Studying Star Formation through Hierarchical Bayesian Modeling of Emission from Astronomical Dust

Brandon C. Kelly (CfA), Rahul Shetty (ITA, Heidelberg), Amelia M. Stutz (MPIA), Jens Kauffmann (JPL), Alyssa Goodman (CfA), Ralf Launhardt (MPIA)

Submitted to the Astrophysical Journal

Using Dust to Study Star Formation

Cold Clouds of gas and dust become unstable, collapse

Dust is important for cooling gas, aids collapse

↓

Eventually, some regions become dense and hot enough to start fusion

↓

Star is formed, disk of material forms around star

↓

Dust grains collide in disk around star, form building blocks of planets

Are the observations consistent with this picture?

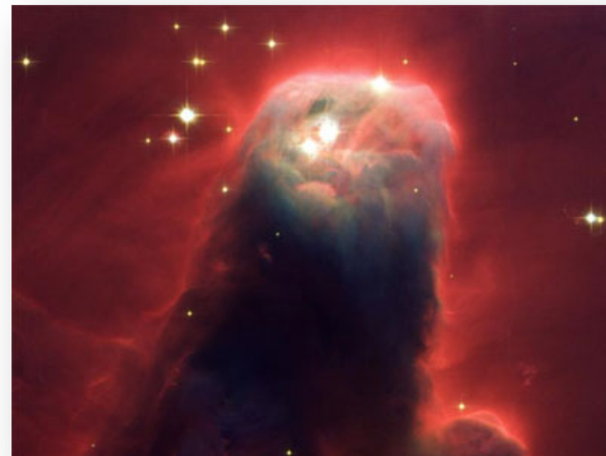
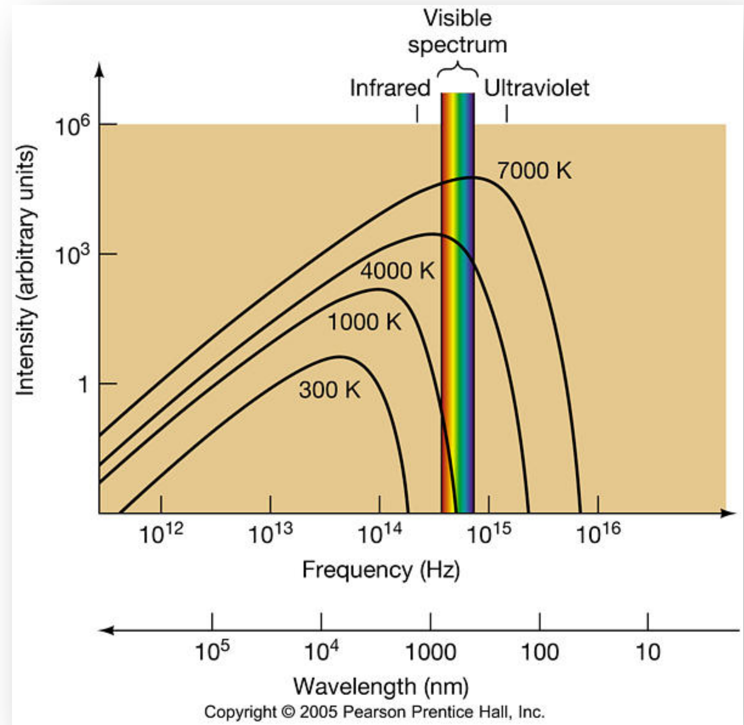


Image Credit: NASA/JPL

Thermal Emission

- Objects with a temperature T emit thermal emission
  - Idealized spectrum of thermal emission is the Planck function
  - Often called `blackbody emission'
  - Cooler object emit more energy at longer wavelengths
- Dust is an efficient radiator of thermal energy
- For dust between stars, emission is dominated by Far-IR to sub-mm wavelengths.



Mode and Normalization of BB increase as temperature increases

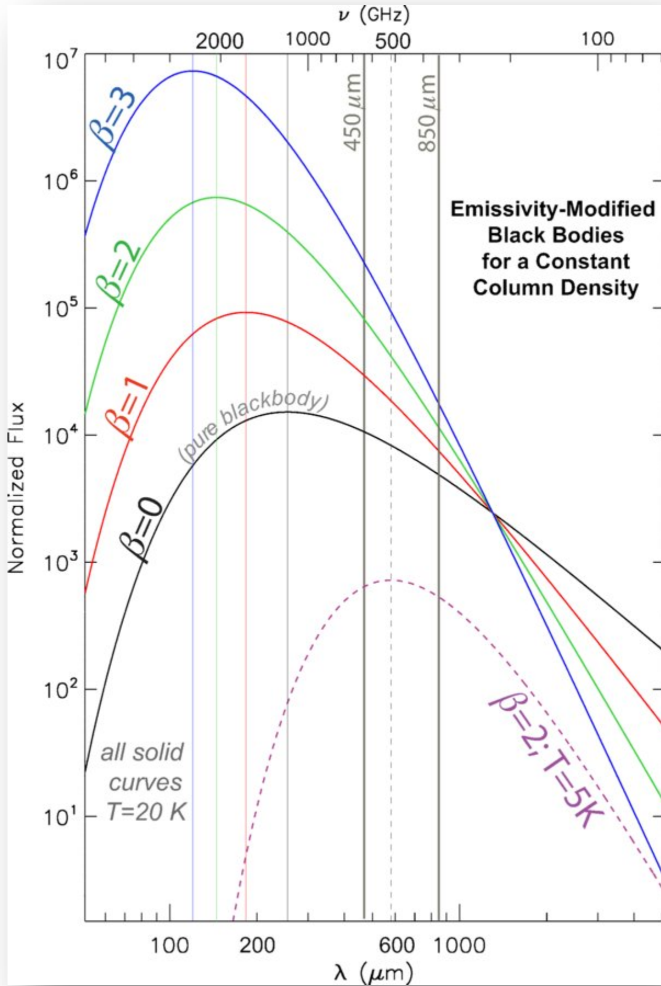
Modeling Dust Emission

- Model dust brightness as a modified black-body:

$$f(\nu) \propto N\nu^{\beta} B_{\nu}(T),\ B_{\nu}(T) \propto \nu^{3}\left[\exp\left(\frac{h\nu}{kT}\right)-1\right]^{-1}$$

- Parameters are the dust temperature, T, the power-law modification index, β, and the column density, N

- β -> 0 as dust grains become larger

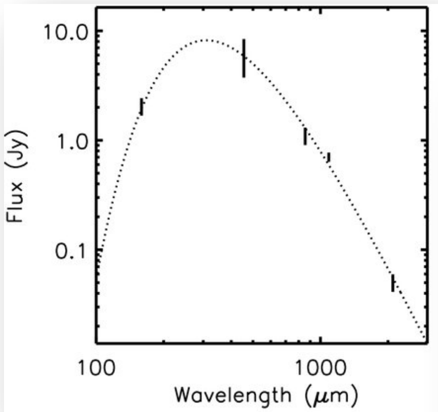


Shetty+2009

Astrophysical Expectations

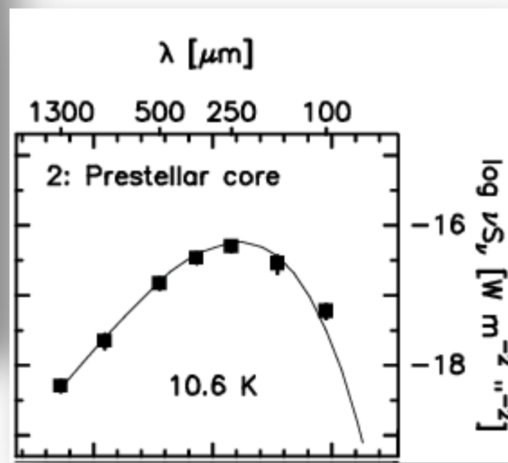
- Dust coagulation: Higher dust agglomeration in dense regions should create larger grains
  - β should decrease toward denser regions
- For dust between stars with silicate and graphite composition, $\beta \approx 2$
- For disks of dust around new stars, observations find $\beta \leq 1$.
- Temperature should decrease in higher density regions
  - Higher density regions more effectively shielded from ambient radiation field

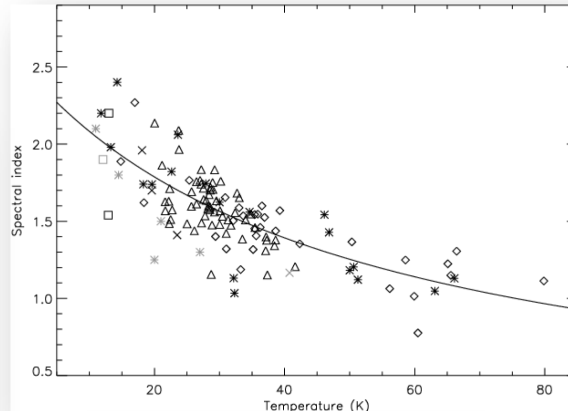
Previous Observational Results



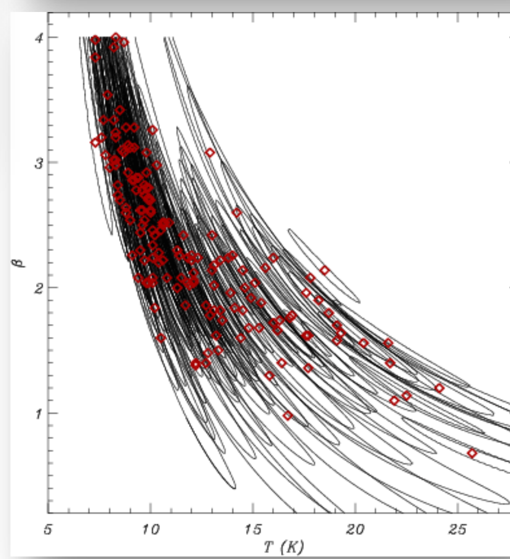
Schnee+2010

Stutz+2010





Dupac +2003



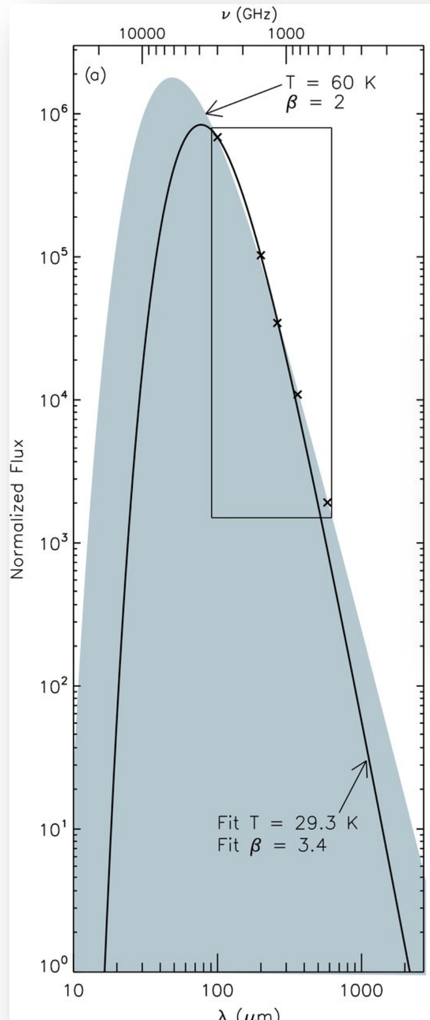
Désert et al., 2008

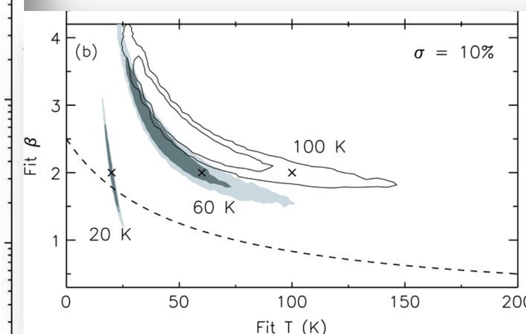
Observations find β and temperature anti-correlated (e.g., Dupac+2003, Desert+2008, Anderson+2010, Paradis+2010,Planck Collaboration 2011)

Anti-Correlation is Unexpected

Unexpected: Why a β–T anti-correlation?

- β and T estimated by minimizing weighted squared error (i.e., χ^2)
- Typically only 5–10 data points for estimating 3 parameters
- Errors on estimated β and T large, highly anti-correlated
- Errors bias inferred correlation, may lead to spurious anti-correlation (Shetty+2009)





Shetty+2009

The Measurement Model

$$y_{ij} = \delta_j S_j(N_i, T_i, \beta_i) + \varepsilon_{ij}$$ ← Measured fluxes

$$S_j(N_i, T_i, \beta_i) \propto N_i \left(\frac{\nu_j}{\nu_0}\right)^{\beta_i} B_{\nu_j}(T_i)$$ ← Astrophysical model for flux

$$\varepsilon_{ij} \sim N(0, \sigma_{ij}^2), \quad \log \delta_j \sim N(0, \tau_j^2)$$ ← Measurement errors

y_{ij} : Measured flux value at j^{th} wavelength for i^{th} data point
$i = 1,\ldots,n \quad j = 1,\ldots,p$
N_i, T_i, β_i : model parameters for i^{th} data point
ν_j : Frequency corresponding to j^{th} observational wavelength
δ_j = Calibration error for j^{th} observational wavelength
ϵ_{ij} = Measurement noise for y_{ij}

Our Hierarchical Model

Joint distribution of log N, log T, and β modeled as a multivariate Normal distribution

$$\mu,\Sigma \sim Uniform$$

$$\log N_i, \log T_i, \beta_i \mid \mu,\Sigma \sim N(\mu,\Sigma)$$

$$\log \delta_j \sim N(0,\tau_j^2)$$

$$y_{ij} \mid N_i, T_i, \beta_i, \delta_j \sim N(\delta_j S_j(N_i, T_i, \beta_i), \sigma_{ij}^2)$$

τ_j and σ_{ij} are considered known and fixed

Calibration Errors are nuisance parameters

Naïve MCMC Sampler

1. Update calibration error, do Metropolis-Hastings (M-H) update with proposal

$$\log \delta_j^{\text{prop}} \sim N(\log \bar{\delta}_j, Var(\bar{\delta}_j)/\bar{\delta}_j^2)$$

$$\bar{\delta}_j = Var(\bar{\delta}_j)\left(\frac{\bar{\delta}_j^{\text{data}}}{v_j^{\text{data}}} + \frac{1}{\tau_j^2}\right), \quad Var(\bar{\delta}_j) = \left(\frac{1}{v_j^{\text{data}}} + \frac{1}{\tau_j^2}\right)^{-1}$$

$$\bar{\delta}_j^{\text{data}} = v_j^{\text{data}} \sum_{i=1}^{n} \frac{y_{ij}}{S_j(N_i,T_i,\beta_i)}\left(\frac{\sigma_{ij}}{S_j(N_i,T_i,\beta_i)}\right)^{-2}, \quad v_j^{\text{data}} = \left(\sum_{i=1}^{n}\left(\frac{\sigma_{ij}}{S_j(N_i,T_i,\beta_i)}\right)^{-2}\right)^{-1}$$

2. Update values of N_i, T_i, and β_i using M-H update with multivariate normal proposal density
3. Do Gibbs update of μ and Σ using standard updates for mean and covariance matrix of Normal distribution

Ancillary-Sufficiency Interweaving Strategy

- Naïve MCMC sampler is very slow due to strong dependence between δ, log N, β, and μ:

$$E(y_{ij} \mid N_i, T_i, \beta_i, \delta_j) \propto \delta_j N_i \left(\frac{\nu_j}{\nu_0} \right)^{\beta_i} B_{\nu_j}(T_i)$$

- Use Ancillary-Sufficiency Interweaving Strategy (ASIS, Yu & Meng 2011) to break dependence

References:

Yu, Y, & Meng, X-L, 2011, JCGS (with discussion), 20, 531

Kelly, B.C., 2011, JCGS, 20, 584

Introduce Sufficient Augmentation

- Original data augmentation is an *ancillary* augmentation for calibration error: δ and (β,N,T) independent in their prior

- Introduce a *sufficient* augmentation for calibration error:

$$\log\tilde{\delta}_i = \log\delta + \mathbf{X}\theta_i$$

$$\mathbf{X} = \begin{pmatrix} 1 & \log\frac{\nu_1}{\nu_0} \\ \vdots & \vdots \\ 1 & \log\frac{\nu_p}{\nu_0} \end{pmatrix}, \quad \theta_i = (\log N_i, \beta_i)^T$$

Sufficient Augmentation Model

- New Model is

$$
\begin{aligned}
&y_{ij} \mid T_i, \tilde{\delta}_{ij} \sim N(\tilde{\delta}_{ij} B_j(T_i), \sigma_{ij}^2) \\
&\log \tilde{\delta}_i \mid \log \tilde{\delta}_k, \theta_i, \theta_k = \log \tilde{\delta}_k - \mathbf{X}(\theta_k - \theta_i), i \neq k \\
&\log \tilde{\delta}_k \mid \theta_k \sim N(\mathbf{X}\theta_{\mathbf{k}}, V_\delta), \quad V_\delta = \mathrm{diag}(\boldsymbol{\tau}_{\mathrm{j}}^2) \\
&(\theta_i, \log T_i) \mid \mu, \Sigma \sim N(\mu, \Sigma)
\end{aligned}
$$

ASIS for this problem

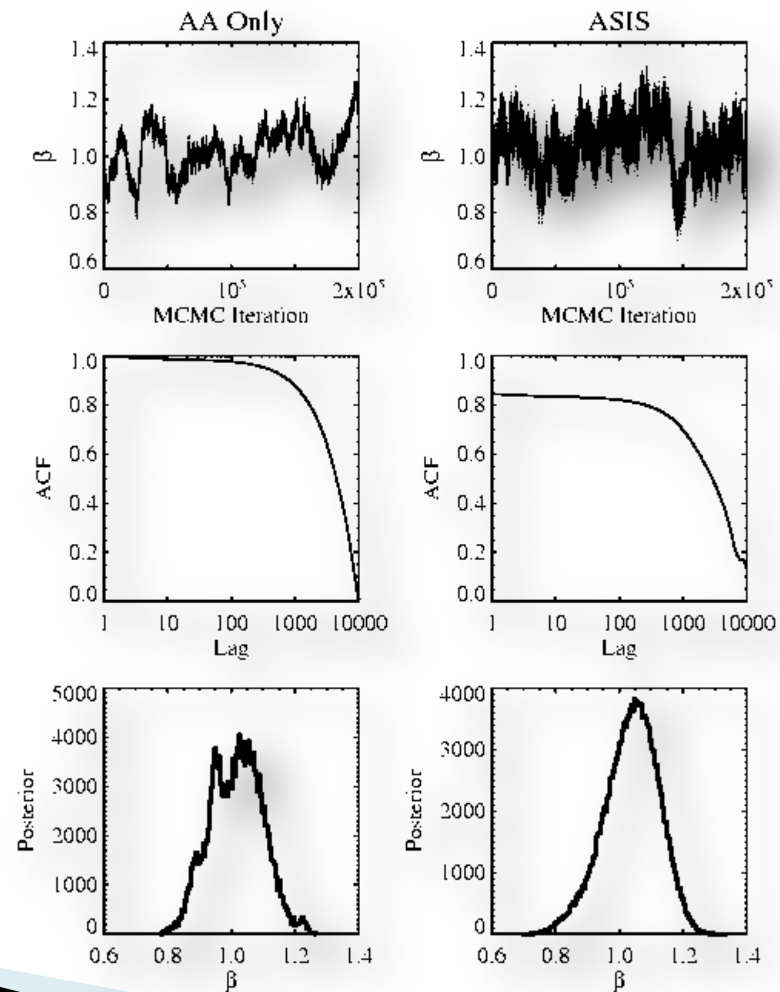
1. Update calibration error as before
2. Update δ, log N, and β under SA:

$$
\begin{aligned}
&\theta_k^{new} \mid \tilde{\delta}_k \sim N(\hat{\theta}_k, V_k) \\
&\hat{\theta}_k = V_k \mathbf{X}^T V_\delta^{-1} \tilde{\delta}_k \\
&V_k = (\mathbf{X}^T V_\delta^{-1} \mathbf{X})^{-1} \\
&\theta_i^{new} \mid \theta_k^{new}, \tilde{\delta}_k, \tilde{\delta}_i = \theta_k^{new} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\tilde{\delta}_i - \tilde{\delta}_k) \\
&\delta = \tilde{\delta}_k - \mathbf{X} \theta_k^{new}
\end{aligned}
$$

(also need to update μ appropriately)

3. Update log N, log T, and β as before under AA
4. Update μ and Σ

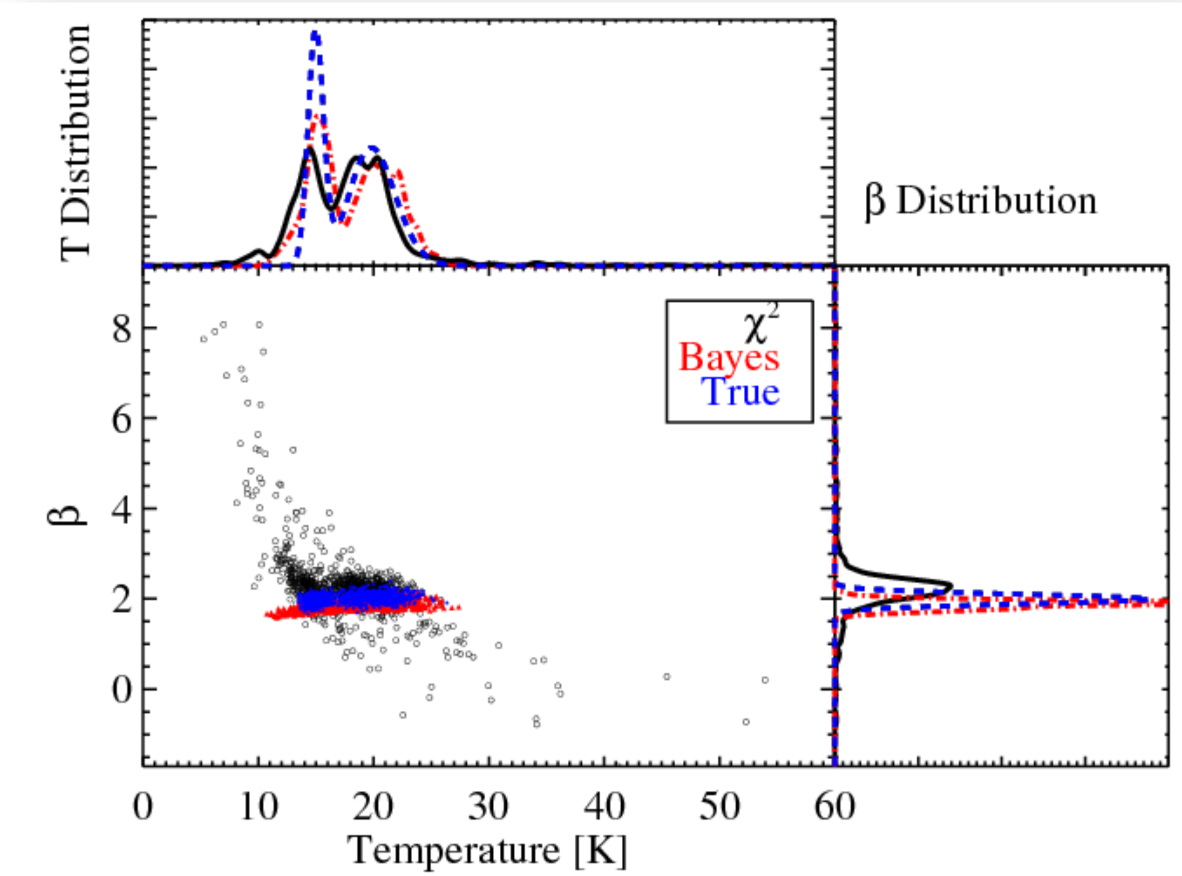
ASIS Significantly Improves MCMC Efficiency for this Problem



MCMC results for data point with highest S/N, 200,000 iterations

Kelly 2011

Test: Comparison of Bayesian and χ^2 methods for Intrinsic Correlation

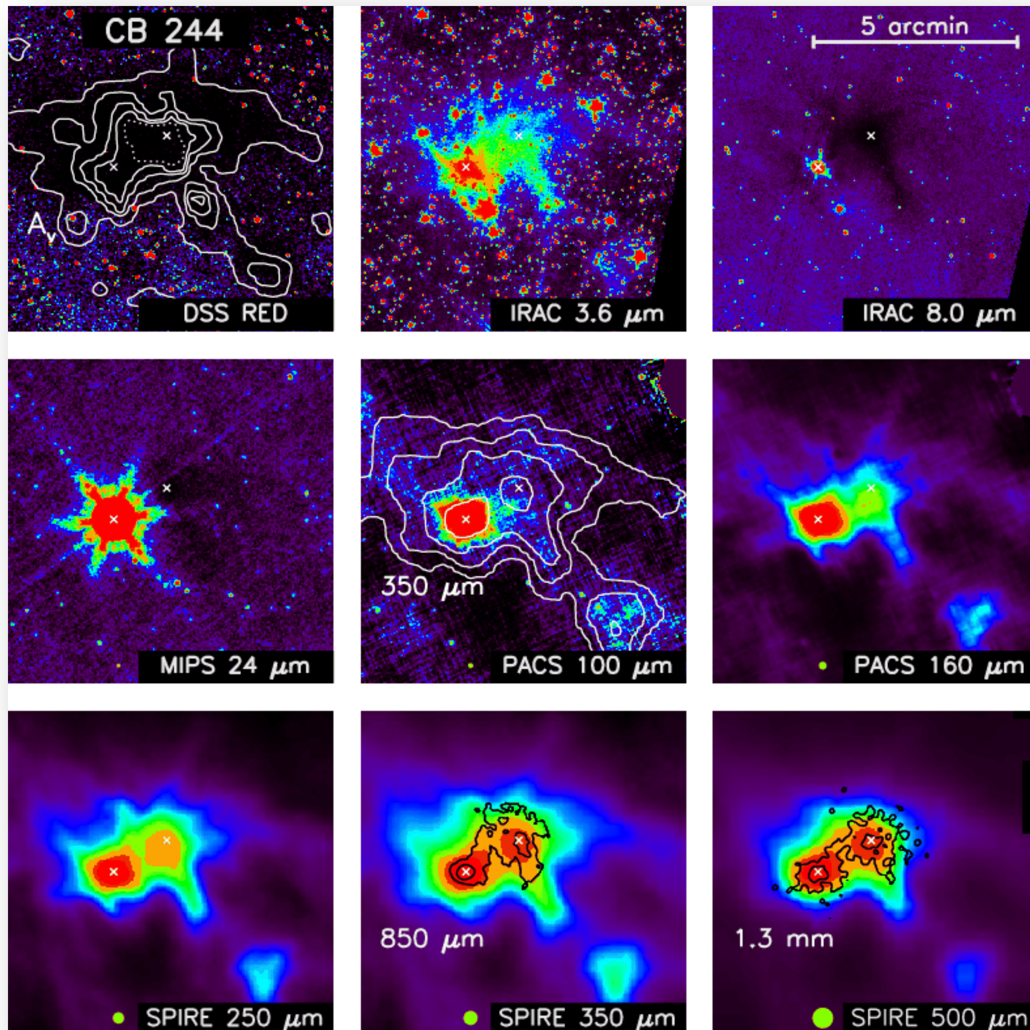


Rank correlation coefficients:

True = 0.33
χ^2 = −0.45±0.03
Bayes = 0.23±0.08

Application: Star-Forming Bok Globule CB244

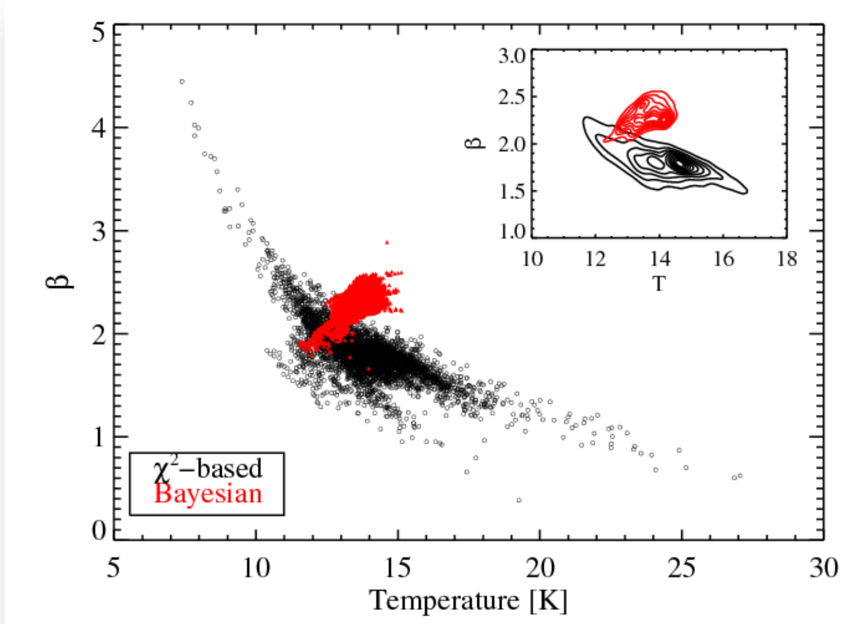
CB244: Small, nearby molecular cloud containing a low-mass protostar and starless core



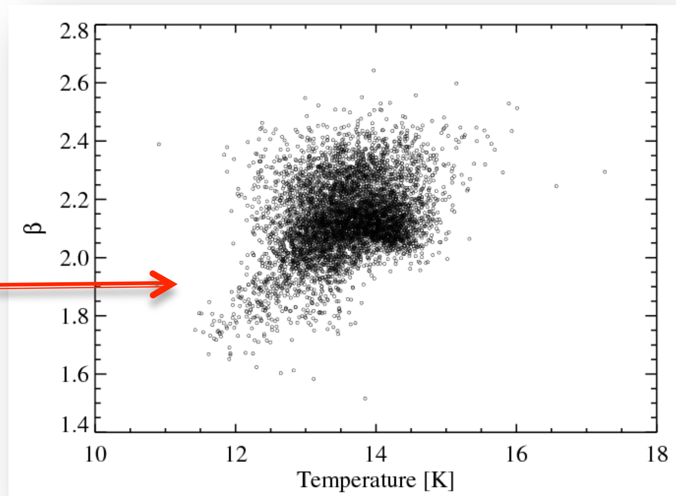
Stutz+2010

CB244 Results:

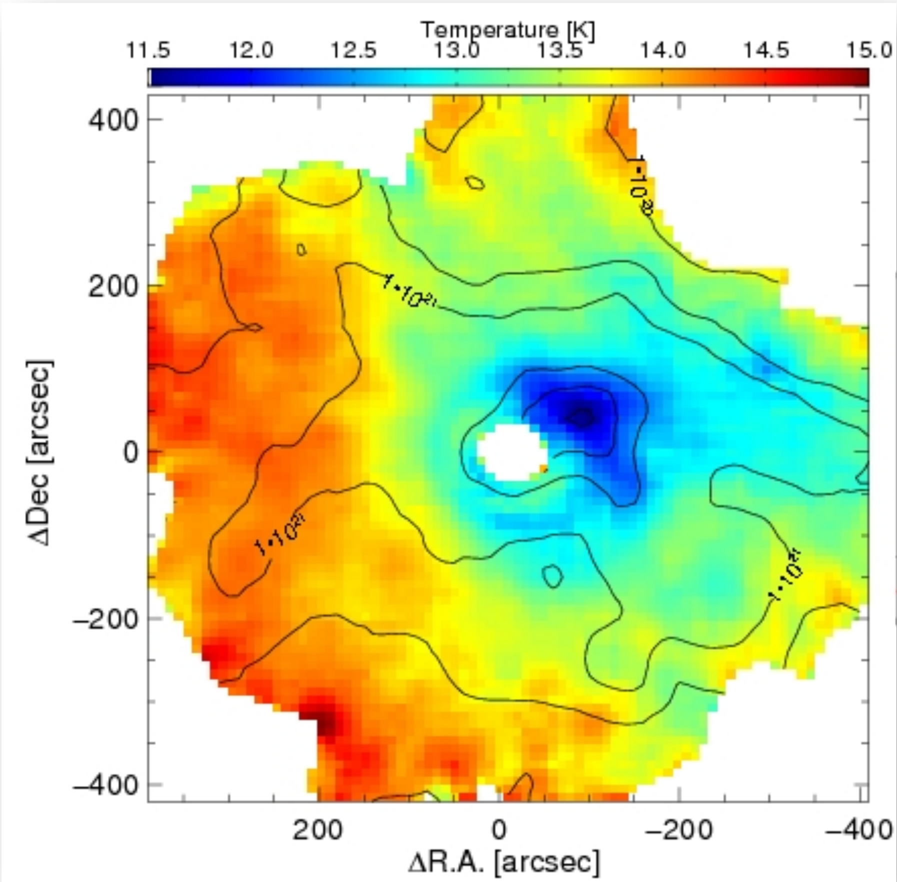
- β and T ***anti-correlated*** for χ^2-based estimates, caused by noise
- β and T ***correlated*** for Bayesian estimates, opposite what has been seen in previous work

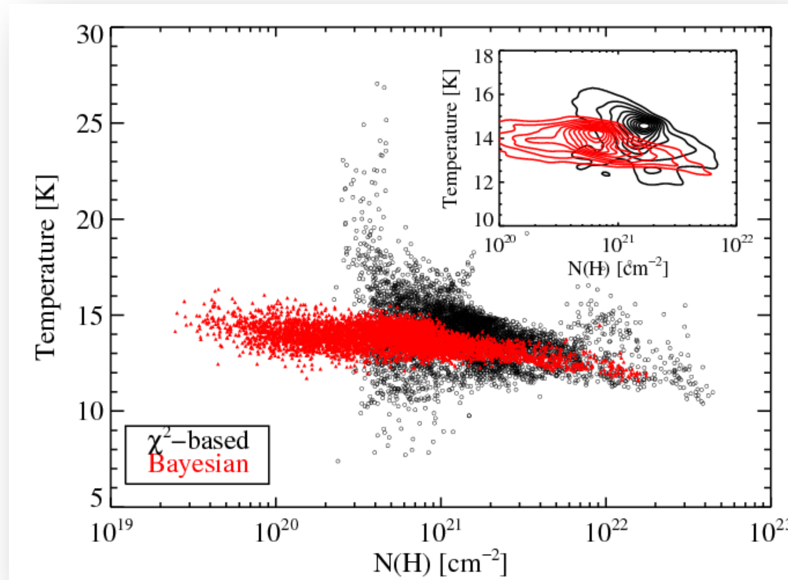


Random Draw from Posterior Probability Distribution



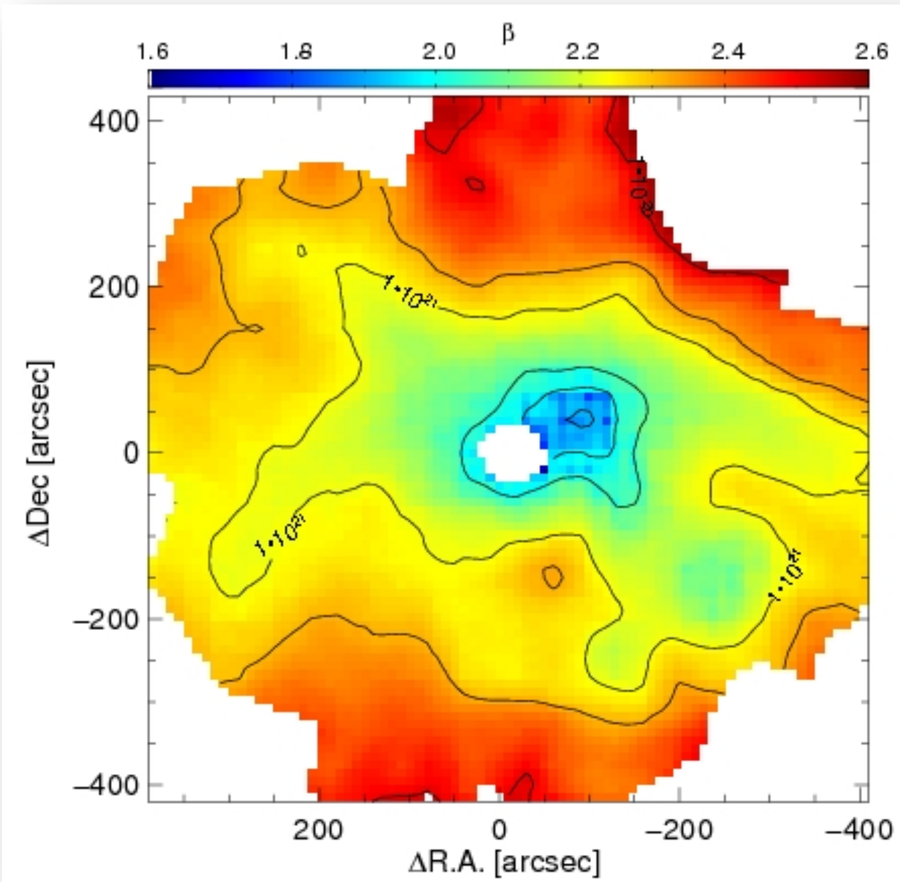
CB244: Temperature and Column Density Maps

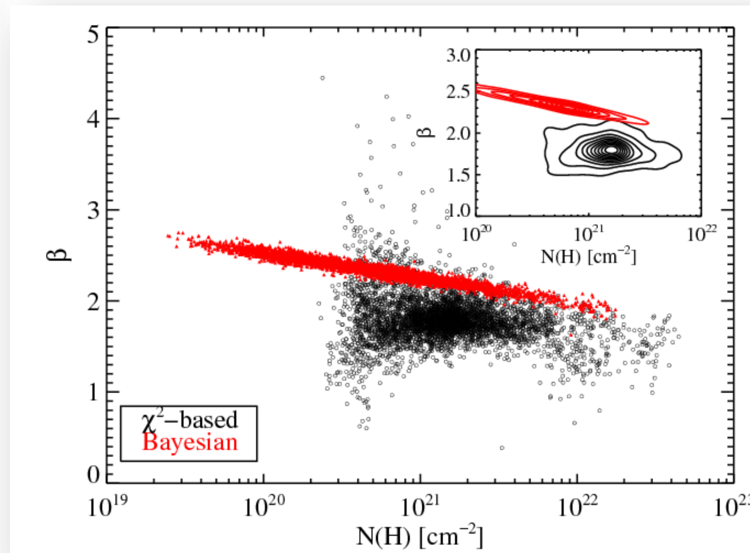




Temperature tends to decrease toward central, denser regions

CB244: Column Density and β Maps





β decreases toward central denser regions

Evidence for Grain Growth?

- β is expected to be smaller for larger grains
- $\beta < 1$ for protoplanetary disks, denser than starless cores
- For CB244 we find β is smaller in denser regions
- Suggests dust grains begin growing on large scales before star forms
- Qualitatively consistent with spectroscopic features seen in mid-IR observations of starless cores (Stutz+2009)

Summary

- Introduced a Bayesian method that correctly recovers intrinsic correlations involving N,β, and T, in contrast to traditional χ^2
- Developed an ancillarity-sufficiency interweaving strategy to make problem computationally tractable
- For CB244, our Bayesian method finds that β and T are correlated, opposite that of χ^2
- For CB244, we also find that β increases toward the central dense regions, while temperature decreases
- Increase in β with column density may be due to growth of dust grains through coagulation
- Bayesian method led to very different scientific conclusions compared to naïve χ^2 method, illustrating importance of proper statistical modeling for complex astronomical data sets