Investigating high-mass star formation with ALMAGAL

Robust temperatures and luminosities of cores in a thousand high-mass cluster-forming regions

Beth Jones

Gary Fuller, Georgie Stroud, Tianwei Zhang, Thomas Möller, Álvaro Sánchez-Monge, Judith Beck, Eleonore Dann, ALMAGAL team

PIs: Sergio Molinari (I), Peter Schilke (D), Cara Battersby (US), Paul Ho (Tw)

ALMAGAL

DFG

Bundesministerium für Bildung und Forschung

Universität zu Köln

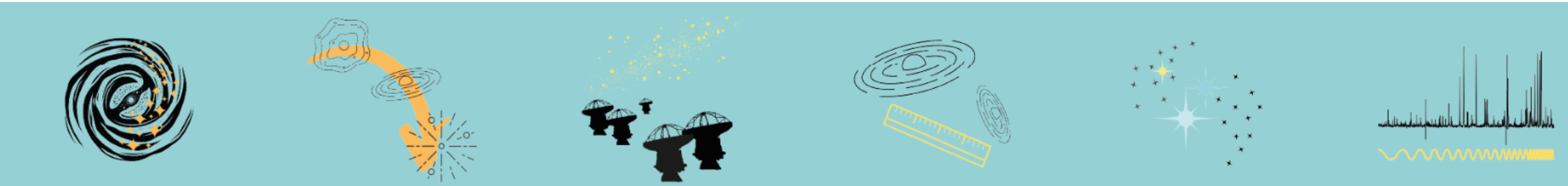
ph1 Universität zu Köln

Background image credit: G. Stroud

The ALMAGAL Survey

Cycle 7 ALMA Large Program
PIs: Sergio Molinari (I), Peter Schilke (D), Cara Battersby (US), Paul Ho (Tw)

The ALMAGAL Survey



1017 high-mass star forming clumps

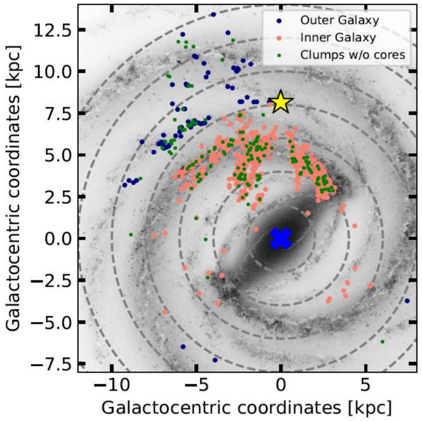
Luminosity-to-mass ratio from 0.05-500 $L_\odot/M_\odot$

3 configurations of ALMA in band 6 (1mm)

Spatial resolution of ~1000 AU

Continuum sensitivity of 0.1 mJy (0.3 $M_\odot$)

10,000 spectral channels at 1.5, 0.3 km/s resolution



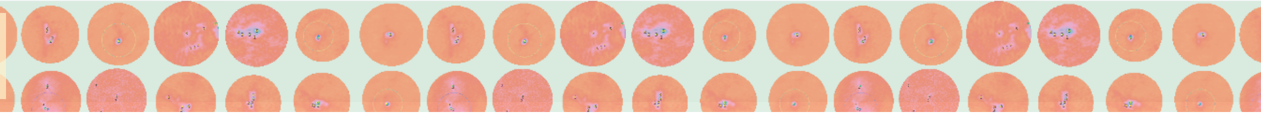
Clumps selected from Hi-GAL / RMS

- $M > 500\ M_\odot$
- $\Sigma > 0.1$ g cm^{-2}
- $F_{350\mu m} > 5$Jy
- $\delta < 0^o$ and $D < 7.5$ kpc

Technical working group lead:
Álvaro Sánchez-Monge

Joint deconvolution of all configurations
3 years
12M CPU hours @ JSC
450TB FITS products

Come and talk to me about array combination!



The ALMAGAL Survey

Cycle 7 ALMA Large Program
PIs: Sergio Molinari (I), Peter Schilke (D), Cara Battersby (US), Paul Ho (Tw)

The ALMAGAL Survey

- 1017 high-mass star forming clumps
- Luminosity-to-mass ratio from 0.05-500 $L_\odot/M_\odot$
- 3 configurations of ALMA in band 6 (1mm)
- Spatial resolution of ~1000 AU
- Continuum sensitivity of 0.1 mJy (0.3 $M_\odot$)
- 10,000 spectral channels at 1.5, 0.3 km/s resolution

Science goals

For a Galaxy-wide sample and as a function of evolution:

- Fragmentation and CMF
- Chemistry
- Clump-level SFH
- Intra-clump core dynamics
- Core(s)-clump feedback
- Outflows & Disks

Motivation for large program

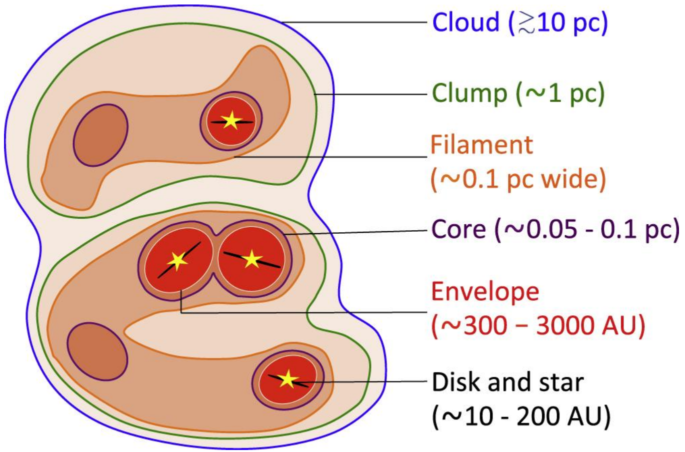
With a large, coherent data set, we can:

- Remove statistical biases
- Break degeneracies due to projection effects
- Determine means + spread and probability distributions
- Correlate with clump properties, environmental effects etc.
- Compare with simulations

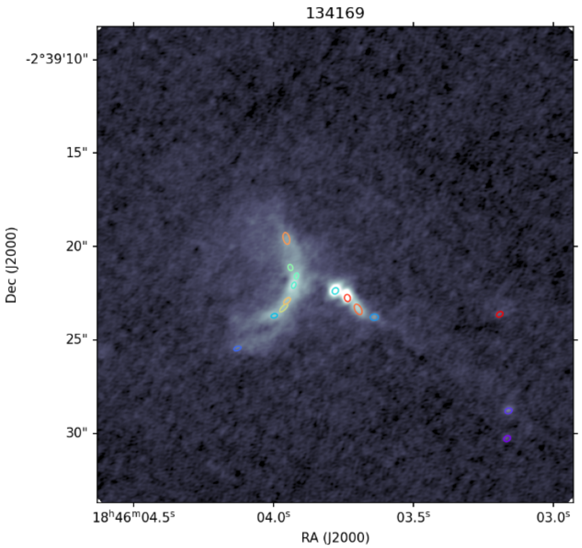
Definitions: what is a core?*

Alessandro Coletta et al. (in prep; Tuesday talk)

* **A** definition relevant for this talk, not **the** definition.



Pokhrel et al. (2018)



Clump = whole ALMA field of view/Hi-GAL compact source:

~pc scale

$N(H_2) \sim 10^5$ cm^{-3}

Core = compact sources:

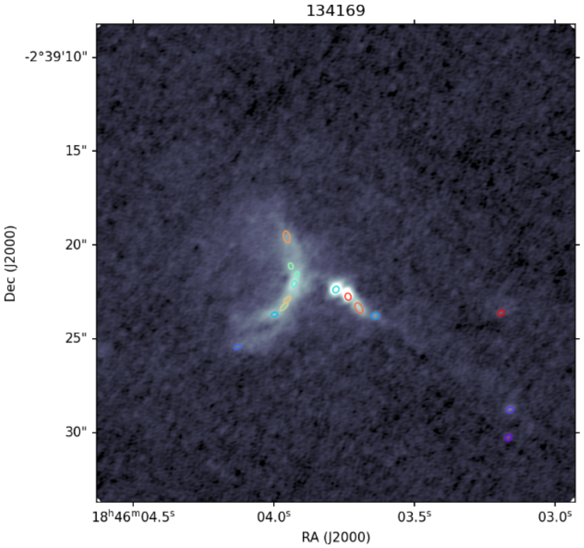
~ 1000 au scales in ALMAGAL

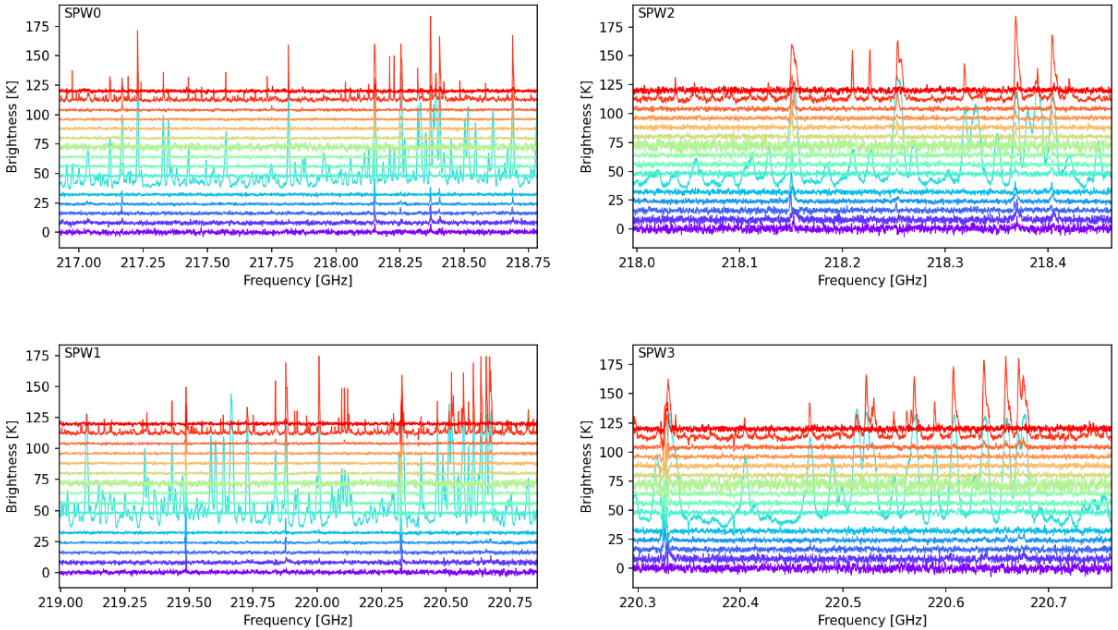
$N(H_2) \sim 10^6 - 10^9$ cm^{-3}

Definitions: what is a core?*

Alessandro Coletta et al. (in prep; Tuesday talk)

* **A** definition relevant for this talk, not **the** definition.





The motivation for temperature determination

$$
\begin{aligned}
M &= \frac{R_{\mathrm{dg}} F_{\mathrm{int}} d^2}{B_\nu(T)\kappa_{1.3}} \\
B_\nu(T) &= \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/k_BT} = 1} \approx \frac{2\nu^2 k_B}{c^2}T \\
\Rightarrow M &\propto \frac{F_{\mathrm{int}}}{\color{red}{T}} \\
L &= 4\pi {\color{red}{r^2}} \sigma_{\mathrm{SB}} {\color{red}{T^4}}
\end{aligned}
$$

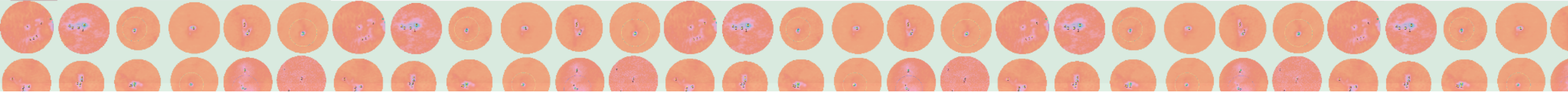
On **clump** scale, full SED can be measured from Hi-GAL (Herschel)

⇒ Luminosity is directly determined by integrating over SED

⇒ Mass can be determined since dust temperature is known from SED

On **core** scales neither the luminosities nor the dust temperatures can be measured directly

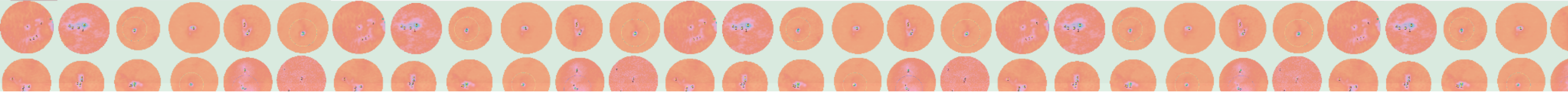
Proxy: **temperatures through lines**



Goal

Provide robust temperatures from spectral lines to reconstruct reliable mass, temperature, and luminosity distributions for embedded protostellar cores in a statistically significant sample of Galactic high-mass star forming clumps

- Uncertainty in **core mass functions,** luminosities for **radiative heating**
- Cores are not co-eval within a clump: **local evolution and heating is important**
- **Understanding cluster evolution:** age distribution, disruption by feedback
- **Comparison with simulations:** temperature/luminosity distribution of embedded cores



This talk: two parts

1. **The Spectral Line Temperature catalogue for ALMAGAL continuum cores**

   What I will present: the temperatures from CH_3CN toward 1000 detected protostellar cores; implied luminosities of the protostellar cores*

   What I want to convince you of: we **can** do advanced spectral line fitting for large samples; transferring between dust and gas is not straightforward; we really **shouldn't** use discrete components for cores

1. **Beyond core boundaries: structure recovery of high-mass protostellar core envelopes**

   What I will present: the proposed methods to move beyond cores into continuous functions; our prototype ML results; our ongoing supercomputing ML project

   What I want to convince you of: we **can** do better; the information **already available** in observations is sufficient to do better; we need **can** evolve our tools alongside our data availability

Part I: The ALMAGAL Spectral Line Core Temperature Catalogue

The ALMAGAL temperature tracers

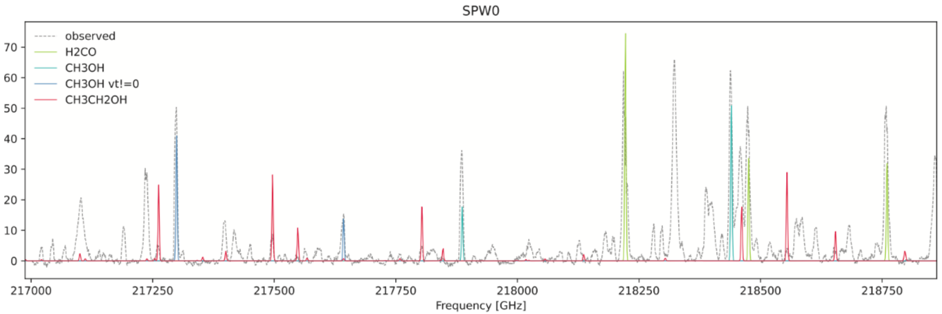
Come and talk to me about suitability of temperature tracers

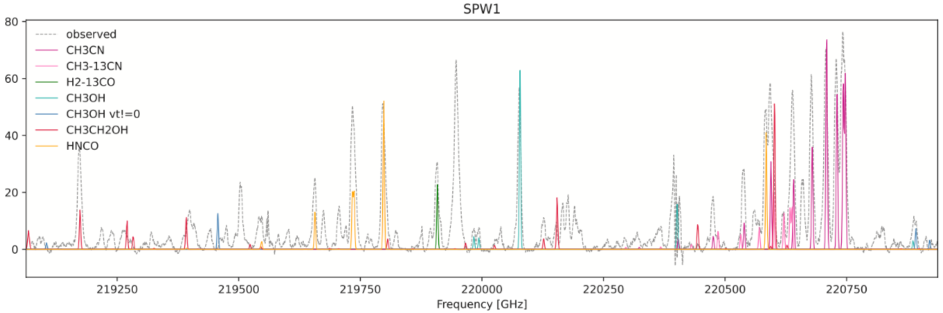
Temperature tracers

Methanol
CH_3OH
T_{ex} 35 - 735K

Formaldehyde
H_2CO, $H_2{}^{13}CO$
T_{ex} 10 - 70K

Methyl Cyanide
CH_3CN, $CH_3{}^{13}CN$
T_{ex} 60 - 920K





Blended species

Ethanol
CH_2CH_3OH
T_{ex} 65 - 1640K

Isocyanic Acid
HNCO
T_{ex} 47 - 1000K

The ALMAGAL temperature tracers

Come and talk to me about outflow identification in ALMAGAL

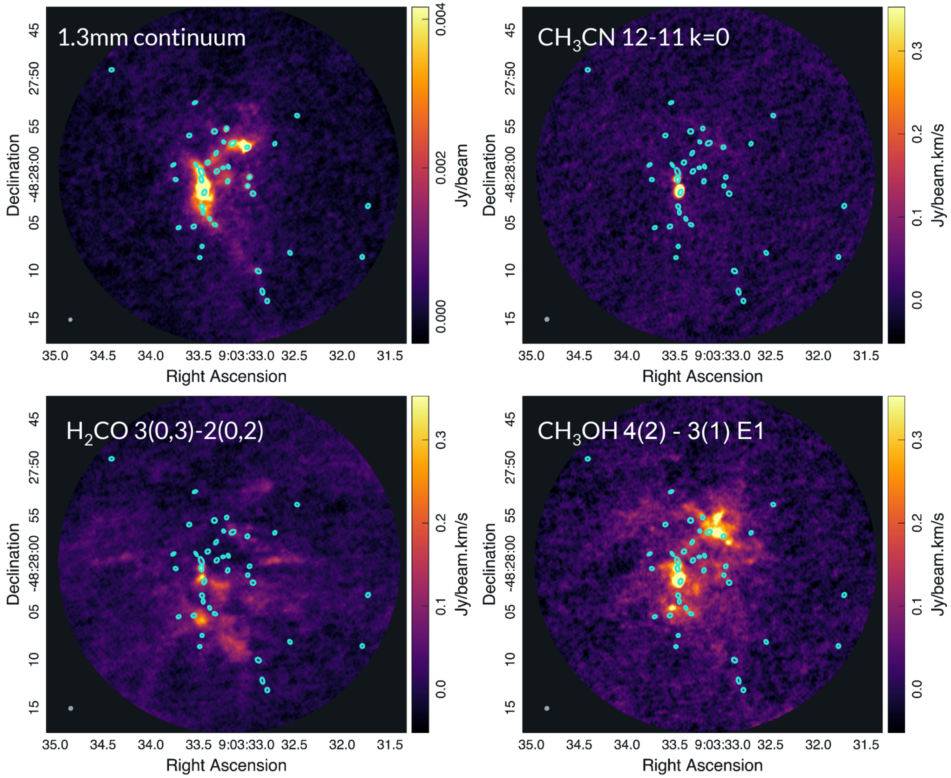
CH_3CN concentrated toward cores

H_2CO, CH_3OH more widespread*

- May peak toward cores in higher excitation energy lines

Use CH_3CN preferentially where available

ALMA-IMF Brouillet et al. (2022): CH_3CN core tracer, CH_3CCH outer layers



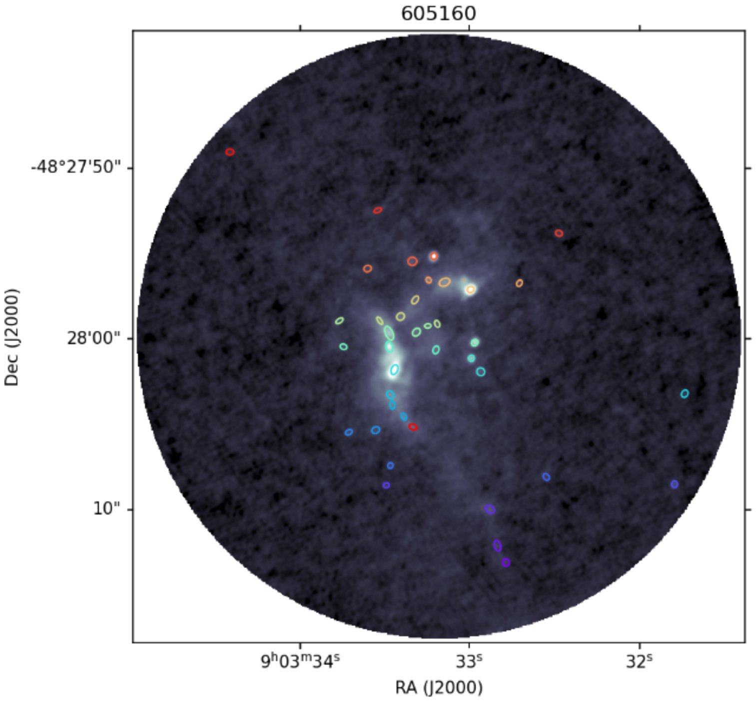
Moment maps and emission distributions:
Minnini et al. (in prep)

Jones et al. (in prep.)

XCLASS ALMAGAL pipeline

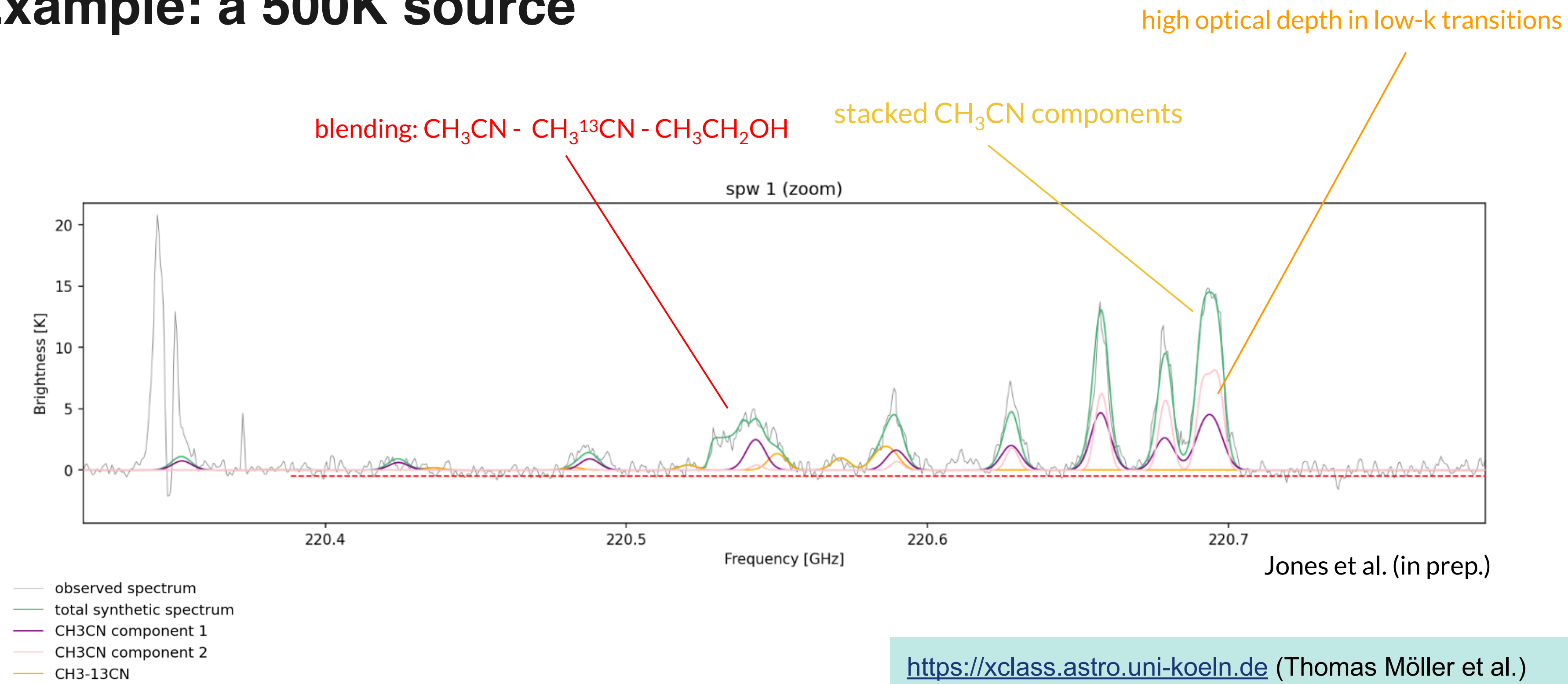
Continuum core catalogue: Coletta et al. (in prep)
Source velocities: M. Benedettini et al.; Molinari et al. (in prep)

- Spectra extracted for each core over FWHM ellipse
- **Simultaneous fitting of multiple species/components**
- **Opacity corrections and sub-beam source description**
- Multiple components:
  - Velocity, excitation, density, line-of-sight
- Accelerated fitting and optimisation
  - **Error estimation with MCMC**
- Links to most modern, routinely improved databases: CDMS, VAMDC



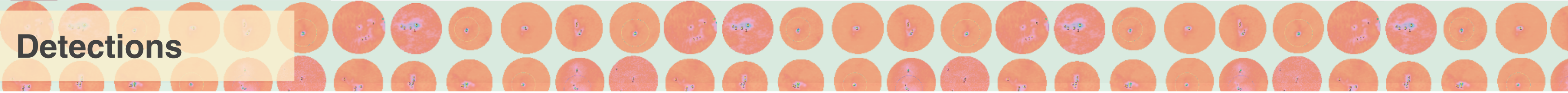
https://xclass.astro.uni-koeln.de (Thomas Möller et al.)

Example: a 500K source



Jones et al. (in prep.)

https://xclass.astro.uni-koeln.de (Thomas Möller et al.)

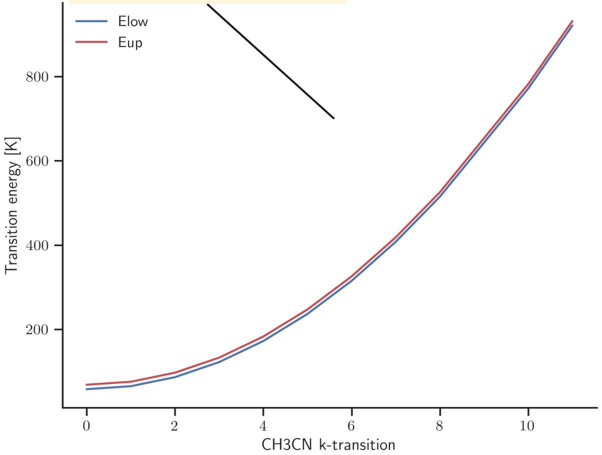


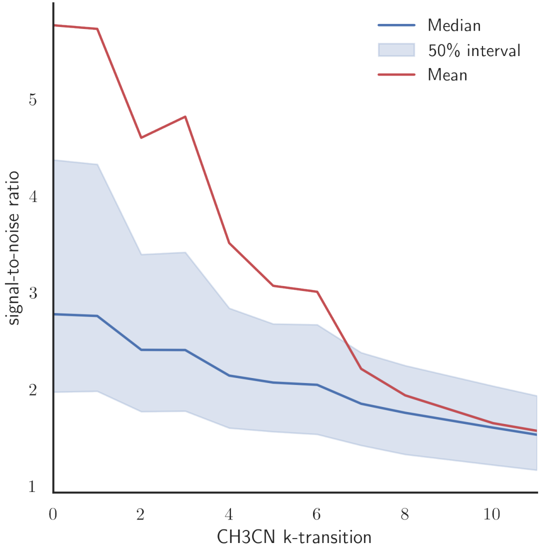
Detections

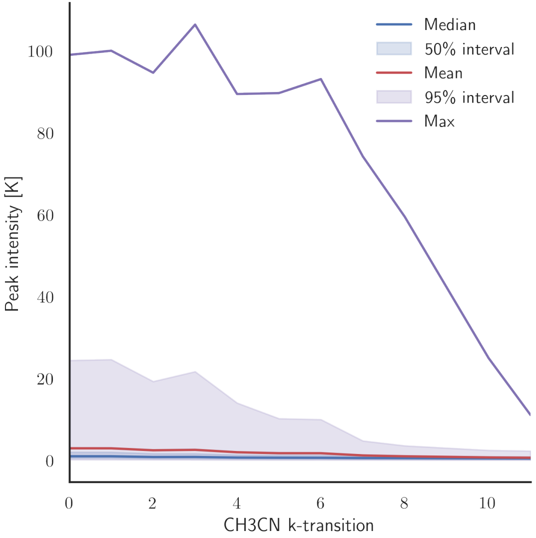
CH3CN detection statistics in plots: S/N and peak intensity of all sources

CH_3CN 12-11:

k = 0 : 60 K
k = 3 : 125 K
k = 6: 320 K
k = 8: 510 K







Jones et al. (in prep.)

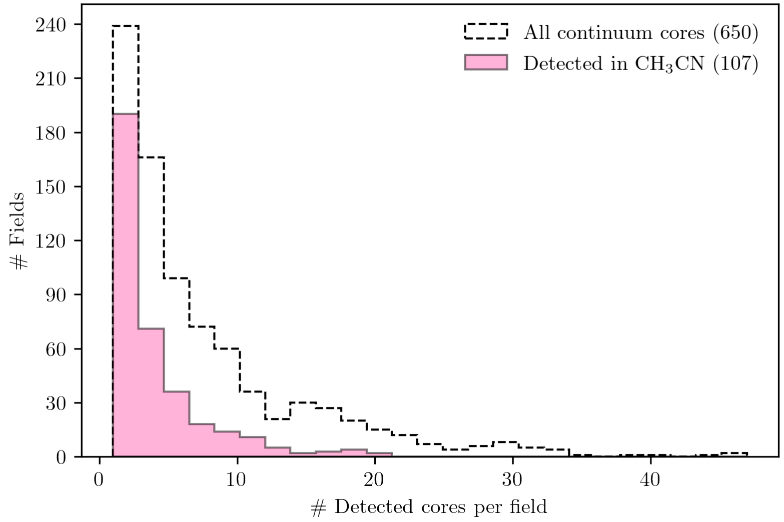
CH3CN detection statistics in numbers of cores

Of the 6286 cores extracted:

21.2% (1335 cores) are detected in CH_3CN above 5-sigma

Of the 837 fields with at least one continuum core extracted:

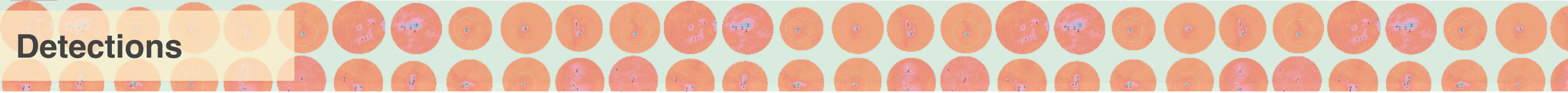
43% of all fields have at least one core detected in CH_3CN



Jones et al. (in prep.)

CH_3CN 12-11:

k = 0 : 60 K
k = 3 : 125 K
k = 6: 320 K
k = 8: 510 K



CH3CN detection statistics: which clumps?

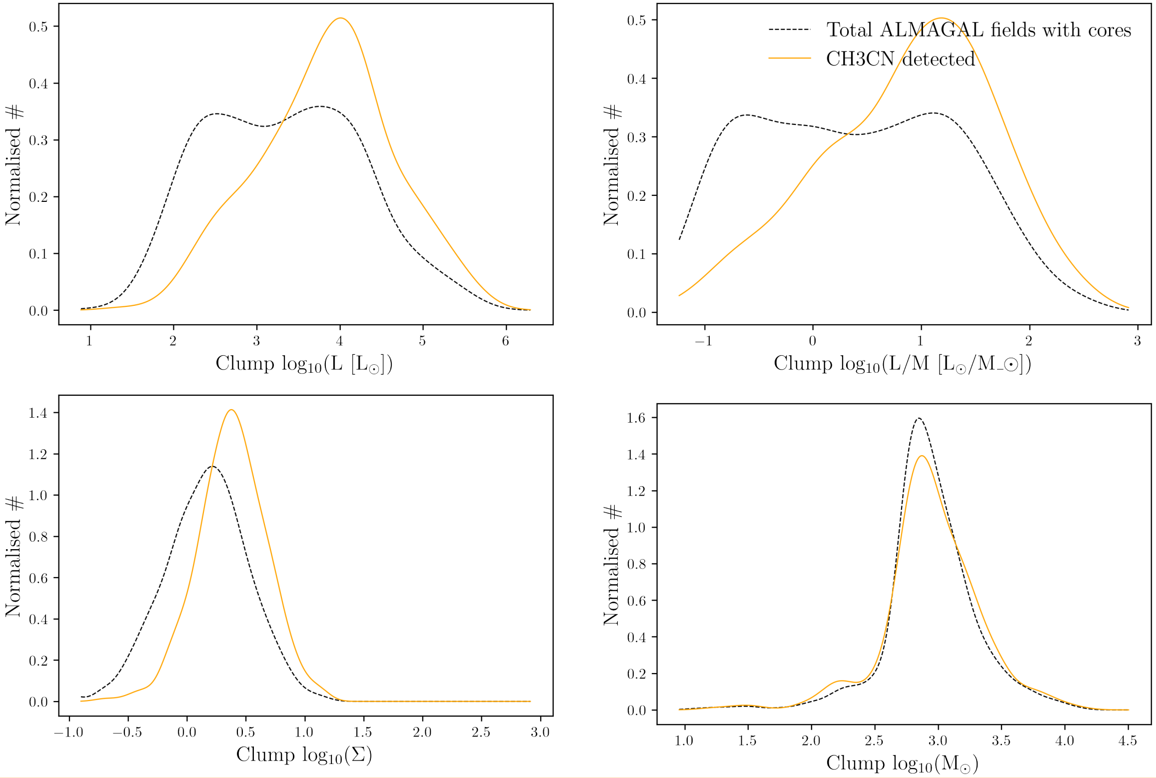
Jones et al. (in prep.)

Of the 837 fields with at least one continuum core extracted:

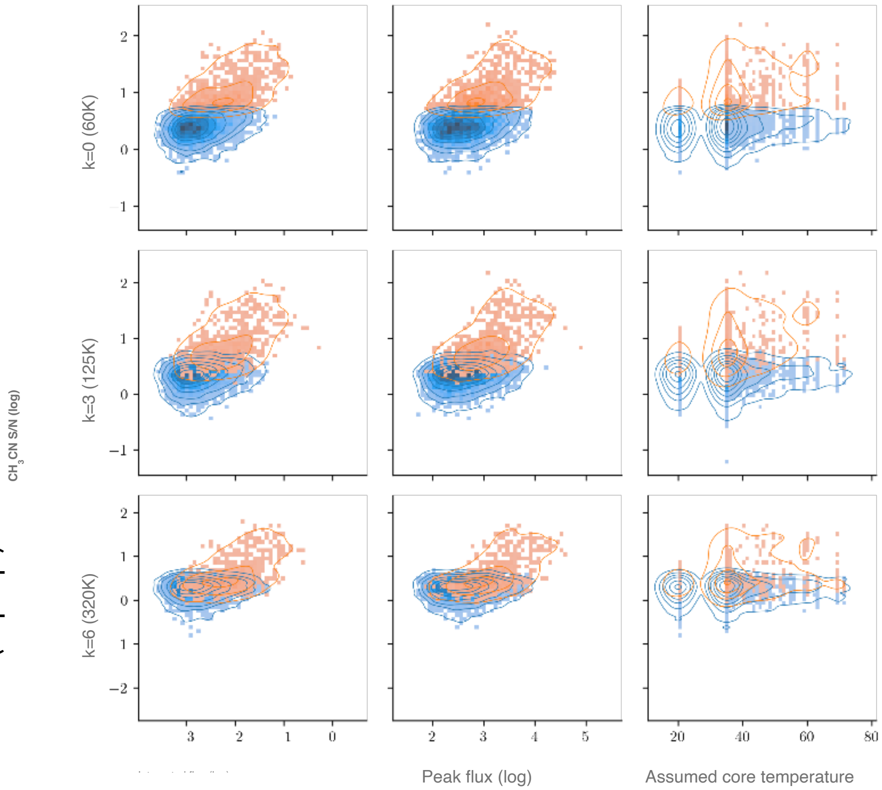
43% of all fields have at least one core detected in CH_3CN

Samples higher clump L, L/M within sample but **does not exclusively trace the more evolved sources**

No significant distinction in clump surface density, clump mass



CH3CN detection statistics: core continuum properties



Jones et al. (in prep.)

Red: detected
Blue: not detected

Of the 837 fields with at least one continuum core extracted:

43% of all fields have at least one core detected in CH_3CN

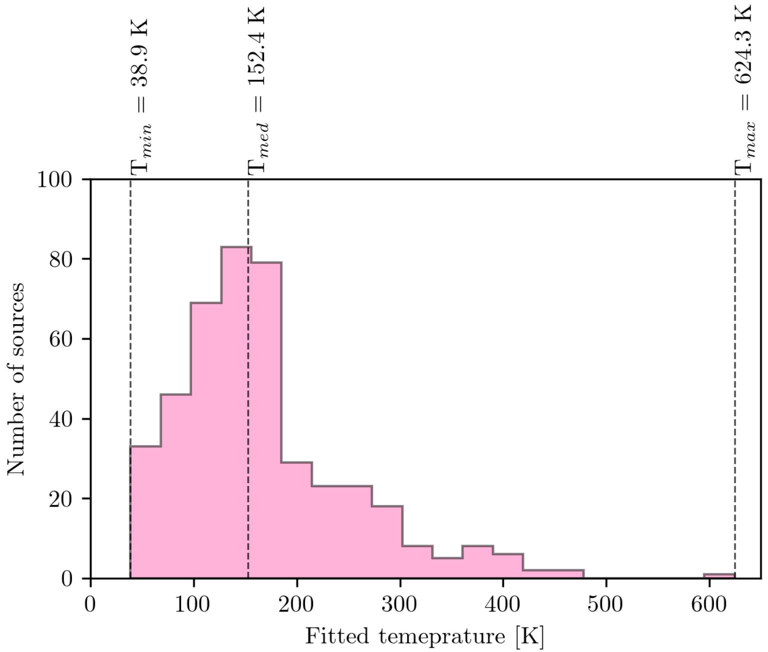
Selects brightest continuum cores but samples most of range except very faintest

No correlation between core temperature assigned from clump-scale assumptions with the S/N in any in CH_3CN transition

$$\Rightarrow M \propto \frac{F_{\rm int}}{T}$$

$$L = 4\pi r^2 \sigma_{\rm SB} T^4$$

Temperature distribution of cores: single component fits for >15-sigma sources



Minimum temperature of 39 K

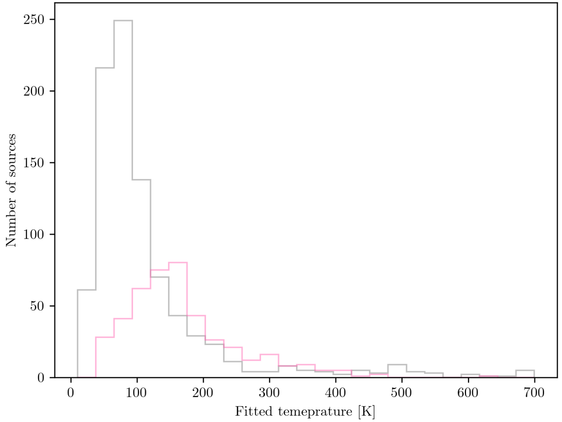
Maximum temperature of 624 K

Median temperature of 152 K

Most sources detected well in CH3CN are between 100 - 200 K with high temperature tail

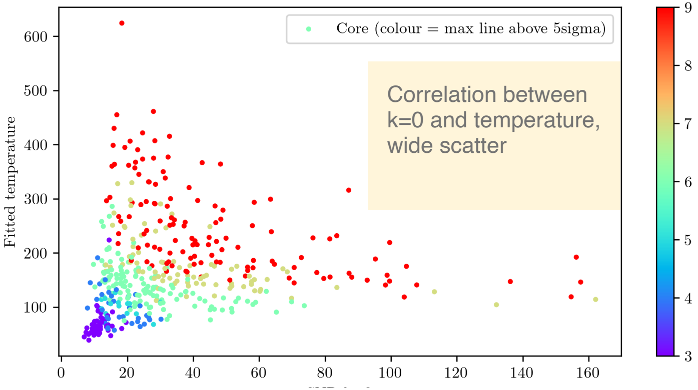
CORE: Gieser et al. (2021)
- 80 to 240K

Threshold on S/N set to exclude low signal to noise sources with large temperature uncertainties - limits provided for these sources

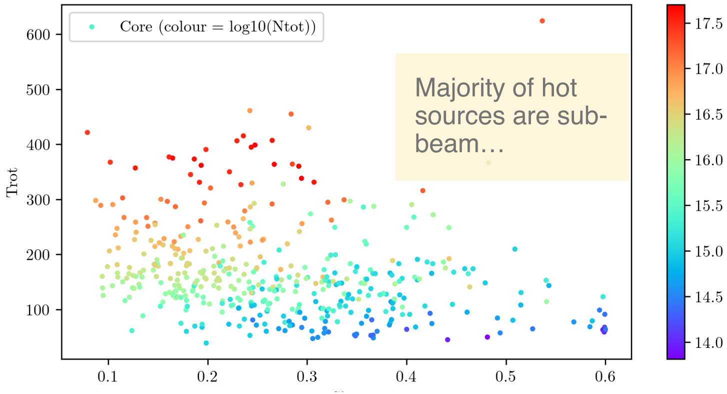


Jones et al. (in prep.)

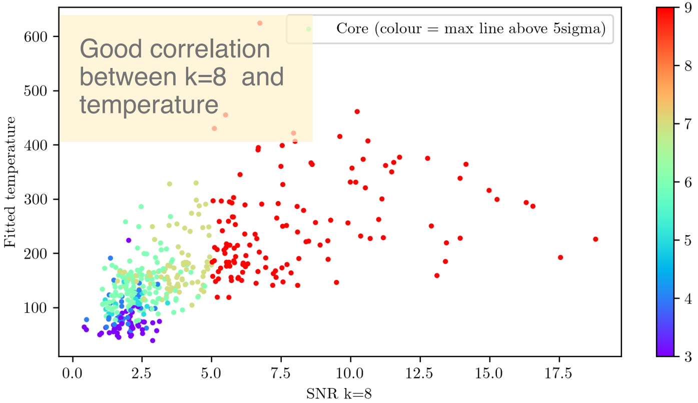
Temperature distribution of cores: single component fits for >15-sigma sources



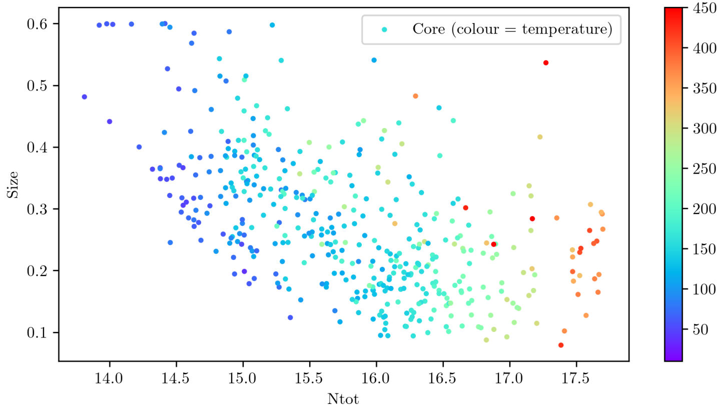
Correlation between k=0 and temperature, wide scatter



Majority of hot sources are sub-beam…



Good correlation between k=8 and temperature



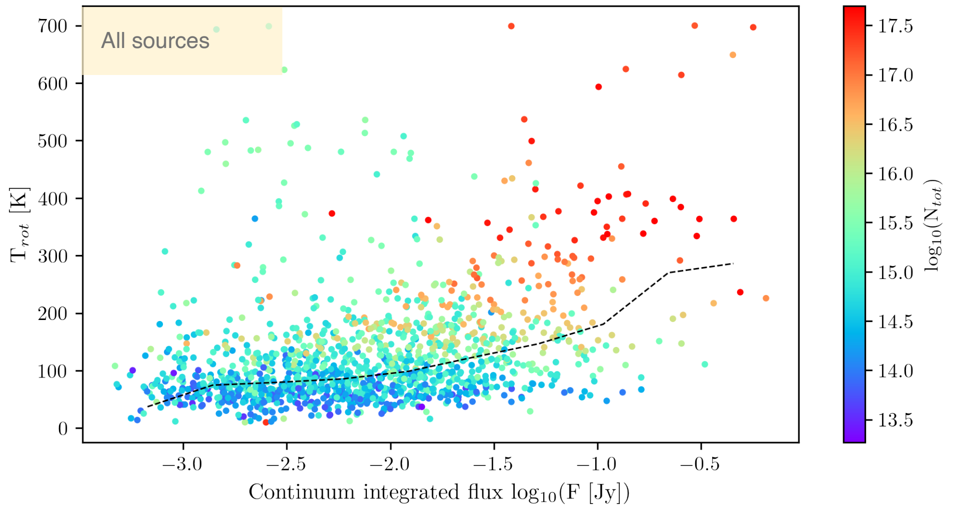
… which are also the highest column density sources

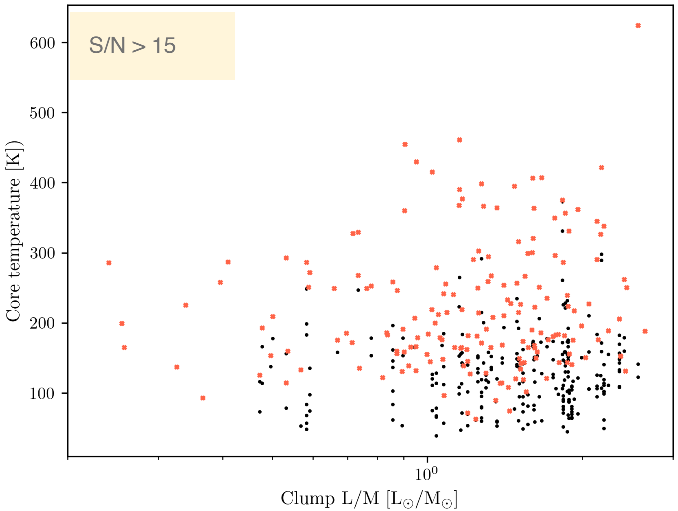
CH_3CN 12-11:

k = 0 : 60 K
k = 3 : 125 K
k = 6: 320 K
k = 8: 510 K

Jones et al. (in prep.)

Core temperature with continuum flux





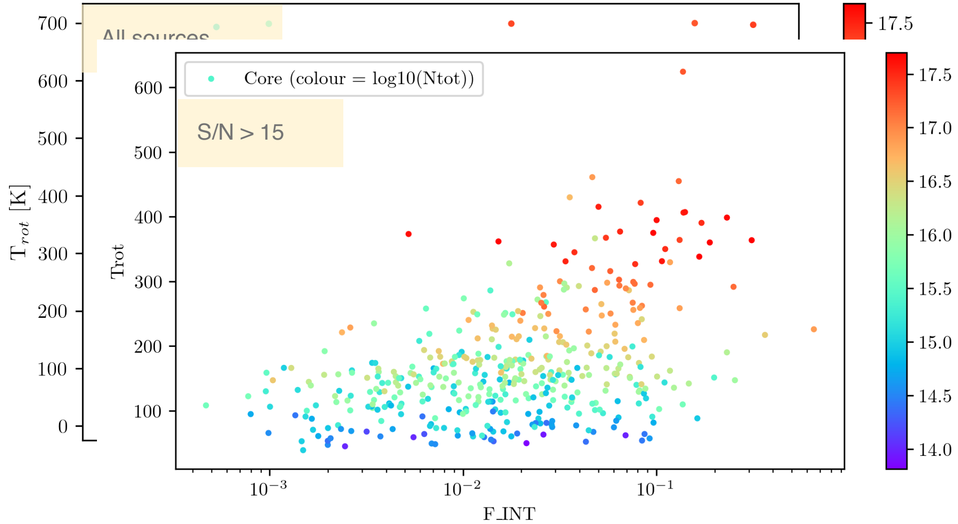
Jones et al. (in prep.)

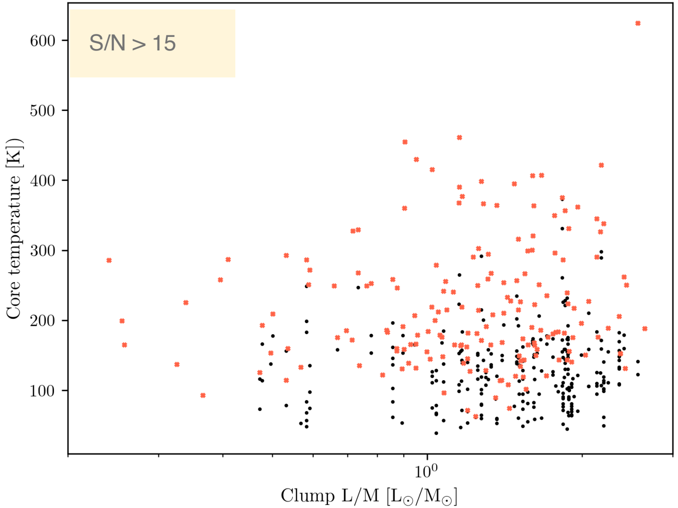
Correlation with continuum core integrated flux

Trend of increase in highest temperature core in clump with evolution **but not exclusively cold**

Also number of warm/hot cores

Core temperature with continuum flux





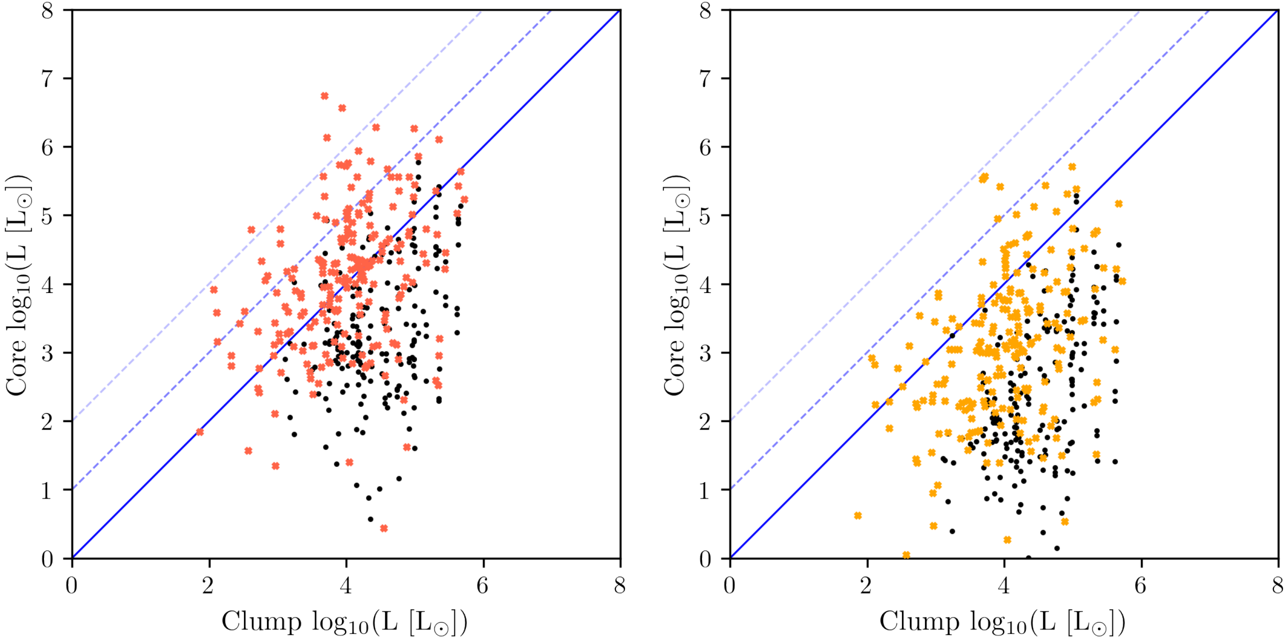
Jones et al. (in prep.)

Correlation with continuum core integrated flux

Trend of increase in highest temperature core in clump with evolution **but not exclusively cold**

Also number of warm/hot cores

Luminosities: as a function of clump luminosity



$$L = 4\pi r^2 \sigma_{SB} T^4$$

Sub-beam sources sizes only

Left: continuum source size

Right: fitted source size

Coloured points: brightest in each field

Blue lines (decreasing saturation): 1x, 10x, 100x

Jones et al. (in prep.) - or maybe not…

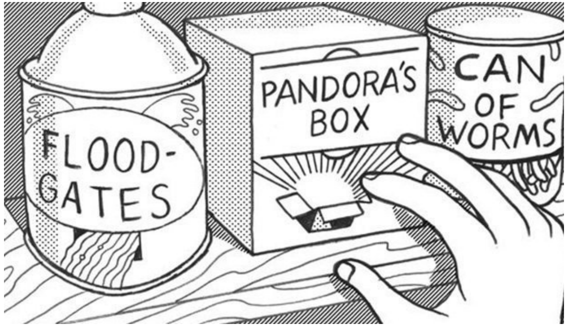
Solving the luminosity problem: transferring line temperatures to continuum dust temperatures in cores

Multi-component fits with well-defined sub-beam source sizes

(B. Jones)

Minimum source size from assuming an optically thick surface

(P. Schilke)



Radiative transfer modelling

(1D: S. Molinari; 3D: P. Schilke/B. Jones with S. Kesselheim/F. Niesel @ FZJ)

Radial profiles from map fitting

(B. Jones)

Comparison with simulations incl. radiative feedback

(B. Zimmermann)

Measure sub-beam source sizes with high-resolution disk data

(PI A. Ahmadi, P. Schilke)

Other causes of discrepancy also addressed

Decoupling between dust and gas temperatures

Radiation leakage/absorption as a function of size scale

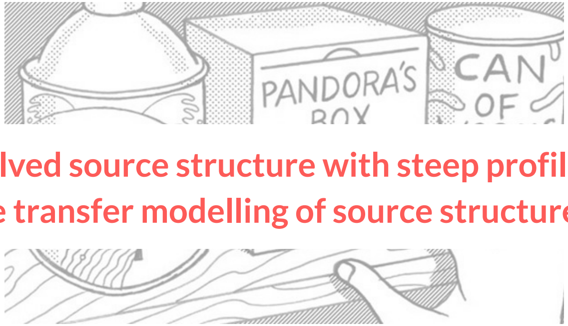
Extremely non-spherical geometries

Beyond 1D: calibrating lines and continuum

Solving the luminosity problem: transferring line temperatures to continuum dust temperatures in cores

Multi-component fits with well-defined sub-beam source sizes

(B. Jones)



Radiative transfer modelling

(1D: S. Molinari; 3D: P. Schilke/B. Jones with S. Kesselheim/F. Niesel @ FZJ)

“Averages” break down due to unresolved source structure with steep profiles and high opacities. Transferring between size scales requires radiative transfer modelling of source structure.

ressed

(P. Schilke)

Decoupling between dust and gas temperatures

Radiation leakage/absorption as a function of size scale

Radial profiles from map fitting

(B. Jones)

Comparison with simulations incl. radiative feedback

(B. Zimmermann)

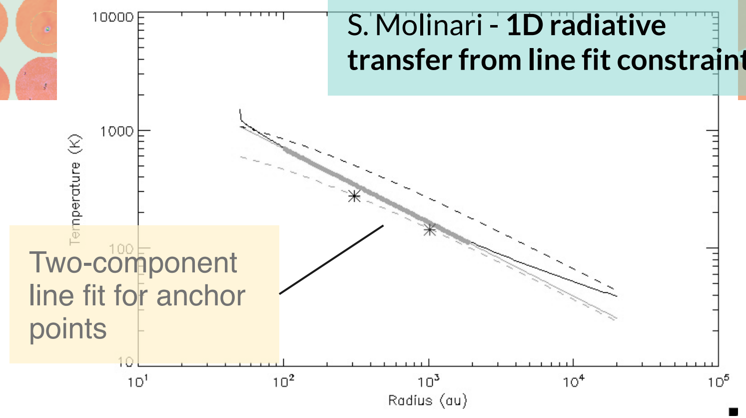
Measure sub-beam source sizes with high-resolution disk data

(PI A. Ahmadi, P. Schilke)

Extremely non-spherical geometries

Beyond 1D: calibrating lines and continuum

S. Molinari - **1D radiative transfer from line fit constraints**



1D radiative transfer modelling

Gives conversion between line temperature and size scale and mass-averaged temperature on continuum size scale

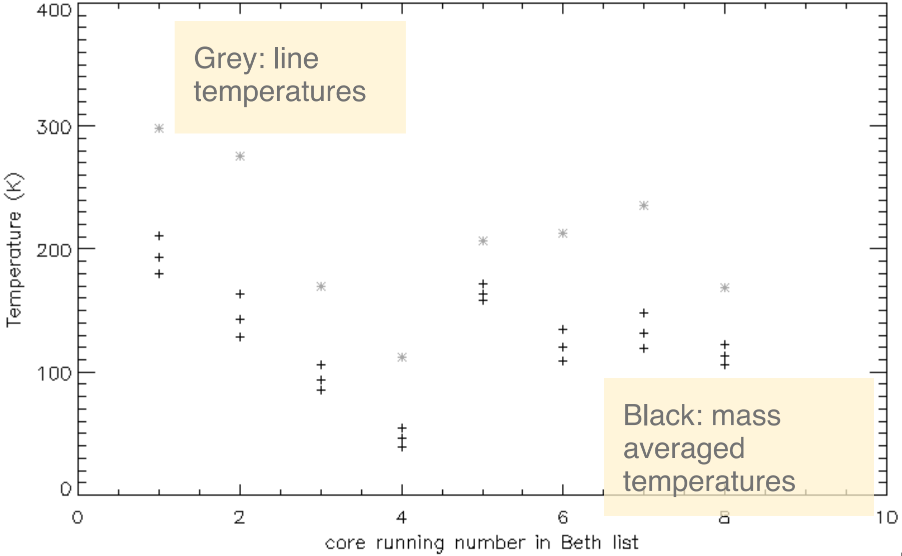
Can run grids and models for representative sources with realistic runtime

More sources available than mapping

Still relies on single LoS component fits (not true 3D for profiles)

Assumes that size scale and temperature returned from discrete component fit are connected

Requires at least one component to be optically thick with well defined source size



Temperature maps

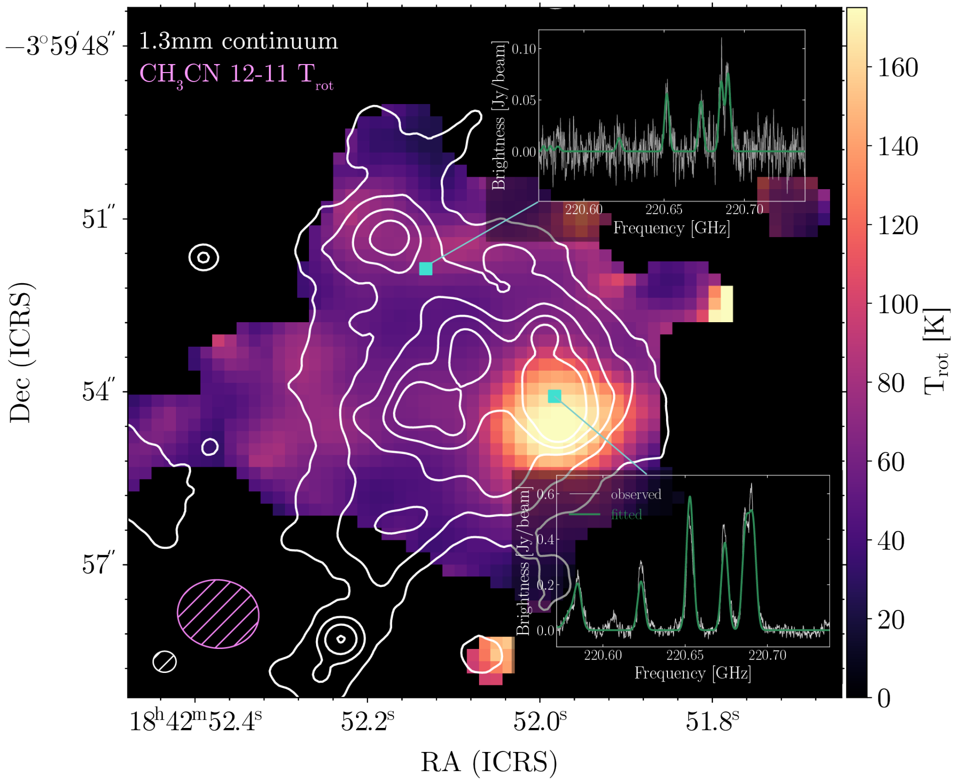
Gives line temperature estimate as function of radius

Possible to implement in realistic runtime

Low resolution required for signal-to-noise per pixel

Still relies on single LoS component fits (not true 3D for profiles)

Unusual to have very widespread emission in multiple CH_3CN transitions



CORE: Gieser et al. (2021) - line radial profiles
+ dust maps - power law index -2

Part II: Beyond core boundaries: structure recovery of high-mass protostellar core envelopes

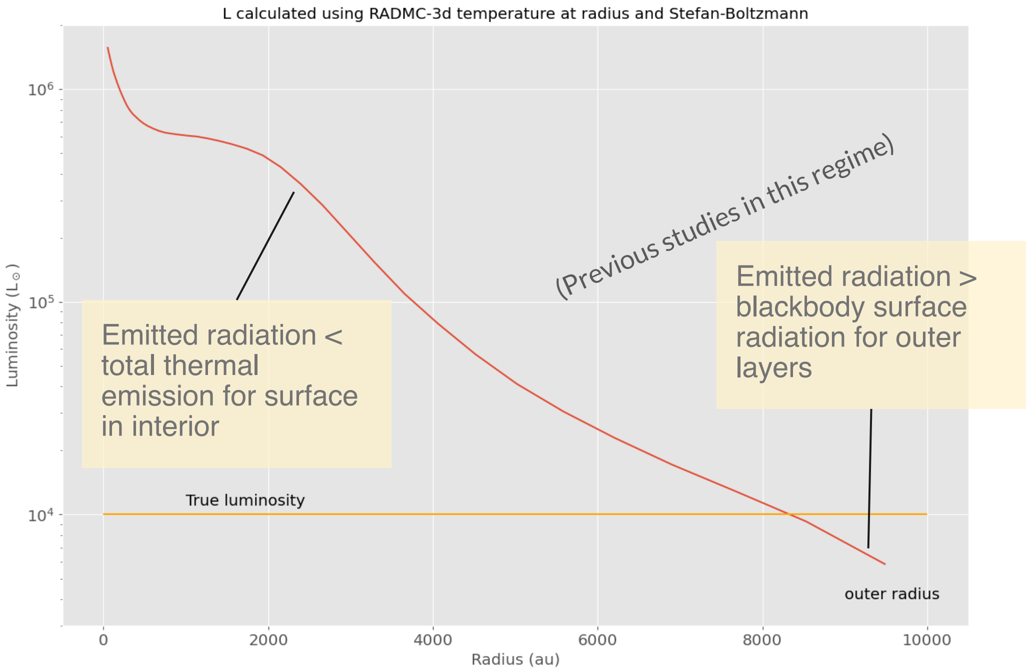
Solving the luminosity problem: transferring line temperatures between scales within cores

$$L = 4\pi r^2 \sigma_{SB} T^4$$

In core interiors, densities are sufficiently high that total emitted radiation is no longer proportional to total radiation in each interior surface for dense interiors (c.f. the Sun)

Ongoing JSC project: 3D radiative transfer modelling

(P. Schilke/B. Jones with S. Kesselheim/F. Niesel @ FZJ)



P. Schilke - **3D self consistent models with RADMC-3D**

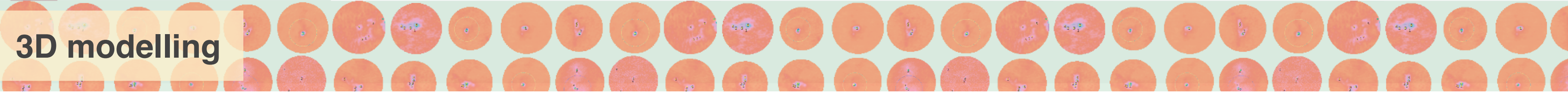
Precision astrophysics: source structure recovery

Assumptions of uniformity break down at core scales and densities (e.g. optical depth, strong gradients, no co-eval cores in a clump)

Current methods for high-precision **modelling is prohibitively computationally expensive** for core densities/gradients

Temperature uncertainty limits accuracy of property distributions (e.g. CMF), and is also unquantified

Observational **boundary definition inhibits comparison simulations**



Precision astrophysics: source structure recovery

Assumptions of uniformity break down at core scales and densities (e.g. optical depth, strong gradients, no co-eval cores in a clump)

Model each core with continuous functions - **XCLASS CubeFit**

Temperature uncertainty limits accuracy of property distributions (e.g. CMF), and is also unquantified

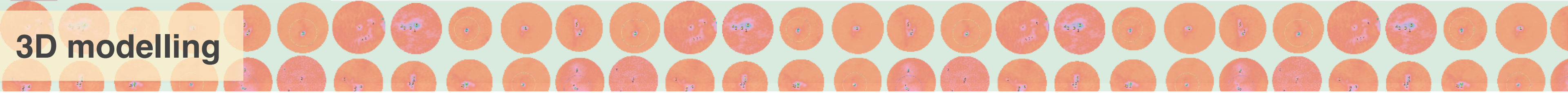
Calibration + uncertainty between discrete line and mass-averaged temperatures - **radiative transfer models, MCMC error estimation**

Current methods for high-precision **modelling is prohibitively computationally expensive** for core densities/gradients

Amortised **machine learning** methods

Observational **boundary definition inhibits comparison simulations**

Define continuous function in observations and simulations; **parameterise 3D structure + line of sight variation**



Precision astrophysics: source structure recovery

Assumptions of uniformity break down at core scales and densities (e.g. optical depth, strong gradients, no co-eval cores in a clump)

Current methods for high-precision **modelling is prohibitively computationally expensive** for core densities/gradients

Model each core with continuous functions - **XCLASS**

Amortised **machine learning** methods

These limitations can all be addressed with the currently available data by developing techniques to all the information available in both the spectral and spatial domain simultaneously

Temperature uncertainty limits accuracy of property distributions (e.g. CMF), and is also unquantified

Observational **boundary definition inhibits comparison simulations**

Calibration + uncertainty between discrete line and mass-averaged temperatures - **radiative transfer models, MCMC error estimation**

Define continuous function in observations and simulations; **parameterise 3D structure + line of sight variation**

Prototype ML: fitting source structure in 3D

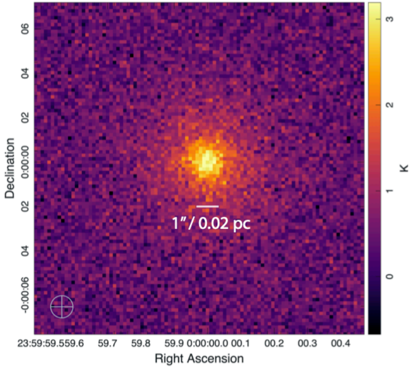
Machine learning project to test model inversion from information encoded in spectral and spatial dimensions available in observations

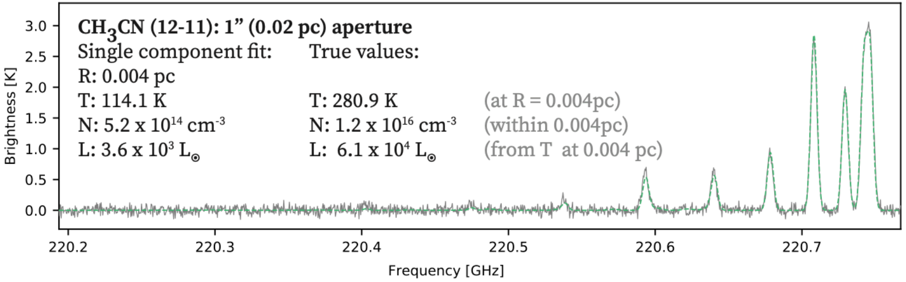
Power-law density and temperature profiles (central plateau)

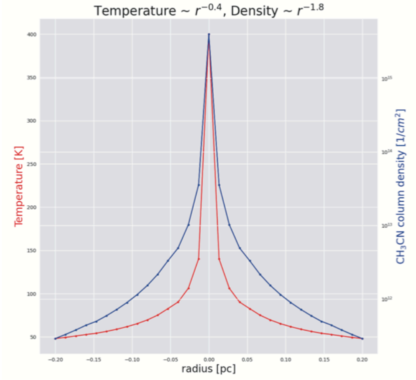
Spherical symmetry

Line radiative transfer with XCLASS

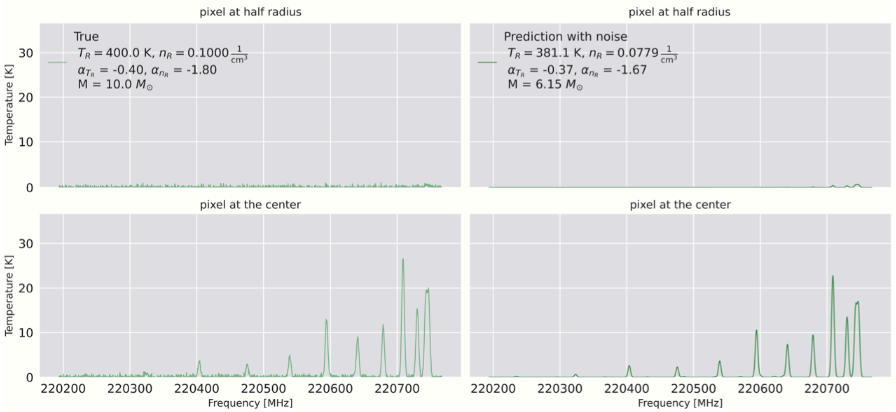
Sampling in three radially sampled spectra







CH_3CN (12-11): NN results from pixels at (0.001, 0.1, 0.2 pc)



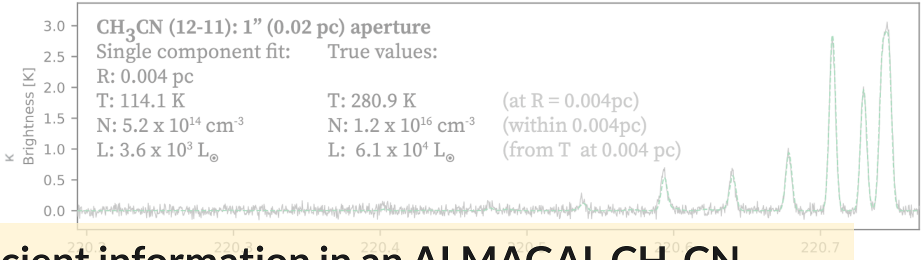
Lead: Judith Beck - Master's project

T. Möller - accelerated XCLASS generator

S. Kesselheim, **F. Niesel** (FZJ)

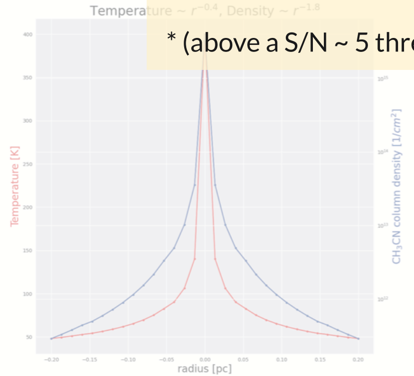
P. Schilke - self consistent models with RADMC-3D

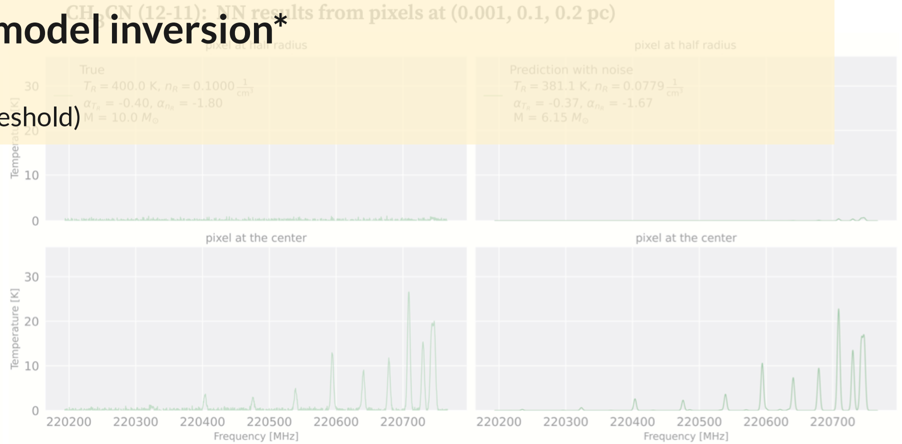
Prototype ML: fitting source structure in 3D



There is sufficient information in an ALMAGAL CH$_3$CN spectrum for model inversion*

* (above a S/N ~ 5 threshold)





Machine learning project to test model inversion from information encoded in spectral and spatial dimensions available in observations

Power-law density and temperature profiles (central plateau)

Spherical symmetry

Line radiative transfer with XCLASS

Sampling in three radially sampled spectra

Lead: Judith Beck - Master's project

T. Möller - accelerated XCLASS generator
S. Kesselheim, **F. Niesel** (FZJ)
P. Schilke - self consistent models with RADMC-3D

Self-consistent radiative transfer models

Training set of full physical models suitable for science application

Plummer-like density profiles and a central irradiating source with RADMC-3D

10,000 au outer radius, variable plateau radius

Density fluctuations and large-scale density asymmetries, abundance variation

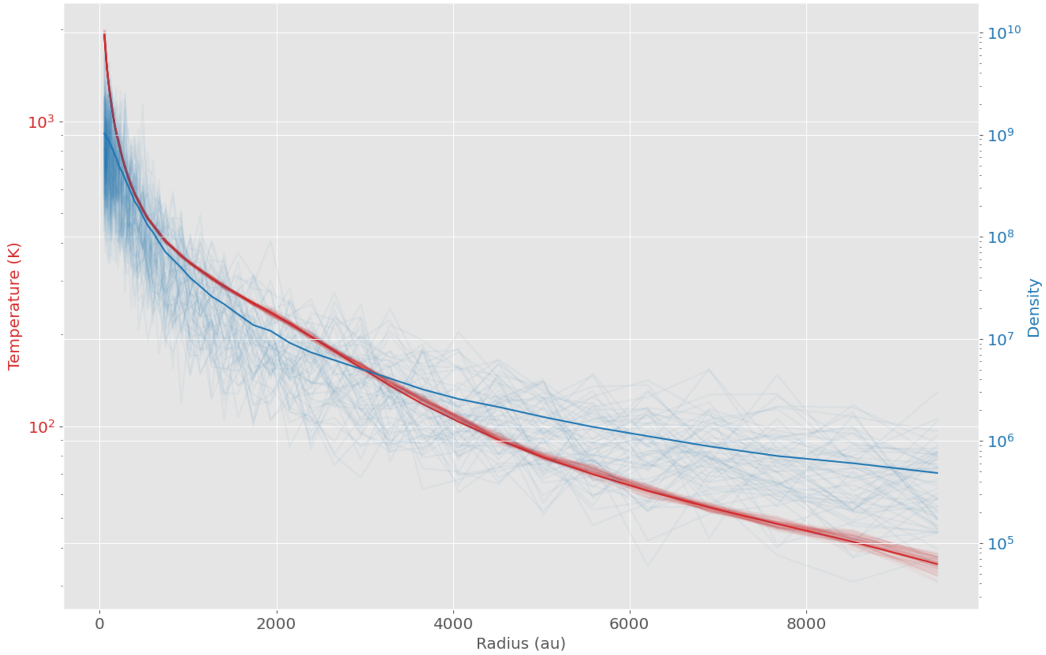
Accompanying development:

Upgrade neural network for generalisation to observational data *(MLP to CNN)*

Comparison to XCLASS CubeFit incl. uncertainty estimation on parameters

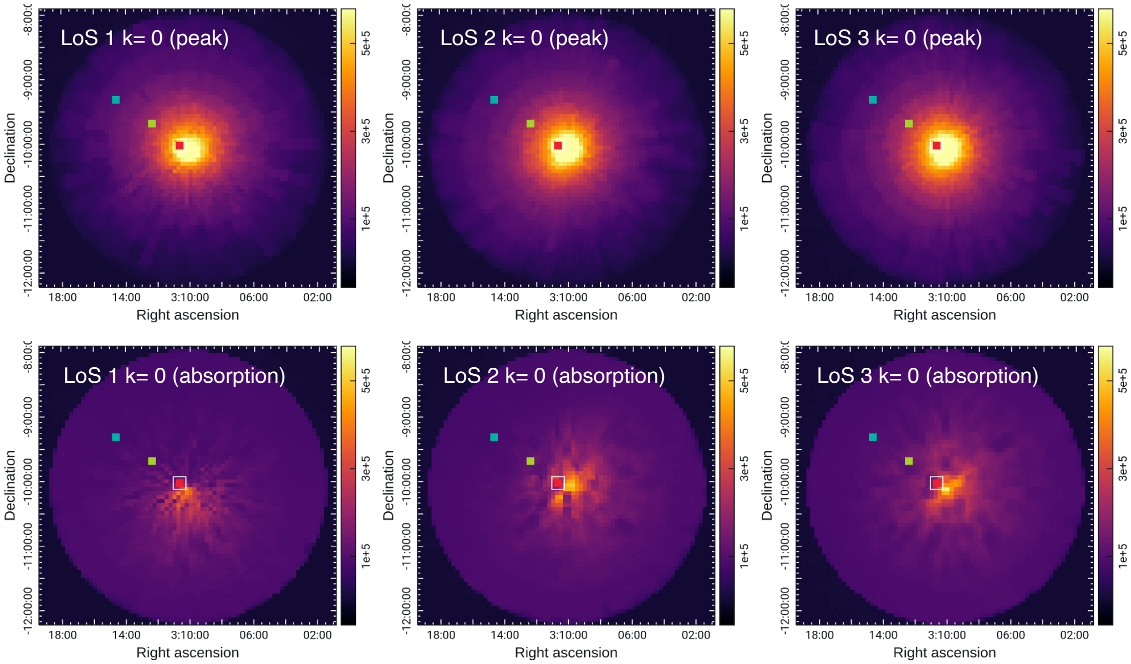
SBI (Simulation Based Inference) to recover posterior distributions instead of point estimates

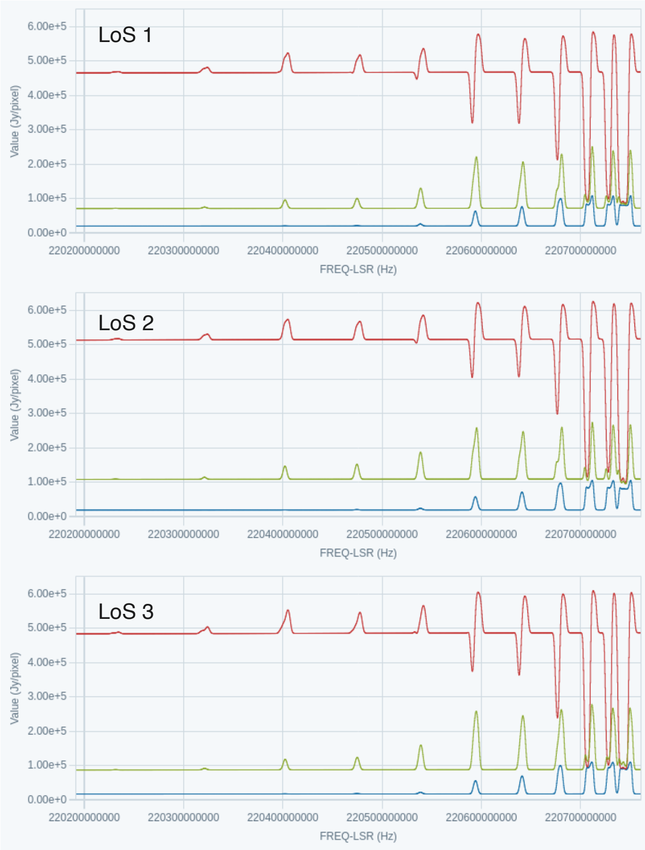
CubeFit results for calibration against conventional methods



J. Beck - Prototype ML; T. Möller - accelerated XCLASS generator; S. Kesselheim, F. Niesel (FZJ); **P. Schilke - self consistent models with RADMC-3D**

Self-consistent radiative transfer models





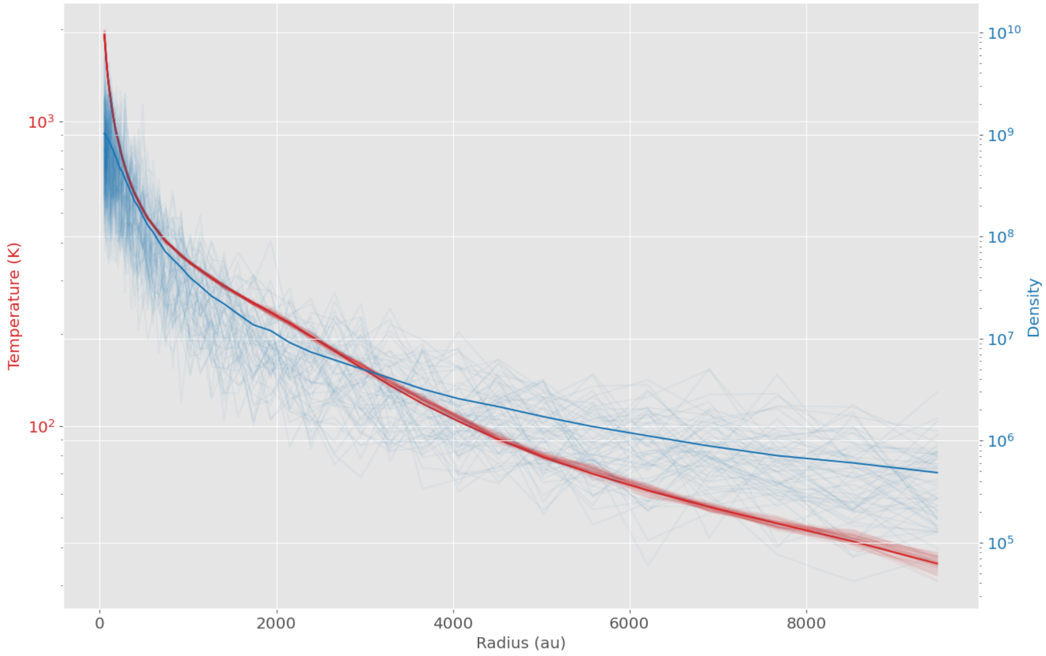
Self-consistent radiative transfer models

First application: spectral line temperature catalogue

Spectral line temperature catalogue: Models to calibrate between discrete component temperatures and true distribution + uncertainty

Second application: ALMAGAL source structure recovery

Train network and apply to ALMAGAL data (with source selection criteria for suitability)

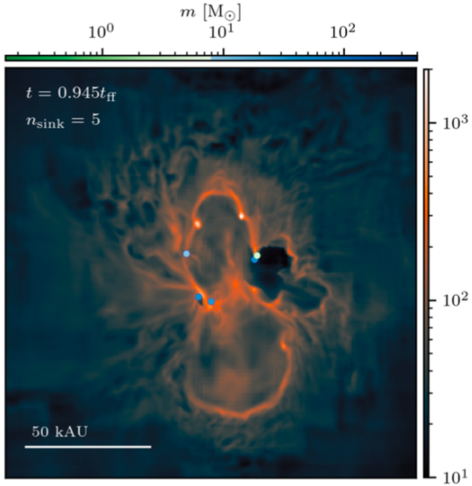


J. Beck - Prototype ML; T. Möller - accelerated XCLASS generator; S. Kesselheim, F. Niesel (FZJ); **P. Schilke - self consistent models with RADMC-3D**

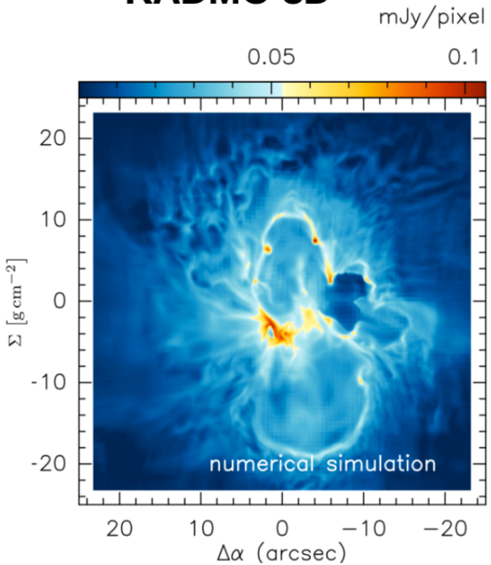
3D modelling: comparison to simulations

Core definition to match the observations…

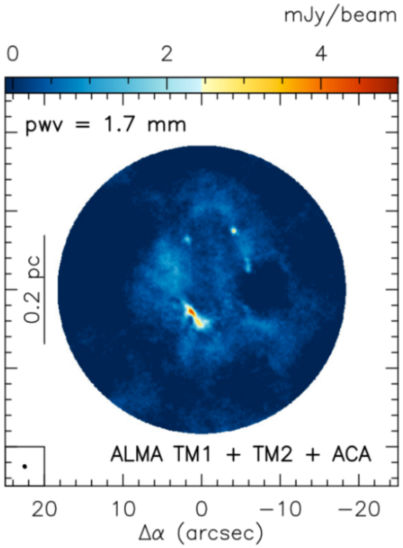
FLASH



RADMC-3D



CASA



Two step process for comparison to observations:

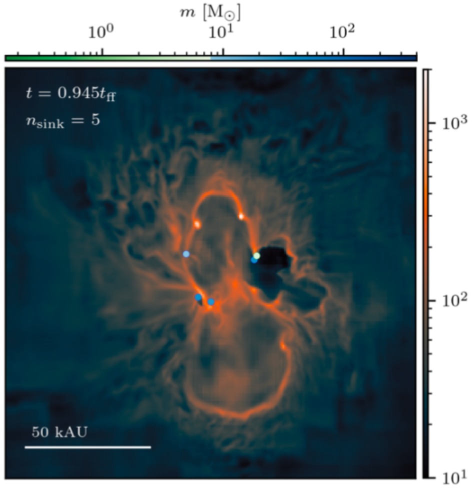
1. Process into dust emission maps
2. Process into observations with matched sensitivity, interferometric filtering and imaging artifacts

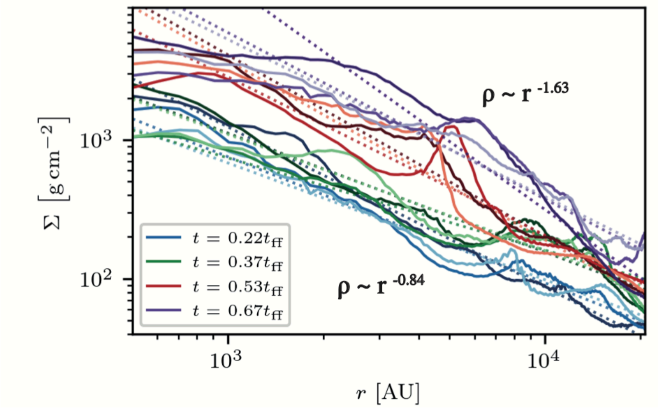
Leads: Birka Zimmermann, Flavia Santos do Amaral
Synthetic observation scripts: Alvaro Sanchez-Monge

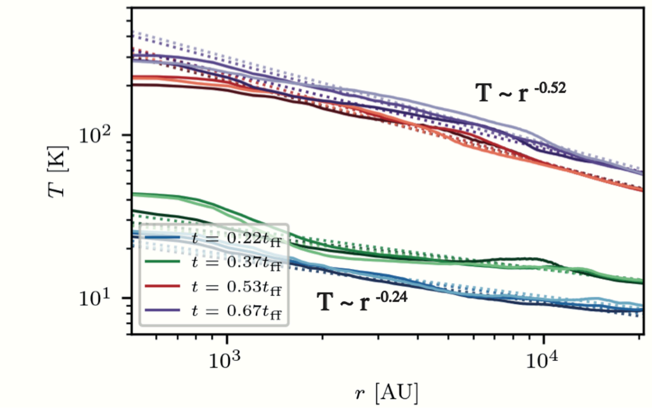
3D modelling: comparison to simulations

… and profile definition to match the simulations

Shown: fiducial run. Full set in Zimmermann et al. (subm.)



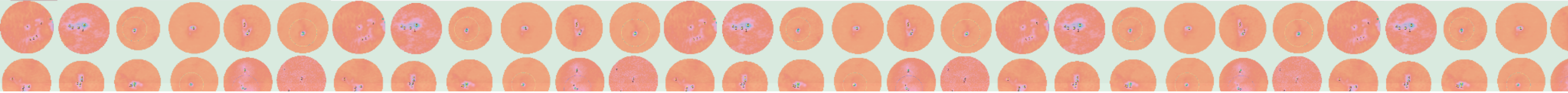




Colours: timesteps; profiles as **function of evolution**
Dotted lines: variation from observed **line of sight**

Steepening with evolution in radial density and temperature profiles as function of model parameters

Radial profiles (and radially averaged profiles) of surface density, temperature from simulated cores -> **radial profiles from observations**



Summary

- Large sample line fitting to over 1000 cores gives core temperatures between 60 and 600K for the ALMAGAL sample detected in CH_3CN
    - Calibration to transfer to continuum + uncertainty measurement (MCMC) ongoing

- Radiative transfer modelling is required to calibrate between continuum and line temperatures for reliable mass and luminosity estimates on 1000au core scales

- **We can do better, and ALMAGAL allows us to do transformational science in high-mass star formation**
    - 3D structure recovery **is** possible for large samples through ML for model inversion
    - There **is** sufficient information in the observations
    - Statistics to untangle for **probability distributions** for parameters
    - We **can** move away from discrete objects and define a 3D parameterisation for comparison to simulations

Precision astrophysics by exploiting ALL the information in the data (not just skimming the top 5%)