The Foundations of Molecular Cloud Population Synthesis

Marissa Perry\textsuperscript{1,2}, Jens Kauffmann\textsuperscript{1}

\textsuperscript{1}MIT Haystack Observatory
\textsuperscript{2}The University of Texas at Austin
REU 2023
Motivation: can we meaningfully decompose this unresolved signal?
Astrochemistry

Spiral Arm in M51

- dense gas
- feedback
- shocks
- diffuse gas

Watanabe et al. 2013

Marissa Perry | MIT Haystack Observatory | UT Austin | Haystack REU Symposium
Astrochemistry

Emission line intensity

\[ T_{\text{MB}} = T_{\text{ex}} \cdot (1 - e^{-\tau}) \]

Molecular density
Molecular abundance

Emission lines probe molecular cloud properties

Spiral Arm in M51

Watanabe et al. 2013
Molecular Cloud Evolution

Chevance et al. 2020

-30 -20 -10 0 10

Time [Myr]

- Gas
- Stars

NGC5068
NGC628
NGC3351
NGC4535

intense star formation and cloud dispersal

early cloud formation

efficient SF

gas dispersion

hot core (e.g. CH₃OH)
dense cold gas (e.g. N₂H⁺)
HII region
PDR (e.g. C₂H)
Population Synthesis: Signal Decomposition

\[ \text{measured cloud} \times \text{weight} + \text{measured cloud} \times \text{weight} + \text{measured cloud} \times \text{weight} + \ldots \]

\[ \sim 10 \text{ pc} \]
Population Synthesis: Signal Decomposition

1) Numerical Optimization
   • SciPy Optimize Minimize

2) Probabilistic Programming (MCMC)
   • PyMC3
Signal Decomposition: Line Luminosity

$T_{\text{MB}} = T_{\text{ex}} \cdot (1 - e^{-\tau})$

$n$ # of molecular emission lines

$\mathbf{L}_i = \begin{pmatrix} L_{i,1} \\ \vdots \\ L_{i,n} \end{pmatrix}$
Signal Decomposition: Galaxy Region

$m$ # of cloud populations

\[
\vec{L}_{\text{gal}} = \vec{L}_a \cdot w_a + \ldots + \vec{L}_m \cdot w_m \\
= \hat{L}_{\text{cloud}} \cdot \hat{w}
\]

Marissa Perry | MIT Haystack Observatory | UT Austin | Haystack REU Symposium
The system is solvable when there are more emission lines ($n$) than molecular clouds ($m$).

$$\vec{L}_{\text{gal}} = \hat{L}_{\text{cloud}} \cdot \vec{w}$$

Synthetic observation → Synthetically generated

$$\vec{L}_{\text{model}} = \hat{L}_{\text{cloud}} \cdot \vec{w}$$

Synthetic observation → Unknown

$$F_L = (\vec{L}_{\text{gal}} - \vec{L}_{\text{model}})^2$$

$m = 5, n = 15$ → Median error: $37.95\%$
\[ \hat{L}_{\text{cloud,obs}} = \hat{L}_{\text{cloud}} + \delta \hat{L}_{\text{cloud}} \]

\[ \vec{L}_{\text{gal,obs}} = \vec{L}_{\text{gal}} + \delta \vec{L}_{\text{gal}} \]

- Synthetic observation
- True value
- Measurement error

\[ m = 5, \ n = 15 \]

\[ \rightarrow \text{Median error: } 5.96\% \]
• Population synthesis of molecular clouds might help us overcome the resolution gap

• In decomposing an unresolved galaxy signal, one must be observing more emission lines (n) than molecular clouds (m)

• Probabilistic programming methods, such as MCMC, are more desirable
Acknowledgements: thank you to Dianne, Nancy, Vincent and Phil for organizing this year’s REU program. Thanks to Jens for the mentorship and to his student Derek Sheen for the work which the numerical optimization experiment was built upon.
Extra Slides
Optimization Problems

Constraining cloud weights values to be non-negative
Testing Various MCMC Models

Ensuring MCMC calculations accounted for measurement errors in both cloud and galaxy luminosity

\[ \hat{L}_{\text{cloud,obs}} = \hat{L}_{\text{cloud}} + \delta \hat{L}_{\text{cloud}} \]

\[ \tilde{L}_{\text{gal,obs}} = \tilde{L}_{\text{gal}} + \delta \tilde{L}_{\text{gal}} \]