Interpreting molecular spectra

with examples

Andrea Giannetti

26 February 2019

INAF - IRA, Italian ALMA Regional Centre

Why molecular lines

Probes of physical conditions: T, n(H_2), ...

Why molecular lines

Probes of physical conditions: T, n(H_2), ...

Selectively trace gas with specific properties: cold/hot, shocked, dense/diffuse, ...

Why molecular lines

Probes of physical conditions: T, n(H_2), ...

Selectively trace gas with specific properties: cold/hot, shocked, dense/diffuse, ...

Star formation: protostellar activity, shocks, radiation field, ... (see Silvia's talk)

Why molecular lines

Probes of physical conditions: T, n(H_2), ...

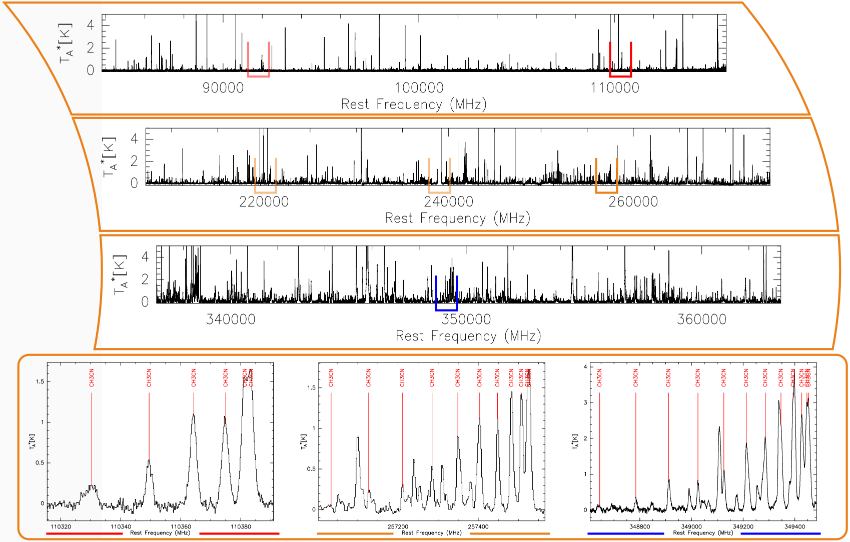
Selectively trace gas with specific properties: cold/hot, shocked, dense/diffuse, ...

Star formation: protostellar activity, shocks, radiation field, ... (see Silvia's talk)

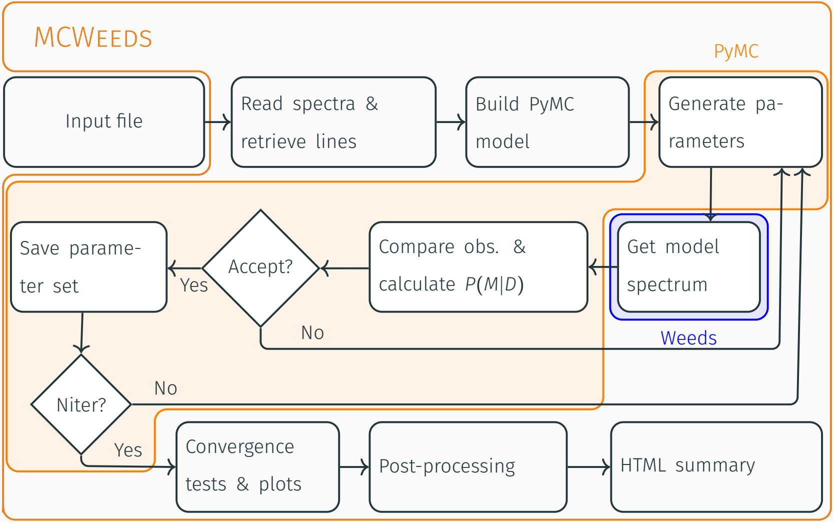
Chemistry of ISM: Formation of prebiotic molecules, chemical ages, evolution of physical properties, ...

New generation of telescopes and instruments

Unveiling excitation conditions and chemistry



Excitation conditions and column densities

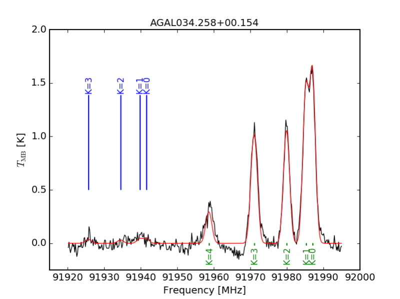


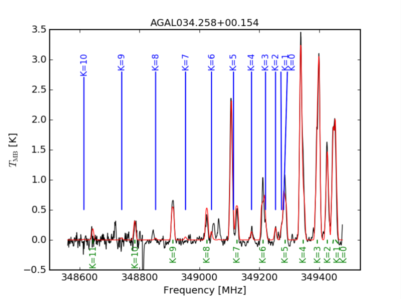
Evolution of high-mass star-forming regions

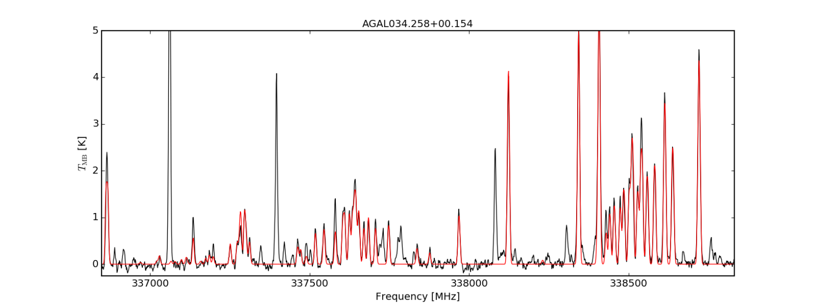
Evolutionary sequence for high-mass clumps

Tracers	Transitions
CH_3CCH: • Symmetric-top • Good T tracer	$(5-4)$: 1000 sources, E_u : $12-128$K $(6-5)$: 400 sources, E_u : $17-197$K $(20-19)$: 100 sources, E_u : $172-1200+$K
CH_3CN: • Symmetric-top • Good T tracer • Hot-core tracer	$(5-4)$: 1000 sources, E_u : $13-128$K $(6-5)$: 400 sources, E_u : $19-197$K $(19-18)$: 100 sources, E_u : $167-2500$K
CH_3OH: • Slightly asymmetric-top • T tracer • Hot-core tracer	$(7-6), \nu_t = 0$: 100 sources, E_u : $65-260$K $(7-6), \nu_t = 1$: 100 sources, E_u : $350-650$K

Examples of the fits







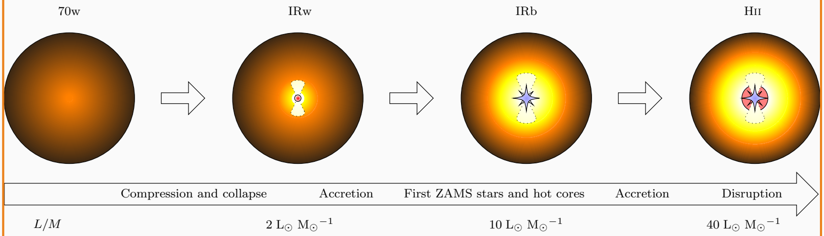
Evolutionary sequence for high-mass clumps

Observation of progressive warm-up due to YSOs in multiple tracers

Evolutionary sequence for high-mass clumps

Observation of progressive warm-up due to YSOs in multiple tracers

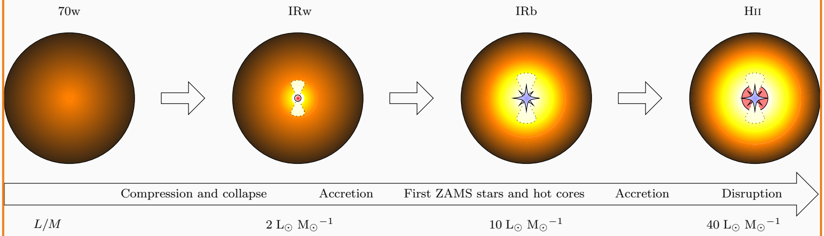
Validation of evolutionary sequence



Giannetti, Leurini+2017, A&A, 603, 33

Observation of progressive warm-up due to YSOs in multiple tracers

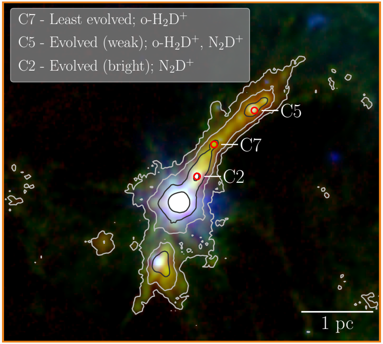
Validation of evolutionary sequence



Giannetti, Leurini+2017, A&A, 603, 33

Identification of most important process in intervals of L/M

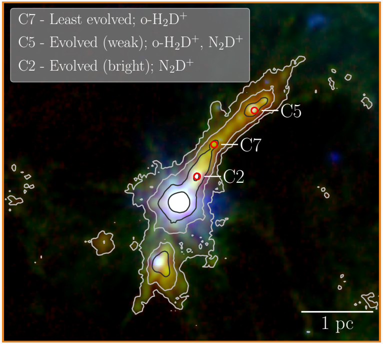
Timing the high-mass star formation process



Different evolutionary phases:

$L/M < 1 - 10\ L_\odot/M_\odot$

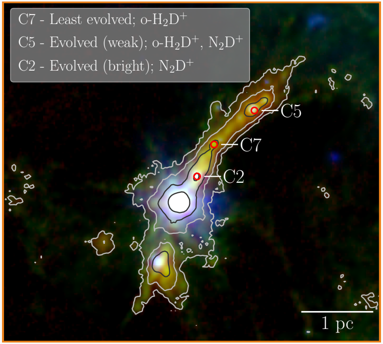
Timing the high-mass star formation process



Different evolutionary phases:
$L/M < 1 - 10\ L_\odot/M_\odot$

Three clumps observed with APEX in o-H_2D^+ and N_2D^+(3–2)

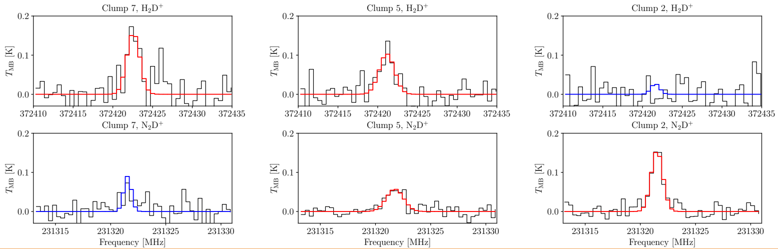
Timing the high-mass star formation process



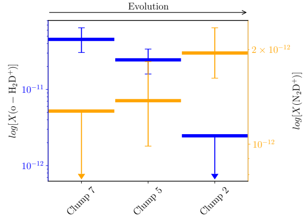
Different evolutionary phases: $L/M < 1-10\ L_\odot/M_\odot$

Three clumps observed with APEX in o-H$_2$D$^+$ and N$_2$D$^+$(3–2)

Opposite behaviour with evolution



Giannetti+2019, A&A,621, L7



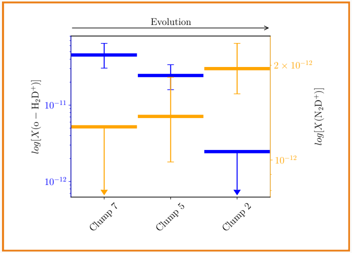
H_2D^+ is formed quickly in cold and dense gas where CO is depleted

Deuterium enrichment:

~~$H_3^+ + CO \rightarrow HCO^+ + H_2$~~

$H_3^+ + HD \leftrightarrow H_2D^+ + H_2 + \Delta E$

Why an anticorrelation?



H_2D^+ is formed quickly in cold and dense gas where CO is depleted

Time lag – $H_2D^+ \rightarrow N_2D^+$

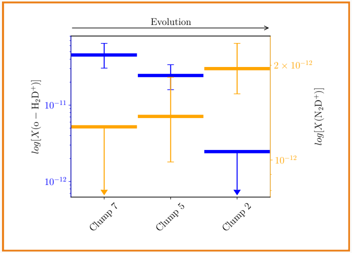
$$H_2D^+ + N_2 \rightarrow N_2D^+ + H_2$$

Deuterium enrichment:

~~$H_3^+ + CO \rightarrow HCO^+ + H_2$~~

$H_3^+ + HD \leftrightarrow H_2D^+ + H_2 + \Delta E$

Why an anticorrelation?



H_2D^+ is formed quickly in cold and dense gas where CO is depleted

Time lag – $H_2D^+ \rightarrow N_2D^+$

$$H_2D^+ + N_2 \rightarrow N_2D^+ + H_2$$

H_2D^+ being converted to multiply-deuterated forms

Deuterium enrichment:

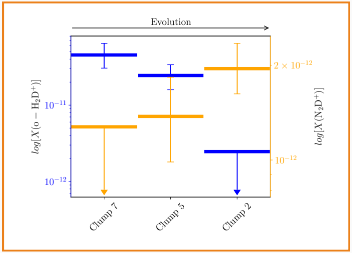
~~$H_3^+ + CO \rightarrow HCO^+ + H_2$~~

$H_3^+ + HD \leftrightarrow H_2D^+ + H_2 + \Delta E$

$H_2D^+ + HD \leftrightarrow D_2H^+ + H_2 + \Delta E$

$D_2H^+ + HD \leftrightarrow D_3^+ + H_2 + \Delta E$

Why an anticorrelation?



H_2D^+ is formed quickly in cold and dense gas where CO is depleted

Time lag – $H_2D^+ \rightarrow N_2D^+$

$$H_2D^+ + N_2 \rightarrow N_2D^+ + H_2$$

H_2D^+ being converted to multiply-deuterated forms

Increased efficiency for N_2D^+ formation

Deuterium enrichment:

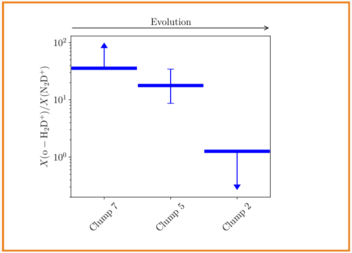
~~$H_3^+ + CO \rightarrow HCO^+ + H_2$~~

$$H_3^+ + HD \leftrightarrow H_2D^+ + H_2 + \Delta E$$

$$H_2D^+ + HD \leftrightarrow D_2H^+ + H_2 + \Delta E$$

$$D_2H^+ + HD \leftrightarrow D_3^+ + H_2 + \Delta E$$

Why an anticorrelation?



Deuterium enrichment:

~~$H_3^+ + CO \rightarrow HCO^+ + H_2$~~

$H_3^+ + HD \leftrightarrow H_2D^+ + H_2 + \Delta E$

$H_2D^+ + HD \leftrightarrow D_2H^+ + H_2 + \Delta E$

$D_2H^+ + HD \leftrightarrow D_3^+ + H_2 + \Delta E$

H_2D^+ is formed quickly in cold and dense gas where CO is depleted

Time lag – $H_2D^+ \rightarrow N_2D^+$

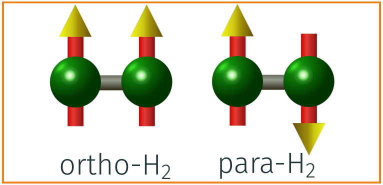
$$H_2D^+ + N_2 \rightarrow N_2D^+ + H_2$$

H_2D^+ being converted to multiply-deuterated forms

Increased efficiency for N_2D^+ formation

From chemical models: age $\lesssim 10^5$ yr for youngest clump

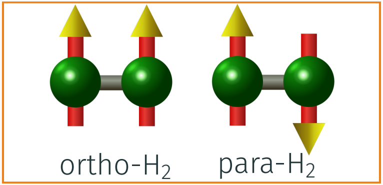
H_2 ORTHO-TO-PARA RATIO

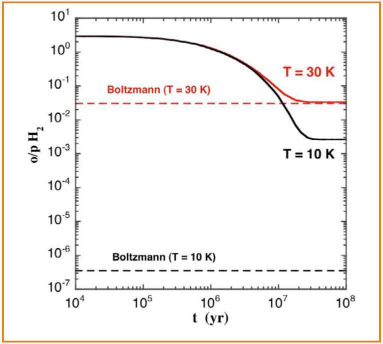


Taken from Flower+2006

Two spin isomers of H_2 exist

H_2 ORTHO-TO-PARA RATIO



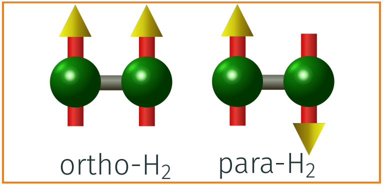


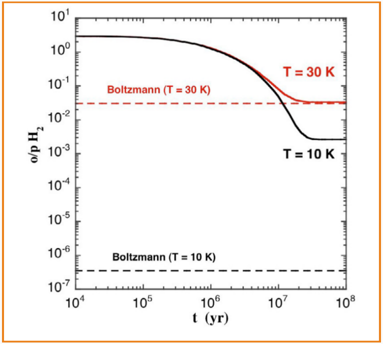
Taken from Flower+2006

Two spin isomers of H_2 exist

OPR steadily decreases with time

H_2 ORTHO-TO-PARA RATIO





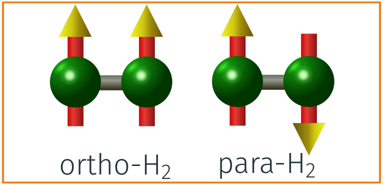
Taken from Flower+2006

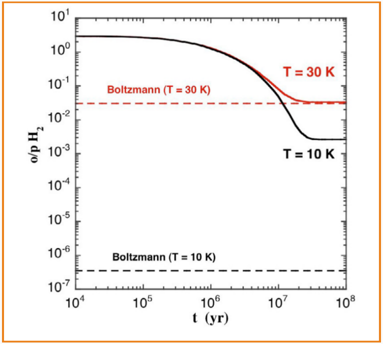
Two spin isomers of H_2 exist

OPR steadily decreases with time

Most efficient age estimator available

H_2 ORTHO-TO-PARA RATIO





Taken from Flower+2006

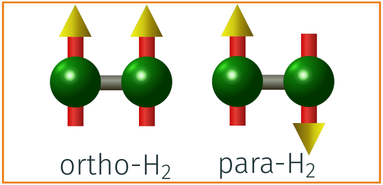
Two spin isomers of H_2 exist

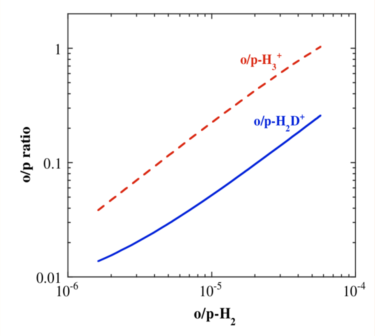
OPR steadily decreases with time

Most efficient age estimator available

Cannot be measured directly

H_2 ORTHO-TO-PARA RATIO





Taken from Walmsley+2004

Two spin isomers of H_2 exist

OPR steadily decreases with time

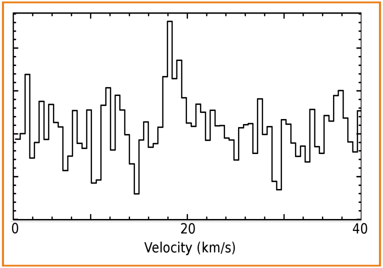
Most efficient age estimator available

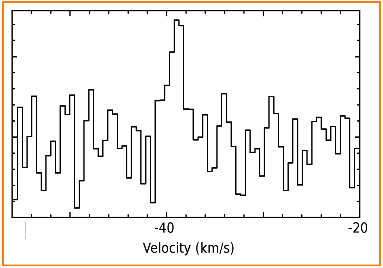
Cannot be measured directly

Connected to the ortho-para ratio of H_2D^+

Survey of o-H_2D^+ in TOP100: 17 detections!

The age of a massive clump with SOFIA

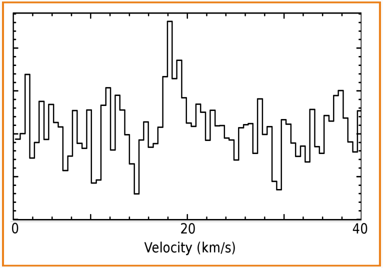


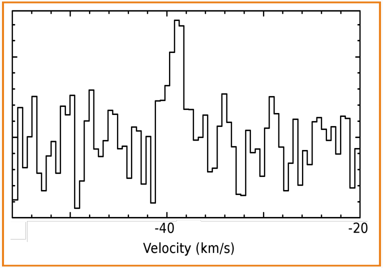


Survey of o-H_2D^+ in TOP100: 17 detections!

Selected strongest clumps at 1.37 THz ($\sim$ 300 Jy)

The age of a massive clump with SOFIA



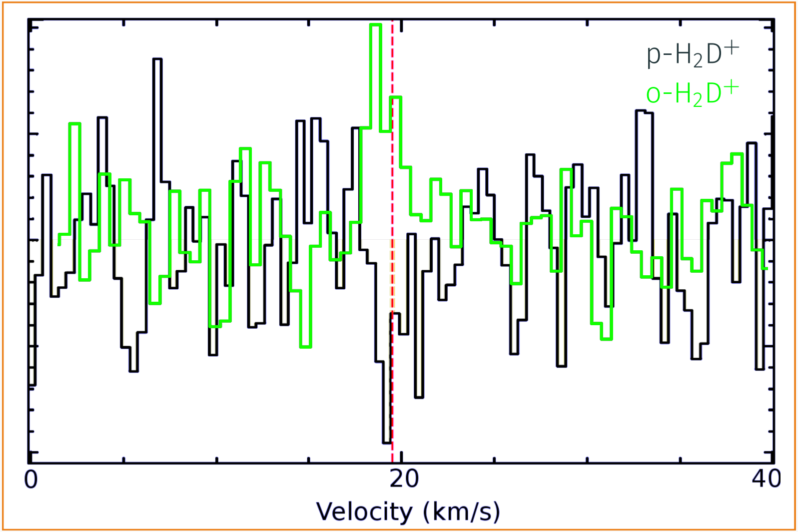


Survey of o-H_2D^+ in TOP100: 17 detections!

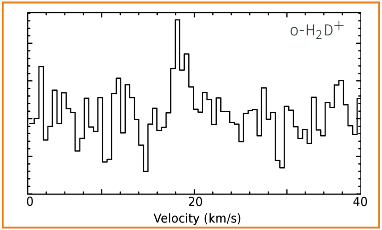
Selected strongest clumps at 1.37 THz ($\sim$ 300 Jy)

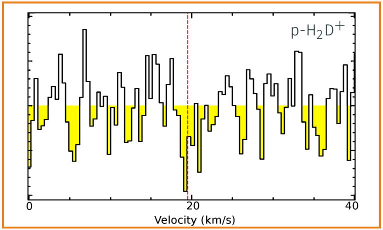
First tentative detection ($\sim 2\sigma$) of p-H_2D^+ in high-mass clump

The age of a massive clump with SOFIA



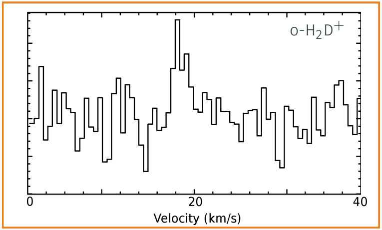
The age of a massive clump with SOFIA



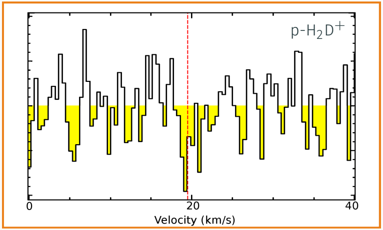


In Cycle 7 obtained 6 hrs to secure the detection

THE AGE OF A MASSIVE CLUMP WITH SOFIA



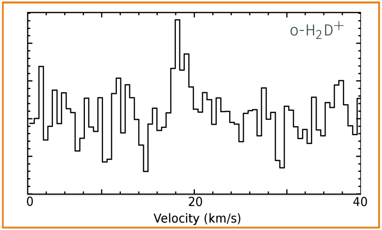
+



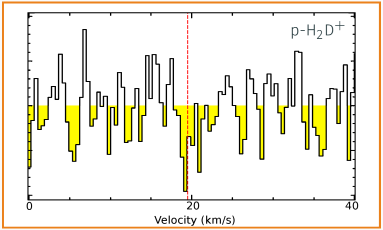
In Cycle 7 obtained 6 hrs to secure the detection

First direct measurement of ortho-para H_2D^+ in high-mass regime, second in literature

The age of a massive clump with SOFIA



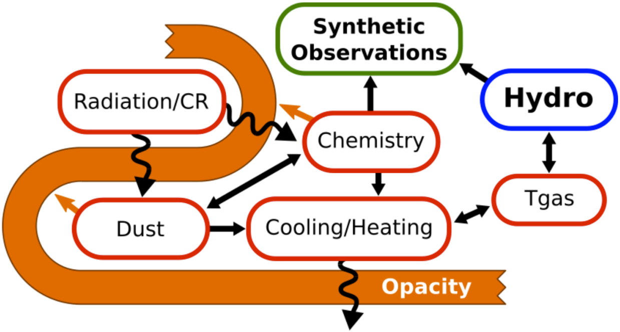
+



In Cycle 7 obtained 6 hrs to secure the detection

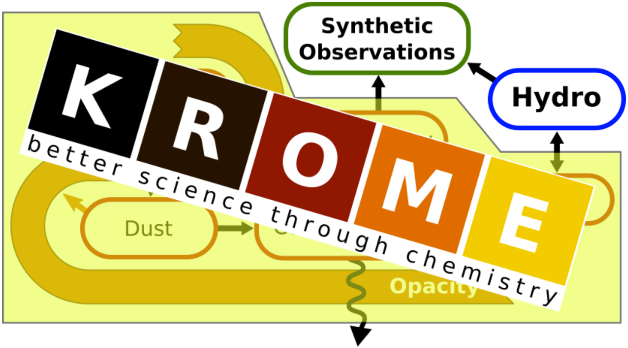
First direct measurement of ortho-para H_2D^+ in high-mass regime, second in literature

One more ingredient needed to estimate clump ages



- Extremely CPU demanding (solving stiff ODEs)
- Many complex and interconnected physical processes
- Needs atomic/molecular and thermochemical data, reaction rate coefficients,...

KROME



- Python pre-processor creates ad hoc optimized F90 modules
- Chemistry, dust-related physics, cooling, heating, photoionization,...
- Open source (Grassi+2014)
- Highly optimized code, based on "fast" solver (DLSODES)
- Hydrocodes-ready, RAMSES, ENZO, FLASH, GASOLINE, GIZMO,...
- $>$ 30 papers (CEMP-stars, molecular clouds, galaxies, BHs, AGBs, pp-disks, . . .)

Simulations of high-mass clumps

3D MHD simulations of collapsing high-mass clumps

Simulations of high-mass clumps

3D MHD simulations of collapsing high-mass clumps

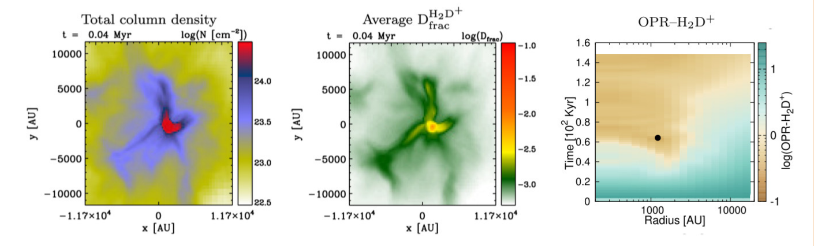
Coupled with chemistry under full depletion hypothesis

Simulations of high-mass clumps

3D MHD simulations of collapsing high-mass clumps

Coupled with chemistry under full depletion hypothesis

Describe evolution of H_3^+ (& deuterated isotopologues)



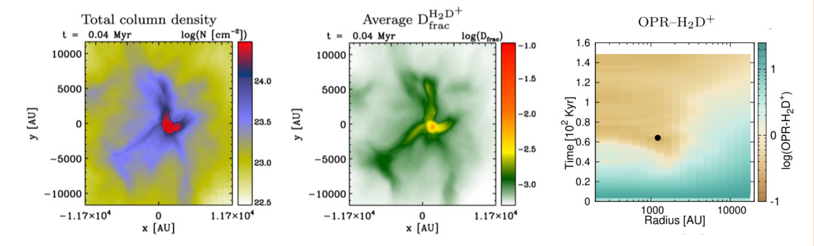
Taken from Körtgen+2017

Simulations of high-mass clumps

3D MHD simulations of collapsing high-mass clumps

Coupled with chemistry under full depletion hypothesis

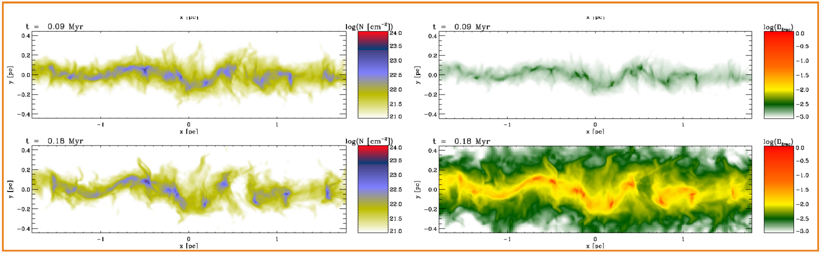
Describe evolution of H_3^+ (& deuterated isotopologues)



Taken from Körtgen+2017

Built postprocessing pipeline to obtain synthetic observations (Zamponi et al., in prep.)

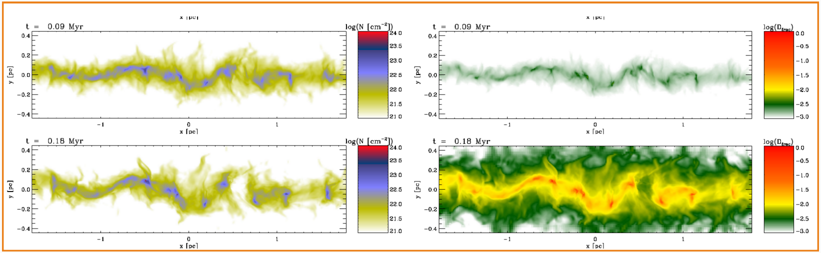
Clumps simulated in their environment



Taken from Körtgen+2018

Simulations of high-mass clumps

Clumps simulated in their environment

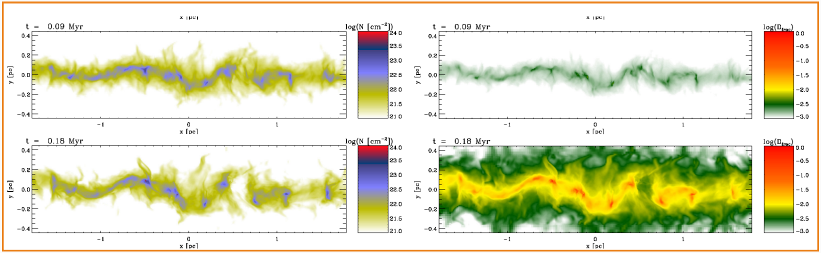


Taken from Körtgen+2018

Improved chemical network to follow time evolution of o-H_2D^+ / N_2D^+

Simulations of high-mass clumps

Clumps simulated in their environment



Taken from Körtgen+2018

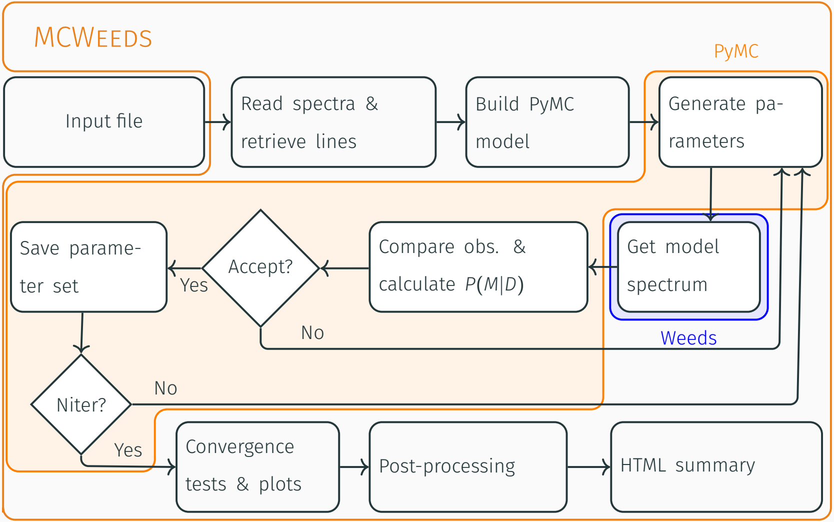
Improved chemical network to follow time evolution of o-H_2D^+ / N_2D^+

Simulations are running

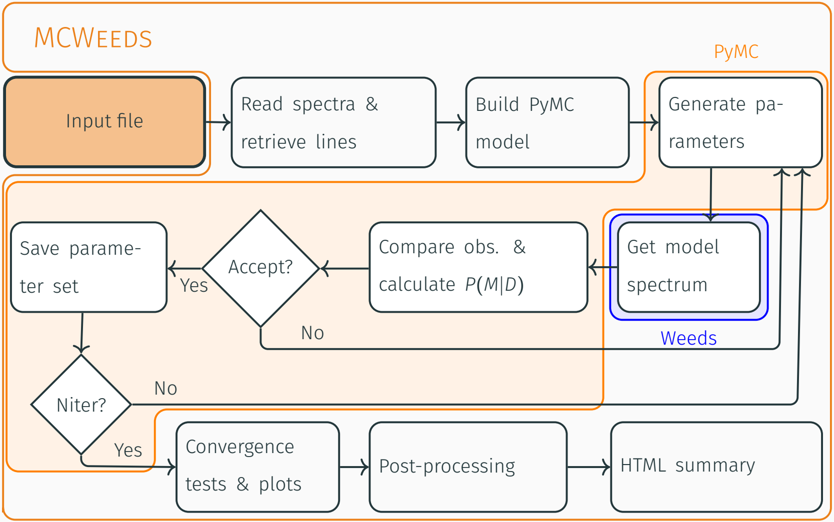
MCWeeds

in (more) detail

MCWeeds flowchart



MCWeeds flowchart



Input file

```
# List of the species that you want to fit with their relative in-
tensity (e.g. useful for isotopologues)
  #                              name            rel. intensity
  species                        CH3CN, v=0      1
  # Frequency ranges,            input_file,     antenna[m]
  frange 91920.0 91995.0         infile1         30
  frange 110340. 110400.         infile2         30

# Database
  db_input     cdms
  db_output    line_database.sql
  db_cached    False
  db_update    True

  source       G35.20-0.74
```

Input file description

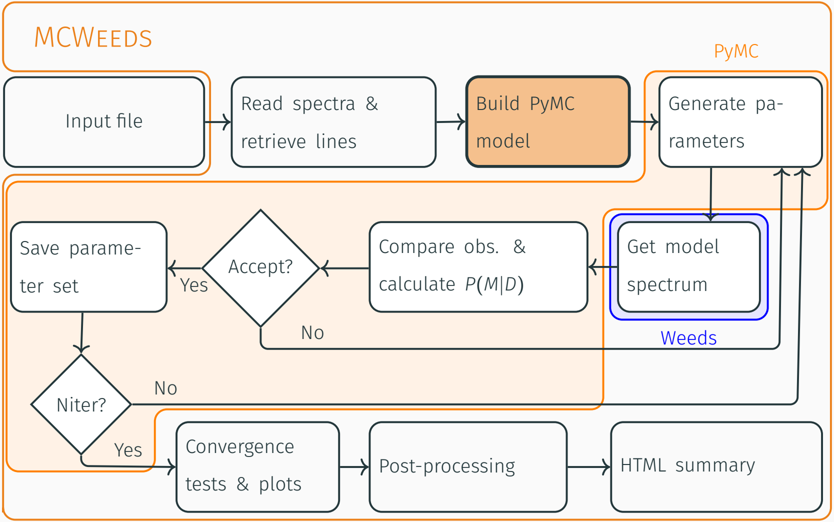
Distributions for the stochastics representing column density, temperature, size, velocity, linewidth

#	stochastic	det_var	init_guess	priors	parameters
spec_par	log10_cd	log10_N	13.	Normal	mu=0. tau=1./2.**2
spec_par	tex	T	20.	TruncatedNormal	mu=0. tau=1./25.**2 a=-12 b=75.
spec_par	size	theta	100.	Fixed	value=100.
spec_par	velocity	V	0.	Normal	mu=0. tau=1./2.**2
spec_par	linewidth	DV	4.	TruncatedNormal	mu=0. tau=1./3.**2 a=-3.5 b=30.

Calibration factors and rms noises

#	stochastic	priors	parameters
model_par	calib_fact_0	TruncatedNormal	mu=0. tau=1./0.07**2 a=-0.3 b=0.3
model_par	rms_spec_0	Fixed	value=0.04
model_par	calib_fact_0	TruncatedNormal	mu=0. tau=1./0.07**2 a=-0.3 b=0.3
model_par	rms_spec_0	Fixed	value=0.04

MCWeeds flowchart



Model file

```
from pymc import Normal, deterministic, Uniform, TruncatedNormal, InverseGamma, Lambda
import numpy as np
import pyclass
import modules.pmodsource_short as pmodsource
import modules.dbutils as dbutils

# ... Definition of variables ...

def get_fit(par_array, antenna, ydata, lines_input, partition_function,dbdata=None):
    (lines, temperature, partfunc) = dbdata
    # get the function values of the fit
    yfit = np.array(pmodsource.main(par_array, antenna, lines,partition_function, ydata, lines_input))
    return yfit

# Set up Priors
linewidth = TruncatedNormal('linewidth',mu=0.,tau=1./3.**2,b=30.,a=-3.5)
rms_spec_1 = InverseGamma('rms_spec_1',alpha=20.,beta=1.)
rms_spec_0 = InverseGamma('rms_spec_0',alpha=20.,beta=1.)
excitation_temperature = TruncatedNormal('T',mu=0.,tau=1./10.**2,b=50.,a=-42)
log10_col_dens = Normal('log10_N',mu=0.,tau=1./5.**2)
calib_fact_0 = TruncatedNormal('calib_fact_0',mu=0.,tau=1./0.07**2,b=0.3,a=-0.3)
velocity = Normal('velocity',mu=0.,tau=1./1**2)
```

Model file

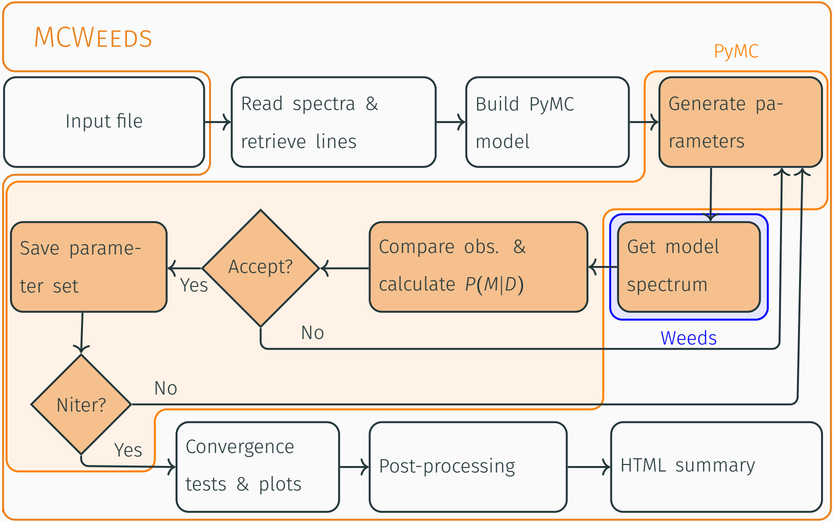
```
@deterministic(plot=True)
def size(temperature=excitation_temperature,dist=distance,lumi=luminosity):
    result = (2.*4.31*lumi**(5./8.)*(temperature+50.)**(-5./2.))/dist*(180./(np.pi)*3600.)
    return result

@deterministic(plot=False)
def get_model_spectrum_0(log10_N=log10_col_dens, T=excitation_temperature, theta=size, V=velocity, DV=linewidth, cf=calib_fact_0, sp_data=spec_data_0):
    par_array = [['CH3CN'],[10**(log10_N+13.)],[(T+50.)],[(theta)],[(V+0.)],[(DV+4.)],[None]]
    fit = get_fit(par_array, antenna_0, sp_data, lines_input_0,partition_function=partition_function['sp_slice_0'], dbdata=dbdata_0)
    model_spec_0 = (cf + 1.) * fit
    return model_spec_0

...

signal_mod_0 = Normal('signal_mod_0', mu=get_model_spectrum_0, tau=(1./rms_spec_0)**2, value=spec_array[0], observed=True)
signal_mod_1 = Normal('signal_mod_1', mu=get_model_spectrum_1, tau=(1./rms_spec_1)**2, value=spec_array[1], observed=True)
```

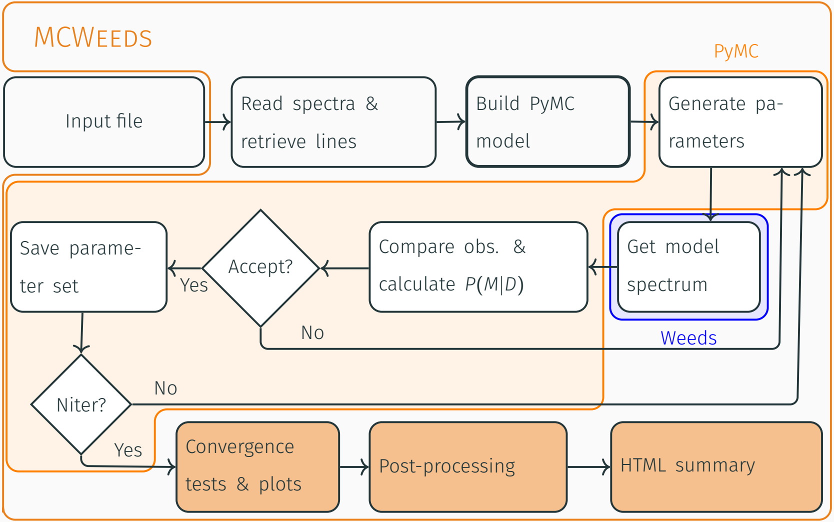
MCWeeds flowchart



Fitting algorithms

1. Maximum a posteriori estimate
    - Fast
    - No estimation of uncertainties
2. Normal Approximation: joint distribution of all stochastic variables is assumed Gaussian
    - Fast
    - Uncertainties are estimated with the above assumption
3. Monte Carlo Markov Chains
    - Slow
    - Full probability distribution function for stochastics

MCWeeds flowchart



Examples

```
andrea@Newton: ~/Documents/mcweeds_group/source
andrea@Newton: ~/Documents/mcweeds_gr...   andrea@Newton: ~/Documents/mcweeds_gro...
[-----------------92%---------------    ] 36985 of 40000 complete in 114.9 sec
[-----------------92%---------------    ] 37135 of 40000 complete in 115.4 sec
[-----------------93%---------------    ] 37284 of 40000 complete in 115.9 sec
[-----------------93%---------------    ] 37436 of 40000 complete in 116.4 sec
[-----------------93%---------------    ] 37587 of 40000 complete in 116.9 sec
[-----------------94%---------------    ] 37740 of 40000 complete in 117.4 sec
[-----------------94%---------------    ] 37891 of 40000 complete in 117.9 sec
[-----------------95%----------------   ] 38041 of 40000 complete in 118.5 sec
[-----------------95%----------------   ] 38190 of 40000 complete in 119.0 sec
[-----------------95%----------------   ] 38341 of 40000 complete in 119.5 sec
[-----------------96%----------------   ] 38492 of 40000 complete in 120.0 sec
[-----------------96%----------------   ] 38642 of 40000 complete in 120.5 sec
[-----------------96%----------------   ] 38794 of 40000 complete in 121.0 sec
[-----------------97%----------------   ] 38942 of 40000 complete in 121.5 sec
[-----------------97%-----------------  ] 39094 of 40000 complete in 122.0 sec
[-----------------98%-----------------  ] 39243 of 40000 complete in 122.5 sec
[-----------------98%-----------------  ] 39397 of 40000 complete in 123.0 sec
[-----------------98%-----------------  ] 39546 of 40000 complete in 123.5 sec
[-----------------99%-----------------  ] 39699 of 40000 complete in 124.0 sec
[-----------------99%-----------------  ] 39849 of 40000 complete in 124.5 sec
[-----------------100%-----------------] 40000 of 40000 complete in 125.0 sec
[-----------------100%-----------------] 40000 of 40000 complete in 125.0 sec
MCMC finished...
Writing summaries...
```

Examples

andrea@Newton: ~/Documents/mcweeds_group/source

andrea@Newton: ~/Documents/mcweeds_gr... | andrea@Newton: ~/Documents/mcweeds_gro...

```
Raftery-Lewis Diagnostic
========================

1825 iterations required (assuming independence) to achieve 0.01 accuracy with 9
5 percent probability.

Thinning factor of 1 required to produce a first-order Markov chain.

2 iterations to be discarded at the beginning of the simulation (burn-in).

1789 subsequent iterations required.

Thinning factor of 1 required to produce an independence chain.
Processing  excitation_temperature

========================
Raftery-Lewis Diagnostic
========================

1825 iterations required (assuming independence) to achieve 0.01 accuracy with 9
5 percent probability.

Thinning factor of 1 required to produce a first-order Markov chain.
```

Examples

```
andrea@Newton: ~/Documents/mcweeds_group/source
andrea@Newton: ~/Documents/mcweeds_gr...    andrea@Newton: ~/Documents/mcweeds_gro...
Thinning factor of 2 required to produce an independence chain.
Could not calculate Gelman-Rubin statistics. Requires multiple chains of equal l
ength.
++++++++++++++++++++++++++++++++++++++++++++++++++++++++
Best fit parameters chain 0 :
++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Species:  CH3CN
===============================
log10_col_dens 13.6442266391 , 13.5921833085 13.6995242542
excitation_temperature 53.8339360278 , 49.4726953484 57.7881884956
size 100.0 None
velocity 0.134295486607 , 0.04175963317 0.22752306975
linewidth 4.94223968611 , 4.721066895 5.1400558639

calib_fact_0 1.03507728941 , 0.929935921351 1.13359633074
calib_fact_1 0.964248074973 , 0.868511588271 1.05773210222
rms_spec_0 0.0431085045918 , 0.0403838829195 0.0464037501529
rms_spec_1 0.0540167672227 , 0.0507941299323 0.0574697529438

Getting intensities...
global   on-screen  inactive filter: FE-------
global   to-mesfile inactive filter: FEWRI--CU
andrea@Newton:~/Documents/mcweeds_group/source$
```

Output example: HTML page