

Modelling Data in CASA

Thomas Möller¹
 Peter Schilke¹
 Michiel Hogerheijde^{2,3,4}
 Ian Stewart^{2,3}
 Reinhold Schaaf^{5,6}
 Daniel Harsono^{2,3}

¹ University of Cologne, Germany

² Leiden University, the Netherlands

³ Allegro ALMA Regional Centre (ARC) Node, Leiden, the Netherlands

⁴ University of Amsterdam, the Netherlands

⁵ Bonn University, Germany

⁶ German ARC Node, Argelander Institute for Astronomy, Bonn University, Germany

The Common Astronomy Software Applications (CASA) package provides a powerful tool for post-processing Atacama Large Millimeter/submillimeter Array (ALMA) and Karl G. Jansky Very Large Array (VLA) observations, but contains only rudimentary functions for modelling the data. In order to derive physical parameters as well as information on the location and the kinematics of the emitting gas, modelling of the observed data is inescapable. Such modelling can take the form of inference — where physical parameters are inferred from the data — or forward modelling — where model calculations are used to produce synthetic observations for comparison with data. We present two interfaces that allow the use of modelling tools for both flavours from within CASA: the eXtended CASA Line Analysis Software Suite (XCLASS); and Adaptable Radiative Transfer Innovations for Submillimeter Telescopes (ARTIST).

XCLASS

XCLASS¹ (Möller, Endres & Schilke, 2017) is a full message passing interface (MPI) parallelised toolbox for CASA (McMullin, et al., 2007), providing new functions for modelling interferometric and single-dish data. It produces physical parameter fits for all molecules in a dataset. This also allows line identification, but also provides much more information.

In contrast to many other radiative transfer programmes, XCLASS always takes all lines of a species in a given frequency range into account, which reduces the risk of misassignment as a result of blends, and allows for the robust identification of species. In order to derive physical parameters, XCLASS models a spectrum by solving the radiative transfer equation for an isothermal object in one dimension (called a detection equation) assuming local thermodynamic equilibrium (LTE). XCLASS is designed to describe line-rich sources which often have high densities, so LTE is a reasonable approximation. Additionally, a non-LTE (NLTE) description requires collision rates which are only available for a few molecules. Molecular data required by XCLASS are taken from an embedded SQLite3 database containing entries from the Cologne Database for Molecular Spectroscopy² (CDMS; Endres et al., 2016, Müller et al., 2005) and NASA Jet Propulsion Laboratory³ (JPL; Pickett et al., 1998) using the Virtual Atomic and Molecular Data Centre⁴ (VAMDC) portal.

XCLASS is able to model a spectrum with an arbitrary number of molecules, where the contribution of each molecule is described by an arbitrary number of components (see Figure 1). Components can be identified as spatially distinct sources such as clumps, hot dense cores, colder envelopes or outflows, and can

usually be distinguished by different radial velocities. They do not interact with each other radiatively but are superimposed in the model. Each component is described by the source size, the temperature, the column density, and the velocity width and offset, all of which have to be defined by the user in an input file. The Splatalogue^a syntax for molecule names can also be employed, to provide compatibility with the rest of CASA.

MAGIX

Owing to the large number of input parameters required by XCLASS, it is essential to use a powerful MPI parallelised optimisation package to achieve a good description of the observational data. Therefore, XCLASS contains the MAGIX package⁵ (Möller et al., 2013), which is a model optimiser that provides an interface between existing codes and an iteration engine. The package attempts to minimise deviations of the model results from observational data using a variety of algorithms, including swarm algorithms to find global minima, thereby constraining the values of the model parameters and providing corresponding error estimates. Many other model programmes can be combined with MAGIX to explore their parameter space and find the set of parameter values that best fits observational data.

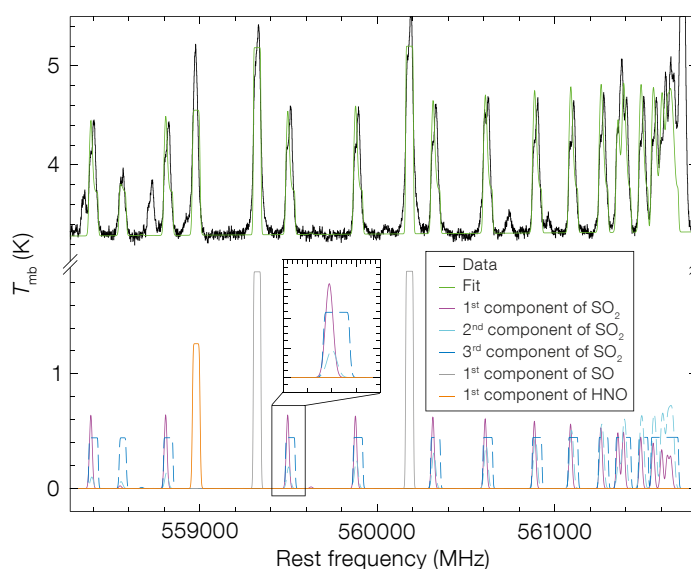


Figure 1. The myCLASS function was used to model HIFI data of Sgr B2(M) (black) using SO₂ (with three different components), SO (with one component), and HNO (with one component). The intensities of each component are shown in the bottom half (Möller et al., 2013).

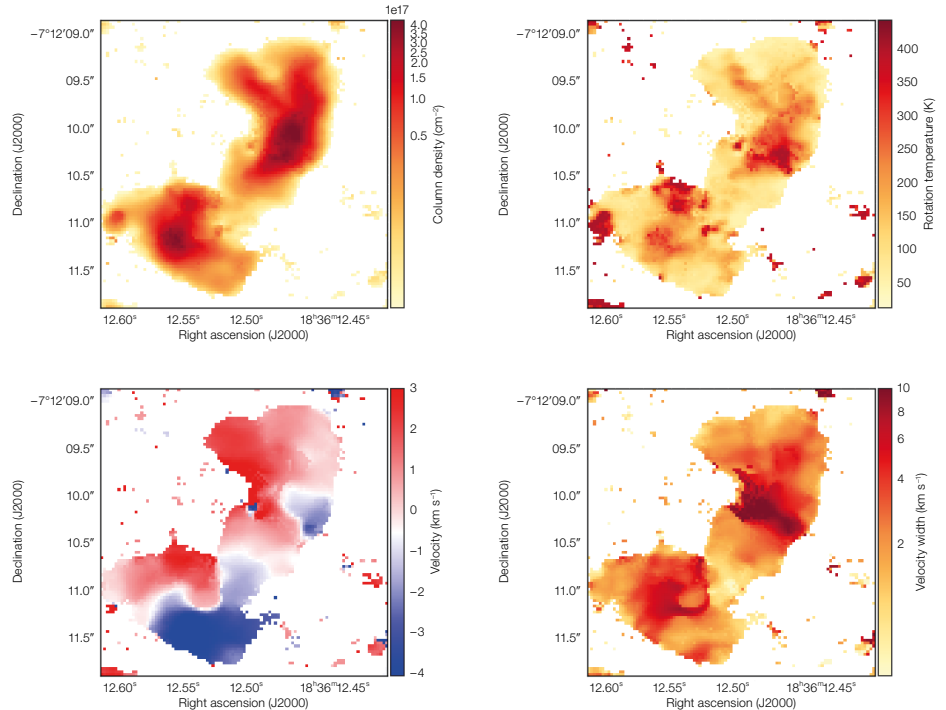
New functions for CASA

XCLASS provides, among other things, two functions (`myXCLASSFit` and `myXCLASSMapFit`) for CASA, which provide the option of fitting the input parameters to observational data. The `myXCLASSFit` function can be used to fit single spectra, i.e., to fit multiple frequency ranges simultaneously in multiple files, and it returns the optimised input file and the corresponding modelled spectra.

In addition to that, the XCLASS interface contains the `myXCLASSMapFit` function, which fits one or more complete (FITS) data cubes. For this, the `myXCLASSMapFit` function reads in the data cube(s), extracts the spectra for each pixel and fits each of these spectra separately. It offers the possibility of limiting the fit to certain frequency ranges of a spectrum and to a user defined region of the cube(s). At the end of the whole fitting procedure, the `myXCLASSMapFit` function creates FITS images for each free parameter of the best fit, where each pixel corresponds to the value of the optimised parameter taken from the best fit for that pixel (see Figure 2). Furthermore, the `myXCLASSMapFit` function creates FITS cubes for each fitted data cube, where each pixel contains the modelled spectrum.

Some applications of this include temperature maps, as well as first and second moment maps, which are based on the simultaneous fitting of many lines, and are fairly robust against line confusion and blending of single lines — this can be a severe issue in many ALMA datasets with line-rich sources.

Finally, XCLASS provides an automatic line identification function to identify species in the given spectra. The automatic line identification (`LineID`) function consists of two parts. The function starts by determining all molecules that have at least one transition within the user-defined frequency range(s). The `LineID` function then performs single molecule fits to calculate the contribution of each molecule. If a molecule covers a defined fraction of the spectrum the molecule is provisionally identified and the corresponding optimised input file is appended to a so-called overall input file, which



describes the contribution of all of the identified molecules. After all single molecule fits are completed, the `LineID` function performs a final (global) fit, using the aforementioned overall input file to account for line blending effects.

ARTIST

The ARTIST⁶ software was developed with the aim of providing an easy entry-level modelling suite for star formation and asymptotic giant branch (AGB) studies with submillimetre telescopes (Padovani et al., 2012). It represents a different approach to source modelling than the one described above. In this one, the user starts with a description of the astrophysical object in question and produces observables for comparison with actual data or as predictions for ALMA proposals. ARTIST combines nine pre-defined, parametrised descriptions of models of star-forming cores and circumstellar disks with a NLTE molecular excitation tool (Line Modelling Engine [LIME]⁷; Brinch & Hogerheijde, 2010). It takes molecular collision rates from the Leiden Atomic and Molecular (LAMDA) database⁸ (Schöier et al., 2005) and it outputs FITS cubes of the molecular line emission. These FITS cubes can then

Figure 2. Example of parameter maps created by the `myXCLASSMapFit` function using nine transitions of CH_3OCHO simultaneously, taken from an ALMA dataset of the core of G24.78 (Moscadelli et al., 2018).

be further processed, for example using CASA to simulate the response of a single-dish telescope or interferometric array.

In its original form, ARTIST employed a graphical user interface to set the model parameters. The new release of ARTIST allows users to set the model parameters directly from the CASA command interface, in the same keyword-driven fashion as other CASA tools, to enable new users to start exploring realistic astrophysical models more easily. It does this by providing two CASA tasks called `limesolver` and `raytrace`.

The source models in ARTIST

ARTIST contains nine popular astrophysical models: the Bonnor-Ebert sphere (Ebert, 1955; Bonnor, 1956); self-similar collapse of an isothermal sphere (Shu, 1977); models of collapsing and rotating or magnetised cloud cores (Ulrich, 1976; Li & Shu, 1996; Allen, Li & Shu, 2003; Mendoza, Tejada & Nagel, 2009);

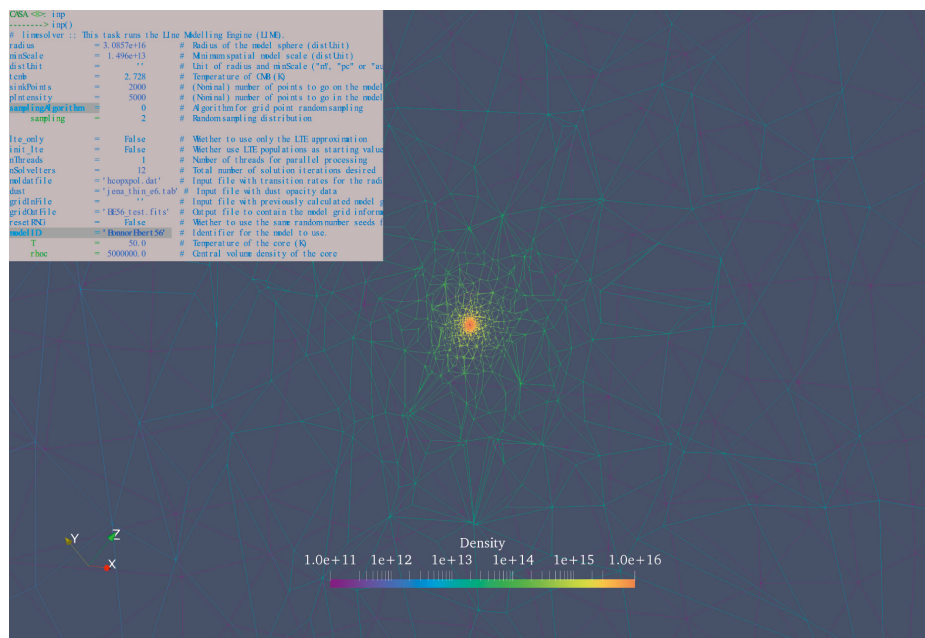


Figure 3. Screen shot of the ARTIST CASA task limesolver with a model setup for a Bonnor-Ebert sphere. In the background, the resulting LIME grid is shown.

Post-processing the ARTIST FITS cubes

The next step in the simulation is provided by the ARTIST CASA task raytrace, which calculates the emission on the sky for a chosen distance to the object and viewing orientation. The user sets the pixel size, field of view, spectral resolution and number of spectral channels, as well as the molecular transition of interest. Typical run times are of the order of a few minutes for a full FITS cube.

The resulting FITS cube is fully compliant with CASA, and can be used as input for further processing. Figure 4 shows the results of a calculation of a Bonnor-Ebert sphere, after running the FITS output through simobserve. The resulting image cube can be compared to actual ALMA observations to judge the similarity of the model to the data, or it can be used to make predictions for ALMA proposals.

It is important to remember that the models included in ARTIST are not fully self-contained. Some models include both a density and a temperature structure; others may only include a density

protoplanetary discs (Chiang & Goldreich, 1997; Dullemond, Dominik & Natta, 2001); and the circumstellar envelopes of evolved stars (Mamon, Glassgold & Huggins, 1988). In addition to these pre-defined models, specification of physical models by tabulated data is possible. The CASA task limesolver allows the user to select the source model and set its parameters. An example of setting the source model parameters through CASA keywords is given in Figure 3.

task. Depending on the size of the grid, the number of energy levels of the molecule, and whether LTE (fast) or NLTE (slower) is requested, the time it takes for the calculation to finish varies between a few minutes to many hours on typical high-end CPUs. The calculation is parallelised and can be run on multiple cores. On completion, LIME will write a file with the molecular excitation at each grid point, and return the CASA prompt.

The molecular excitation and radiation transfer inside ARTIST: LIME

Within limesolver, the user also provides the necessary information to solve the molecular excitation. To do this, ARTIST uses the LIME code. LIME is a versatile NLTE-accelerated Monte Carlo code that employs flexible gridding (Brinch & Hogerheijde, 2010), and comes packaged with ARTIST. The information that LIME needs includes the molecule of interest, the collision rate file from the LAMBDA database, any dust properties, the number of grid points and iterations, and whether LTE or NLTE calculations are requested. Once all of these are defined, the calculation starts. The progress of LIME can be monitored via the CASA logger just like any other CASA

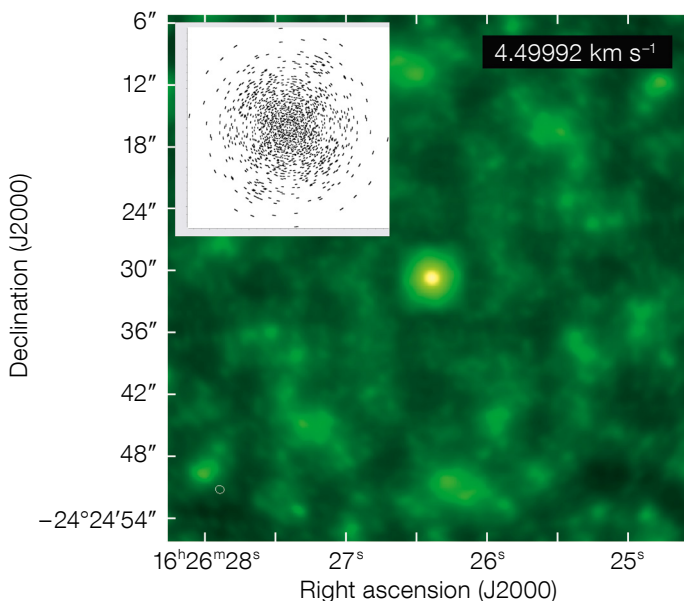


Figure 4. Output from the ARTIST CASA tasks limesolver and raytrace, after postprocessing through simobserve. The inset shows the (u, v) coverage. The model is the same as in Figure 3; a Bonnor-Ebert sphere, and a single velocity channel is shown.

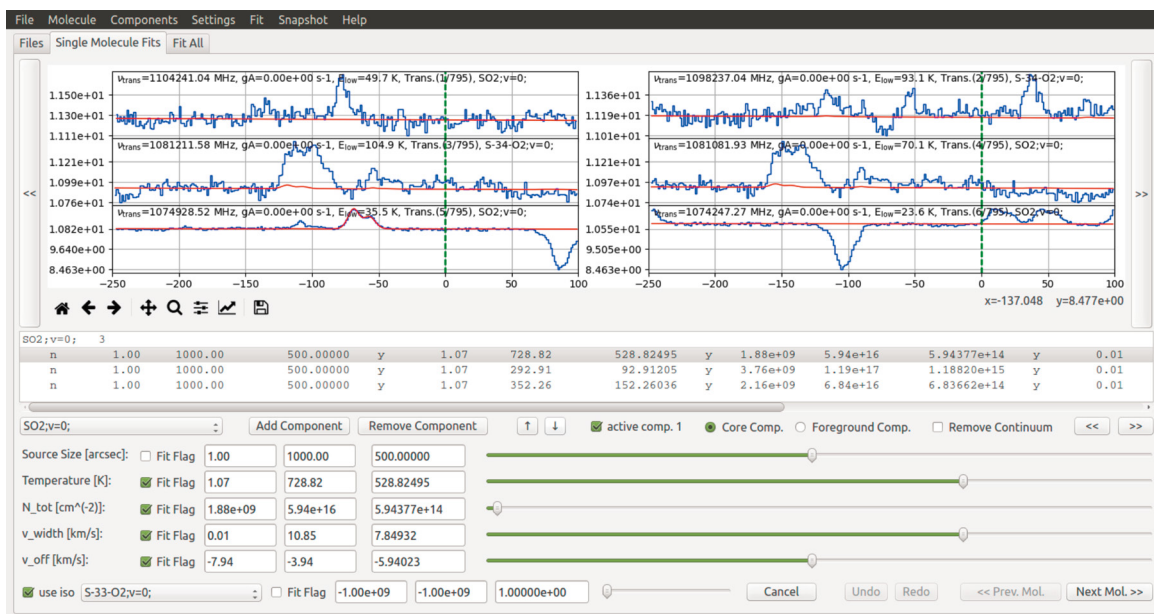


Figure 5. Currently developed user-friendly XCLASS GUI.

structure, leaving it to the user to provide a simple parameterised temperature field. However, the code does not check if this temperature field is consistent with the luminosity of any central star and the density structure of the source. ARTIST does not aim to provide such complete source modelling. Instead, ARTIST aims to provide easy access to popular models that allow users to quickly explore ALMA observables. Further, more detailed modelling requires more realistic source models and the use of the stand-alone LIME package. However, ARTIST would be the code of choice for users who want to quickly obtain a rough indication of the observables.

Outlook

In one of the next XCLASS releases, we will include a user-friendly graphical user interface (GUI), which is currently under development (see Figure 5 for a prototype), to lower the threshold of training required for new users of XCLASS. Additionally, we are working on several extensions, including NLTE, radio recombination lines, (in LTE and NLTE), non-thermal continuum, non-Gaussian line shapes (Lorentz, Voigt and Horn), local overlap of neighbouring lines, and a better source description. Furthermore, we will include a new XCLASS function called CubeFit, which offers the possibility of describing

a data cube using a physical model, and fitting the model parameters. Further developments for the LIME radiation transfer engine inside ARTIST will be implemented as they become available. Additional astrophysical models will be included in ARTIST, and we encourage users to make suggestions about which models are the most useful for their research.

Acknowledgements

We acknowledge support from VBF/BMBF Projects 05A11PK3 and 05A14PK1 for the German ARC Node, from the Netherlands Organisation for Scientific Research (NWO) for the Allegro ARC node, and from the ESO ALMA development project 56787/14/60579/HNE.

References

- Allen, A., Li, Z.-Y. & Shu, F. H. 2003, *ApJ*, 599, 363
 Bonnor, W. B. 1956, *Zeitschrift für Astrophysik*, 39, 143
 Brinch, C. & Hogerheijde, M. R. 2010, *A&A*, 523, A25
 Chiang, E. I. & Goldreich, P. 1997, *ApJ*, 490, 368
 Dullemond, C. P., Dominik, C. & Natta, A. 2001, *ApJ*, 560, 957
 Ebert, R. 1955, *Zeitschrift für Astrophysik*, 36, 222
 Endres, C. P. et al. 2016, *J. Mol. Spectrosc.*, 327, 95
 Li, Z.-Y. & Shu, F. H. 1996, *ApJ*, 472, 211
 Mamon, G. A., Glassgold, A. E. & Huggins, P. J. 1988, *ApJ*, 328, 797
 McMullin, J. P. et al. 2007, *Astronomical Data Analysis Software and Systems XVI (ASP Conf. Ser. 376)*, ed. Shaw, R. A., Hill, F. & Bell, D. J. (San Francisco, CA: ASP), 127
 Mendoza, S., Tejada, E. & Nagel, E. 2009, *MNRAS*, 393, 579

- Möller, T., Endres, C. & Schilke, P. 2017, *A&A*, 598, A7
 Möller, T. et al. 2013, *A&A*, 549, A21
 Müller, H. S. P. et al. 2005, *J. Mol. Struct.*, 742, 215
 Padovani, M. et al. 2012, *A&A*, 543, A16
 Pickett, H. M., Poynter, R. L. & Cohen, E. A. 1998, *J. Quant. Spectr. Rad. Transf.*, 60, 883
 Schöier, F. L. et al. 2005, *A&A*, 432, 369
 Shu, F. H. 1977, *ApJ*, 214, 488
 Ulrich, R. K. 1976, *ApJ*, 210, 377

Links

- ¹ XCLASS software including a manual is available at: <https://xclass.astro.uni-koeln.de/>
² Cologne Database for Molecular Spectroscopy (CDMS): <http://www.cdms.de>
³ Molecular spectroscopy at JPL: <http://spec.jpl.nasa.gov>
⁴ Virtual Atomic and Molecular Data Centre: <http://www.vamdc.eu>
⁵ The MAGIX package can be downloaded from <https://magix.astro.uni-koeln.de>
⁶ The ARTIST package can be found at: <http://www.alma-allegro.nl/artist/>
⁷ LIME: <http://github.com/lime-rt/lime>
⁸ LAMDA: <http://www.strw.leidenuniv.nl/~moldata>

Notes

- ^a The name splatalogue is derived from “spectral line catalogue” and is a database for astronomical spectroscopy that contains information on nearly six million spectral lines.