PHYSICS 201 FINAL PROJECT

CATHERINE ZUCKER

1. Introduction

Given its unique standing as our "home" galaxy, the Milky Way is an ideal test case for studying the formation and evolution of galaxies, and much is already known about its physical makeup. In order to place this knowledge into a larger cosmological context, we must understand its underlying structure. However, in order to delineate this spiral structure and better refine our knowledge of its stars and gas, we need accurate distance measurements. There are currently two major distance methods employed in our Galaxy: kinematic distances and photometric distances, with kinematic being much more common. The kinematic method uses a combination of an object's radial velocity, along the with the assumption of purely circular rotation, to derive a distance. For instance, carbon monoxide (CO) emission, which is the primary tracer of spiral arm structure, emits a spectral line at 115 GHz. This line is shifted to various frequencies, which can then be converted to a radial velocity via the doppler effect. From this information we can create spectral cubes, which show a spatial template of the CO emission along the same line-of-sight at various velocity slices. If we know this radial velocity, along with the longitude of the gas cloud, the angular velocity of the Sun, and an assumed rotation curve, we can calculate a kinematic distance to each slice (Roman-Duval et al. 2009). Unfortunately, very rarely do stars and gas travel on purely circular orbits. Particularly within the solar neighborhood and towards the Galactic center, motion is dominated by large non-circular and peculiar velocities, which leads to inaccurate distances (Reid et al. 2009). In contrast, the second distance method employed does not rely on a rotation curve: it is a photometric, dust-based distance method, which uses stellar optical and near-infrared photometry within a Bayesian framework to determine probabilistic distance and reddening towards a billion stars in the Milky Way (Green et al. 2014, 15).

2. Experiment

In this paper, we use joint stellar posteriors on distance and reddening derived from the photometric distance method to refine the distance to the CO cloud Cepheus given by the kinematic distance method. In doing so, we ultimately hope to refine our knowledge of the Galactic rotation curve for the Milky Way, thus allowing us to extract better kinematic distances from spectral line information. We choose the Cepheus molecular cloud because it is centered at a Galactic longitude and latitude, $(l, b) = (109°, 13°)$, so it is located away from the confusion of the Galactic plane. It is also a relatively compact molecular cloud, which is less computationally expensive than a cloud that is angularly broad. Finally, other works have used the photometric distance method to determine a distance to the Cepheus cloud (e.g. Schlafly et al. 2014) and have found the cloud to be broken up into several discrete distance components, depending on the line-of-sight. We would like to see how our distance determinations, produced through parallel-tempering Markov Chain Monte Carlo, compare to existing distances from the literature, both kinematic and photometric.

3. Statistical Modeling

We would like to infer the distance to twelve velocity slices of a spectral cube, covering the molecular cloud complex Cepheus. We need two datasets. First, we need the photometry of stars towards Cepheus, which have been used in Green et al. 2015 to derive joint posteriors on distance and reddening for each star. Second, we need the intensity of carbon monoxide emission in each velocity slice of a spectral cube towards this same region; we use the Galaxy-wide CO survey from Dame et al. 2001.

This is what we ultimately wish to infer:

$$p(\alpha \mid m, I)$$

where

$$\alpha = (d_1, d_2, ..., d_{12})$$

and...

d=distance to slice (12 d's for 12 velocity slices)
I= gas emission intensities that will act as dust templates, and which will be converted to extinction with a gas-to-dust coefficient
m= set of stellar photometry for all stars

From Bayes Rule, we know that, up to a normalizing constant:

$$p(\alpha \mid m, I) \propto p(m, I \mid \alpha) \times p(\alpha)$$

with the first term being our likelihood function and the second term being our prior function; we're assuming a flat prior for α.

As explicated in Green et al. 2015, we can show that our likelihood function for each star is simply the line integral over the joint posterior on distance and reddening for that star, taken along some specified reddening profile (cumulative reddening as a function of distance). We introduce a nuisance parameter, θ, the stellar type, and we re-express the likelihood function as an integral over the distance modulus μ and the stellar type θ:

$$p(m, I \mid \alpha) \propto \int d\mu \, d\theta \, p(m, I, \mu, \theta \mid \alpha)$$

We can factor this into two terms:

$$\propto \int d\mu \, d\theta \, p(m, I \mid \mu, \theta, \alpha) p(\mu, \theta \mid \alpha)$$

And, if we assume that μ and θ are both independent of the reddening profile, defined by the α coefficients, we get

$$\propto \int d\mu \, d\theta \, p(m, I \mid \mu, \theta, E(\mu; \alpha)) \, p(\mu, \theta)$$

The integrand in the last equation is basically (up to a normalizing constant) a prior times a likelihood, so our likelihood for one star is essentially the integral of the posterior density $p(\mu, E, \theta \mid m)$. We marginalize out stellar type, so we're left with the integral over μ. The posterior density for each star has already been precomputed in Green et al. 2015. Given that the likelihood for all stars is given by the product of the individual likelihoods, our posterior for α is given by:

$$p(\alpha \mid I, m) \propto p(\alpha) \times \prod_j \int p(\mu_j, E_j(\mu_j; \alpha) \mid m_j, I) \, d\mu_j$$

Now we need a model for the reddening profile, which defines the path of our line integral through the stellar posterior arrays. We parameterize the line-of-sight reddening profile towards an individual star, E_j, with the following model:

$$E_j(\mu_j, \alpha) = \sum_{k \to d_k < \mu_j} (I_{j,k} + \delta I_{j,k}) \times c$$

with $I_{j,k}$= intensity of gas emission in the pixel corresponding to the jth star in the kth velocity slice, $\delta I_{j,k}$= uncertainty in the intensity measurement, and c= gas to dust conversion coefficient from the literature. We're summing over the emission in all the velocity slices such that the distance to the velocity slice is less than the distance to our star μ. The form of the reddening profile will be piecewise linear. Our coefficient converts CO intensity in units of Kelvin to reddening in units of magnitudes. To derive this, we convert from CO intensity to atomic hydrogen column density (Dame et al. 2001) and then from atomic hydrogen column density to reddening (Schlegel, Finkbeiner, and Davis et al. 1998). Finally, we multiply by the width of each spectral channel in km s^{-1} to get units of mag/K:

$$\mathrm{c} = (3.6 \times 10^{20} \ \mathrm{H\ cm^{-2}\ K^{-1}\ km^{-1}\ s}) \times (\frac{1 \ \mathrm{mag\ E(B-V)}}{8 \times 10^{21} \ \mathrm{H\ cm^{-2}}}) \times (\frac{0.65 \ \mathrm{km\ s^{-1}}}{\mathrm{channel}})$$
$$= \frac{0.03 \ \mathrm{mag}}{K}$$

The intensity noise $\delta I_{j,k}$, is equivalent to the Gaussian rms noise per velocity channel, given by 0.16 K (Dame et al. 2001). While each slice has the same noise, the total noise for the cumulative reddening will increase as you increase the number of slices summed. Converting to units of reddening, the noise of the first slice is a Gaussian with standard deviation equivalent to 0.16 K $\times$ 0.03 mag/K = 0.005 mag. For the combination of the first and second slices, the total noise is given by the convolution of two Gaussians of the same standard deviation. Since the variances add in quadrature, $\sigma_{total}^2 = \sigma_1^2 + \sigma_2^2$. Thus, the standard deviation of the noise in the nth slice is given by $\sqrt{n \times 0.005^2}$. We apply this uncertainty to the reddening profile by considering the probability not only in the bins defining the central reddening ledge, but also in the bins within one and two rows adjacent to them. We define a reddening "ledge" as a plateau of constant reddening between two of our distance slices. We weight the central, first, and second bins by calculating the area under the Gaussian within 0.005, 0.015, and 0.025 in reddening, respectively. Since the width of each reddening bin in the stellar posterior array is 0.01 mag, this corresponds to the probability of the reddening value falling within the center, ±one, and ±two bins. The weighting provided to each row is proportional to the area of the Gaussian falling within that row, calculated via a cumulative distribution function.

4. Results and Discussion

We apply parallel tempering Markov Chain Monte Carlo analysis to a small 0.4 sq. deg subset of the Cepheus CO cloud, containing over 3000 stars. We employ fifty walkers at five different temperatures, which each run 500 steps each. We discard the first 300 steps of each chain as burn-in, and only consider the coldest chain in our analysis, corresponding to a $\beta = 1$ value. Total CPU time was $\approx$ 15 hours. We set off our walkers for all slices at the kinematic distance for Cepheus given by the literature, $\mu = 9.3$ (Blaauw 1964), corresponding to 730 pc. As the velocity gradient over this region is small ($\approx$ 8 km s^{-1}) we can place the bulk of this cloud at this kinematic distance. In contrast to the kinematic distances, we find that our dust-based distances show significantly more dispersion, varying from $\mu = 7.71 - 9.87$ or 348-941 pc (see Table 1 below). This is perhaps not surprising, as kinematic distances are known to be unreliable when applied to clouds near the Sun, as Cepheus is. In purely circular rotation, clouds near the Sun would have almost zero angular velocity relative to the Sun. While the local clouds are orbiting the Galactic center at similar angular velocity as the Sun, they are dominated by large peculiar velocities, because the Sun itself is not moving in a purely circular orbit. Thus kinematic distances will usually be inaccurate for clouds within 1 kpc of the Sun. While generally inconsistent with the kinematic distance approach, we find that our results *are* consistent with other distances derived from the photometric method. Schlafly et al. 2014 estimate the distances to dozens of local CO clouds through the clouds' reddening and absorption of starlight. In a two step process, they first determine the reddening and distances to stars within a certain sightline towards the cloud and then combine this information to derive the reddening as a function of distance for all the stars in that sightline. We compile all the Schlafly et al. 2014 distances to Cepheus for sightlines within or adjacent to our 0.4 sq. deg region of interest. They span a wide range of distances, including 369 pc, 396 pc, 678 pc, 859 pc, 883 pc, and 971 pc. We conclude that, consistent with other dust-based distance estimates, Cepheus is indeed broken up into several distance components, ranging from $\approx 400 - 900$pc, and that kinematic distances are unreliable in pinpointing its location.

Table 1

Slice	Distance Modulus	Distance (pc)
d1	$8.41 + 0.16 - 0.43$	480
d2	$8.26 + 0.37 - 0.52$	453
d3	$7.71 + 0.73 - 0.24$	348
d4	$9.18 + 0.08 - 0.48$	685
d5	$7.65 + 0.33 - 0.26$	339
d6	$7.84 + 0.27 - 0.24$	367
d7	$8.69 + 0.40 - 0.58$	547
d8	$8.46 + 0.26 - 0.98$	492
d9	$9.87 + 0.75 - 0.10$	941
d10	$8.06 + 0.53 - 0.66$	409
d11	$9.10 + 2.11 - 0.26$	660
d12	$9.26 + 0.10 - 0.84$	711

5. Conclusion

While our results are promising, significantly more work needs to be done to verify them. Ideally, we would run our code on a larger region of the cloud (3x3 deg) with thousands of more stars (50,000+), which would provide more robust estimates and reduce degeneracies: the more stars we include, the fewer possible combination of distance estimates that would produce a reddening profile with maximum probability. Moreover, in the future, we plan to let our gas-to-dust conversion coefficient c float, rather than being fixed by the literature value of 0.03 mag/K. There is no evidence to suggest that the conversion coefficient should be constant, so letting it float will give our fit the freedom to explore new, potentially higher likelihood, reddening profiles. Finally, upcoming photometric data from near infrared telescopes like GAIA and UKIDSS will only further improve our distance resolution, as near-infrared wavelengths in particular can probe further into dusty CO clouds, maximizing the number of stars available for study towards regions of interest closer to the Galactic plane.

6. References

Blaauw 1964, *ARA&A*, 2, 213
Dame et al. 2001, *ApJ*, 547, 792
Green et al. 2014,*ApJ*, 783, 114
Green et al. 2015 *ApJ*, 810, 25
Reid et al. 2009, *ApJ*, 700, 137
Roman-Duval et al. 2009, *ApJ*, 699, 1153
Schlafly et al. 2014, *ApJ*, 786, 29
Schlegel et al. 1998 *ApJ*, 500, 525