be detected with ALMA, using a one-hour integration time. However, the sensitivity reached by the SMA is not enough to detect our disks with reasonable integration times. We conclude that ALMA will play an important role in testing observationally our models and our statistical results.

We wish to thank the LAEFF staff for allowing the use of their personal computers to carry out the calculations of the molecular line emission models. PD acknowledges grants from Conacyt and PAPIIT, México. The model calculations were performed in the cluster at CRyA-UNAM acquired with Conacyt grants 36571-E, 47366-F, and UNAM-PAPIIT grant 110606. FG and IdG are supported by Junta de Andalucía grant FQM-1747. JFG also acknowledges partial support from grants AYA2005-08523-C03-03 of the Spanish MEC (including FEDER funds) and TIC-126 of Junta de Andalucía.

REFERENCES

Aikawa, Y., & Herbst, E. 1999, A&A, 351, 233
Aikawa, Y., & Nomura, H. 2006, ApJ, 642, 1152
Aikawa, Y., Miyama, S. M., Nakano, T., & Umebayashi, T. 1996, ApJ, 467, 684
Aikawa, Y., Umebayashi, T., Nakano, T., & Miyama, S. M. 1997, ApJ, 486, L51
Beckwith, S. V. W., Henning, T., & Nakagawa, Y. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Rusell (Tucson: Univ. Arizona Press), 533
Beckwith, S. V. W., & Sargent, A. I. 1991, ApJ, 381, 250
Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Guesten, R. 1990, AJ, 99, 924
Butler, B., & Wootten, A. 1999, ALMA Memo, 276
Calvet, N., Magris, G. C., Patino, A., & D'Alessio, P. 1992, RevMexAA, 24, 27
Calvet, N., Patino, A., Magris, G. C., & D'Alessio, P. 1991, ApJ, 380, 617
Carr, J. S., Tokunaga, A. T., & Najita, J. 2004, ApJ, 603, 213
Chiang, E. I., & Goldreich, P. 1997, ApJ, 490, 368
D'Alessio, P., Calvet, N., & Hartmann, L. 2001, ApJ, 553, 321
D'Alessio, P., Calvet, N., Hartmann, L., Lizano, S., & Cantó, J. 1999, ApJ, 527, 893
D'Alessio, P., Cantó, J., Calvet, N., & Lizano, S. 1998, ApJ, 500, 411
Dartois, E., Dutrey, A., & Guilloteau, S. 2003, A&A, 399, 773
Draine, B. T., & Lee, H. M. 1984, ApJ, 285, 89
Dutrey, A., Guilloteau, S., & Ho, P. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson: Univ. Arizona Press) 495
Glassgold, A. E., Najita, J., & Igea, J. 2004, ApJ, 615, 972
Gómez, J. F., & D'Alessio, P. 2000, ApJ, 535, 943
Guilloteau, S., & Dutrey, A. 1998, A&A, 339, 467
Gullbring, E., Hartmann, L., Briceño, C., & Calvet, N. 1998, ApJ, 492, 323
Jonkheid, B., Faas, F. G. A., van Zadelhoff, G.-J., & van Dishoeck, E. F. 2004, A&A, 428, 511
Kaiser, H. F. 1958, Psychometrika, 23, 187
__________. 1960, Educational and Psychological Measurement, 20, 141
Kamp, I., & Dullemond, C. P. 2004, ApJ, 615, 991
Kenyon, S. J., Dobrzycka, D., & Hartmann, L. 1994, AJ, 108, 1872
Koerner, D. W., Sargent, A. I., & Beckwith, S. V. W. 1993, ApJ, 408, L93
Miyake, K., & Nakagawa, Y. 1995, ApJ, 441, 361
Nomura, H., & Millar, T. J. 2005, A&A, 438, 923
Pearson, K. 1908, Biometrika, 6, 59
Piétu, V., Guilloteau, S., & Dutrey, A. 2005, A&A, 443, 945
Pollack, J. B., Hollenbach, D., Beckwith, S., Simonelli, D. P., Roush, T., & Fong, W. 1994, ApJ, 421, 615
Qi, C., Kessler, J. E., Koerner, D. W., Sargent, A. I., & Blake, G. A. 2003, ApJ, 597, 986
Qi, C., et al. 2004, ApJ, 616, L11
__________. 2006, ApJ, 636, L157
Raman, A., Lisanti, M., Wilner, D. J., Qi, C., & Hogerheijde, M. 2006, AJ, 131, 2290
Sargent, A. I., & Beckwith, S. V. W. 1991, ApJ, 382, L31
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Snedecor, G. W. 1934, Calculation and Interpretation of Analysis of Variance and Covariance (Ames, Iowa: Collegiate Press)
Thurstone, L. L. 1947, Multiple Factor Analysis (Chicago: Univ. Chicago Press)
van Zadelhoff, G.-J., van Dishoeck, E. F., Thi, W.-F., & Blake, G. A. 2001, A&A, 377, 566
White, G. J., & Sandell, G. 1995, A&A, 299, 179

Paola D'Alessio: Centro de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, Campus Morelia, Apartado Postal 3–72, 58090, Morelia, Michoacán, México (p.dalessio@astrosmo.unam.mx).
Itziar de Gregorio-Monsalvo: European Southern Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile (idegrego@eso.org).
José F. Gómez: Instituto de Astrofísica de Andalucía (CSIC), Apartado 3004, E-18080 Granada, Spain (jfg@iaa.es).