Radial Velocity & Markov Chain Monte Carlo Simulations

Recovering Orbital Parameters of Binary Systems

William Pugsley, Nicole Xu

Friday 15th April, 2022

Abstract

Generally, exoplanets can't be directly observed due to the high brightness contrast between the two objects, and limited telescope resolution. However, it's relatively easy to obtain radial velocity measurements by observing Doppler Shift. These radial velocity measurements can be used to recover information about the binary system, including orbital period, eccentricity, and mean velocity of the system's CoM. We modeled the posterior (parameters given data) with Bayes' Theorem and sample this posterior distribution with a Markov Chain Monte Carlo (MCMC). We found that the orbital parameters can be accurately recovered for both simulated data (including noise) as well as for radial velocity data from real known Binary Systems.

1 Introduction

Under Newton's laws of motion, when two objects that interact only through gravity form a closed orbit this orbit takes the shape of an ellipse [5]. In the center of mass frame, each object orbits about the center of mass in an elliptical fashion. Both of these orbits have the same eccentricity, e, and orbital period, T, but different semi-major axes. These results hold in an inertial reference frame moving relative to this center of mass with velocity v_0. The only difference is that the orbital plane of the two objects is also moving with this velocity; the objects are still confined to this orbital plane. From the perspective of an observer in this inertial reference frame, the motion of either of these objects can be decomposed in radial motion, directly towards or away from an observer along their line of sight, and into perpendicular motion, in the plane perpendicular to their line of sight.

If a star of mass M has an orbiting object of mass m, then the star will orbit about its center of mass with the object. It is expected that at some times during this orbit the motion of the star will be adding to or subtracting from the radial velocity of the system as seen by an observer. This will cause fluctuations about v_0. The radial velocity, v_r, of a star as a function of time, t, in terms of previously defined parameters is [8]:

$$v_r(t) = \kappa \cdot [\cos(f + \omega) + e \cdot \cos\omega] + v_0, \tag{1}$$

where κ and f are related to the parameters of the system by

$$\tan(f/2) = \sqrt{\frac{1+e}{1-e}} \tan(u/2), \tag{2a}$$

$$u - e \cdot \sin u = \frac{2\pi}{T}(t - \tau), \tag{2b}$$

$$\kappa = \frac{(2\pi G)^{1/3} m}{(M+m)^{2/3}} \cdot \frac{\sin I}{\sqrt{1-e^2}}. \tag{2c}$$

In these equations, G is the gravitational constant, I is the inclination of the orbital plane with respect to the line of sight, ω is the angle of the pericenter as measured from the ascending node, and τ is the time of passage through the pericenter [4] [8].

Given a radial velocity curve, it is possible to use Eq. (1) to determine possible values for the orbital parameters of the system. This is especially useful when the orbiting object is, for whatever reason, not directly observable.

2 Methods

2.1 Markov Chain Monte Carlo

Bayes' Theorem provides a solid mathematical framework for parameter inference. It allows us to write the posterior distribution, the probability distribution of specific parameters values, $\vec{\theta}$, given the data, as follows:

$$p(\vec{\theta}|data) = \frac{p(data|\vec{\theta})p(\vec{\theta})}{p(data)}, \tag{3}$$

where $p(data|\vec{\theta})$ is the likelihood function that links our parameters to our data, $p(\vec{\theta})$ is our prior belief on our parameters, and $p(data)$ is our evidence, which is the probability of measuring the data. This can be hard to find, but if we just want to know what the shape of $p(\vec{\theta}|data)$ looks like, we can simplify Bayes' Theorem to:

$$p(\vec{\theta}|data) \propto p(data|\vec{\theta})p(\vec{\theta}). \tag{4}$$

If we assume Gaussian errors, we get the following likelihood function:

$$p(v_i|\vec{\theta}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{(v_i - model(t_i))^2}{2\sigma_i^2}\right], \tag{5}$$

where $model(t_i)$ is Eq. (1) evaluated at t_i, v_i is the i'th radial velocity data point with error σ_i, and n is the total number of data points.

For our prior distribution of a given parameter, we used a uniform distribution:

$$p(\theta_i) = \begin{cases} \dfrac{1}{high - low} & low \leq \theta_1 \leq high \\ 0 & otherwise \end{cases}$$

with *low* and *high* being the upper and lower limits of the allowed values of each parameter, detailed in Table 1. We then took the log of each distribution. Since each parameter is independent, $p(\vec{\theta}) = p(\theta_1) \cdot p(\vec{\theta_2}) \cdot ... \cdot p(\vec{\theta_m})$ is the product of all the individual priors for each parameter.

However, with high dimensional parameter spaces, you would need to plot a *very* large number of data points to plot the posterior by brute force. Since our problem involves a 6-dimensional parameter space, this was not computationally feasible. Instead, we can use an Markov Chain Monte Carlo (MCMC) to sample the posterior distribution!

An MCMC is a sampling algorithm that takes a biased random walk through parameter space, where the walker spends more time in regions of higher probability. The result is a chain of positions for the walker, which we can plot in a histogram that represents out posterior distribution. To implement an MCMC, we used the `emcee` Python package, which uses the Ensemble sampler MCMC algorithm. This package works with the natural logarithm of Eq. (4). To implement this, we had to select a number of walkers and a burnout phase. We selected 60 walkers to be above the minimum recommended threshold of 2-4x the dimension of

the parameter space (in this case, 6). Based on our trace plots, we estimated the burn-in phase to be $\tilde{5}00$ steps when running the simulation with $40,000$ iterations, which we truncated from our chain before plotting our data.

2.2 Parameters & Priors

In the case of an exoplanet orbiting a main sequence star, m is generally much smaller than M. In this approximation $M + m \approx M$ so κ reduces to $\frac{(2\pi G)^{1/3} m}{M^{2/3}} \cdot \frac{\sin I}{\sqrt{1-e^2}}$. This poses a problem as m, M, and I are now degenerate; changes in the value of one parameter can be accounted for by re-scaling another. Reparametrizing Eq. (1) with

$$\mu = \frac{(2\pi G)^{1/3} m}{M^{2/3}} \cdot \sin I \tag{6}$$

reduces the dimensionality of the problem by two and handles the degeneracy issue.

Table 1: Summary of the range of possible values of each parameter.

Parameter	Lower Limit	Upper Limit
μ	$-1.246059 \cdot 10^6$ m/s$^{2/3}$	$1.246059 \cdot 10^6$ m/s$^{2/3}$
e	0	1
ω	0	2π
T	3282.3503 s	$3.46896 \cdot 10^{13}$ s
τ	3282.3503 s	T
v_0	$-100,000$ m/s	$100,000$ m/s

To establish the range of the priors for e and ω it suffices to understand their definitions. Closed orbits have non-negative eccentricities strictly less than one. ω is an angle that can range from 0 to 2π. To find an appropriate range on μ and T consider the most extreme of their values for already known exoplanets [4] [7] [9]. The time of passage through the pericenter is any time at which the orbiting body is nearest to the star. Since orbits are periodic take the smallest of such possible values. This constrains τ to be less than T. The possible values of v_0 are assumed to the broad so as to not restrict the simulation. The priors are summarized in Table 1.

It is not feasible to sample the phase space over T and τ in a reasonable number of iterations due the large range of possible values. When randomly sampling these parameters the algorithm will yield a disproportionately large probability of finding T on the order of 10^{12} to 10^{13}. However, previous observations do not favor such large periods [3]. To address this problem the MCMC used in this project will sample over the logarithm of T and τ.

2.3 Simulating Data and Uncertainty

In addition to running the MCMC on real data, radial velocity data was created through a simulation to test our MCMC. We randomly generated values for each parameter from our pre-determined allowed values. Then, we used these values in our model to calculate the radial velocity of the simulated binary system at a range of time values.

To simulate values for uncertainty, we first examined the relationship between radial velocity and the uncertainty values.

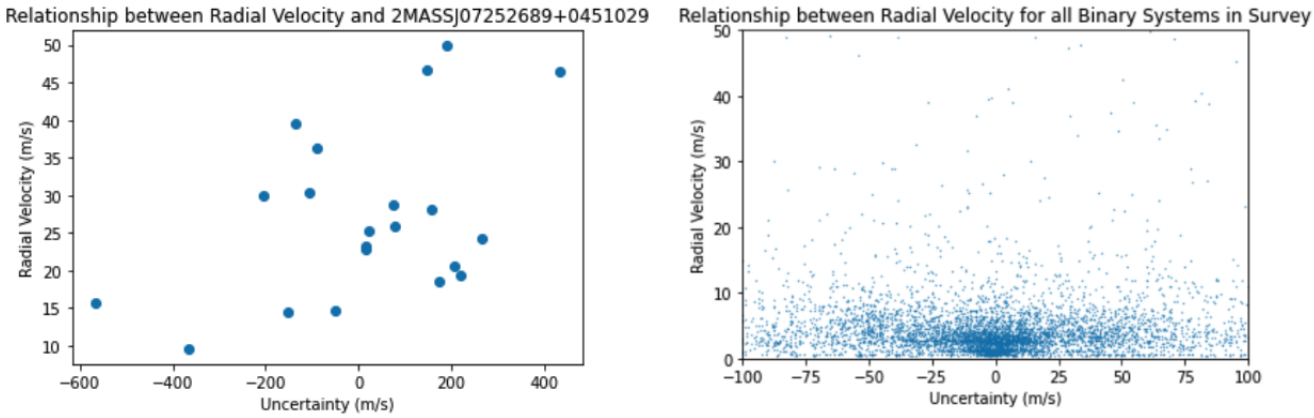

Figure 1: Relationship between radial velocity and uncertainty. Left plot contains data points from a sample Binary System from data. Right plot has all data points from all Binary Systems in data

We found no distinct correlation between the radial velocity value and its associated uncertainty. We then made the assumption that the uncertainty value was uniquely determined by the instrument used to measure the radial velocity data.

For each simulation iteration, we randomly selected an instrument to measure our simulated binary system. We then created an array of all uncertainty values measured with that instrument from the CalTech Exoplanet Data. We then drew randomly from the distribution of uncertainty values to assign an uncertainty value to each radial velocity value.

Finally, we used these uncertainty values to add noise to our radial velocity values. For each radial velocity measurement, we drew from a Gaussian distribution centered at 0 with a standard deviation of 1, then multiplied this by the uncertainty value (to scale it) to simulate noise for that data point. This noise value was then added to the model's radial velocity value to obtain the final simulated radial velocity measurement.

3 Results & Limitations

3.1 HD 129445

The HD 129445 star is known to have at least one exoplanet [2]. This exoplanet was discovered using the radial velocity method with 17 data points [6]. Assuming there are no other bodies in this star system, it is possible to apply the MCMC to find probability distributions for the orbital parameters of this system. The MCMC was run with 60 walkers and 40,000 iterations. The corner plot of the results is shown in Fig. 4 in Appendix A.

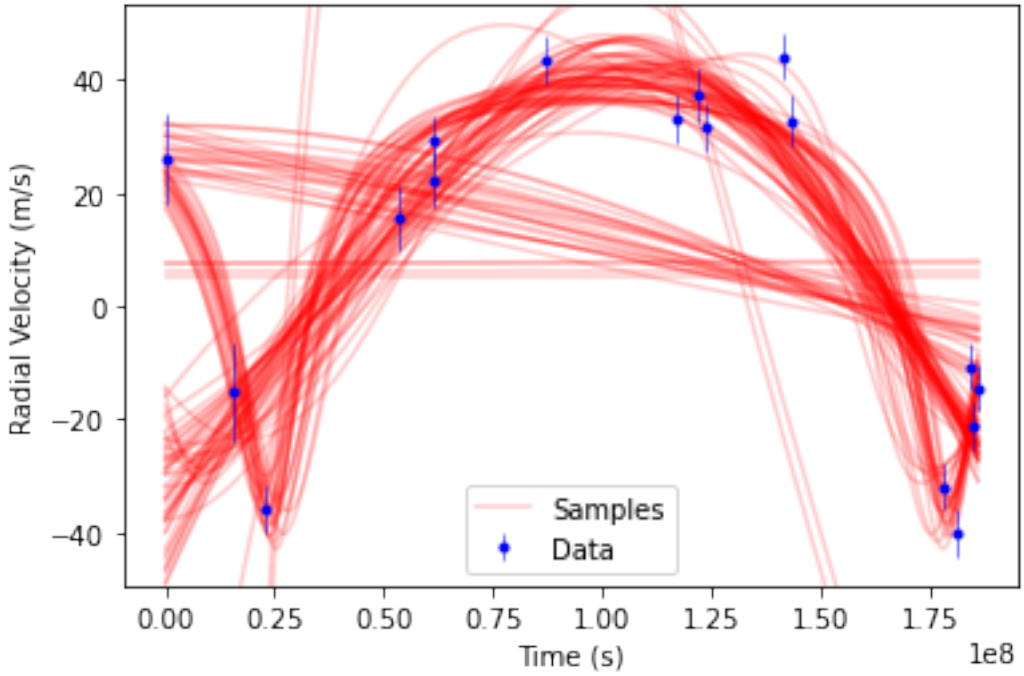

Figure 2: Radial velocity data points for HD 129445 along with 100 sample curves from the MCMC. Time is measured in seconds with the first data point recorded at $t = 0$.

The actual values for the parameters are $\mu = 13,800^{+400}_{-300}$ m/s$^{2/3}$, $e = 0.7 \pm 0.1$, $\omega = 2.8^{+0.3}_{-0.2}$ rads, and $\log T = 8.20 \pm 0.01$ [1]. The MCMC accurately predicts the values of μ and T for the system. A significant failure of this MCMC is that it does not predict the eccentricity of the orbit. This is important because $v_r(t)$ is highly dependent on small shifts in e [8] and because different values of e describe very different situations geometrically.

3.2 No Exoplanet

Consider the scenario where an observed system is not binary, it is solely composed of a star with no other large bodies within gravitational range. Mathematically, this corresponds to m, and therefore μ, being equal to 0. Eq. (1) then simply becomes a constant. Setting $\mu = 0$ and randomly choosing the other parameters, 25 radial velocity data points were generated using the procedure outlined in Section 2.3. The MCMC was run with 60 walkers and $40,000$ iterations. The corner plot of the results is shown in Fig. 5 in Appendix A.

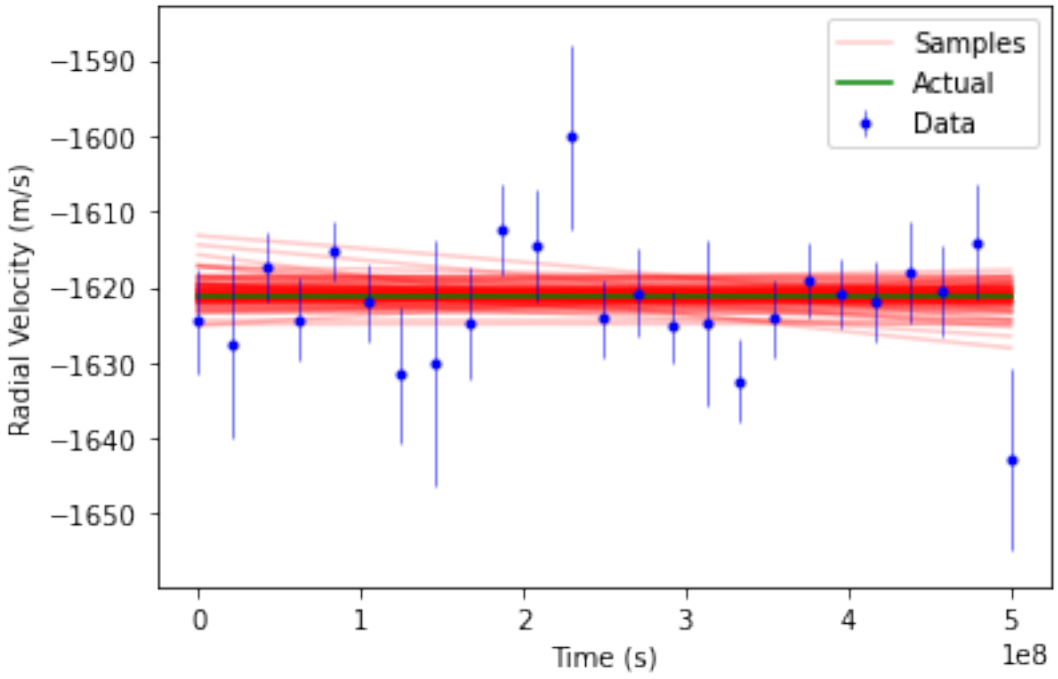

Figure 3: Generated data points along with 100 sample curves from the MCMC. The actual value of $v_r(t)$, $v_0 = -1,621$ m/s, is shown in green. Time is measured in seconds with the first data point recorded at $t = 0$.

The histogram for μ is a Gaussian peaking at 0, which is to be expected by construction. v_0 has a very sharp peak at the actual value used to generate the data. The histograms for the other parameters are uniform along with randfom noise. This is to be expected as when $\mu = 0$ changing e, ω, T, or τ in Eq. (1) does not change the value of $v_r(t)$. This means that the MCMC will not favor any one value of these parameters over another. Note that these results do not seem to apply to T. This because at very large periods $v_r(t)$ is approximately a straight line even when $\mu \neq 0$ over the range time range given. So when evaluating the likelihood there is no reason for the MCMC not to favor this scenario over $\mu = 0$.

3.3 Limitations

Numerically solving Eq. (2b) for u means that limiting cases on e, near 0 or 1, are very sensitive to small changes in other parameters. Extremely elliptical closed orbits are rare [3] so the scenario where $e \approx 1$ is not encountered much outside of the simulated data in this project. On the other hand, $e \approx 0$ is much more common. In this approximation $u \to \frac{2\pi}{T}(t-\tau)$, $f \to u$, and Eq. (1) reduces to a sine wave. However, tangent is not defined over the entire real line so when numerically computing $\tan(u/2)$ there are still regions that are not well-defined.

Just as a change in m could be accounted for by a change in M and vice versa, so can a change in ω be accounted for by v_0. The second cosine term is Eq. (1) is constant with respect to time. Since the scale of the fluctuations of $v_r(t)$ is generally smaller than v_0, it follows that $\kappa \cdot e \cdot \cos\omega$ is also smaller so the distribution for v_0 does not change appreciably. This explains why the MCMCs do not, in general, accurately predict the posterior distribution for ω but still predict that for v_0.

4 Conclusion

Markov chain Monte Carlo algorithms provide a computationally efficient way to explore phase spaces that would otherwise be too large to calculate through direct evaluation. This is particularly useful for determining the parameters of a system given data of that system, as expressed through Bayes' theorem. As an application to astronomy, we can determine the characteristics of the orbit of an exoplanet about a star. Based on our results, we found that we can accurately predict the parameters of a Binary System using radial velocity values, both for simulated data and real binary systems.

However, MCMCs are not without downsides. For parameters with a large number of possibilities on their values, for example T and τ in this project, the algorithm may overrepresent the large numbers when sampling. This happens because the random number selection step in the MCMC chooses from a uniform distribution, but there are more numbers on the order of 10^{13} than 10^5, for example. One way around this is to use information about the system from other sources to place more restrictive constraints on the priors. Furthermore, an MCMC is only as good as the model being used in Eq. (5). As seen in this project, degenerate parameters can make it more difficult to establish reliable posteriors. And models that are difficult to evaluate numerically can impose heavy restrictions both on the computational efficiency of the MCMC and the reliability of results.

References

Arriagada, Pamela et al. "Five Long-period Extrasolar Planets in Eccentric Orbits from the Magellan Planet Search Program". In: 711.2 (Feb. 2010), pp. 1229–1235. DOI: 10.1088/0004-637x/711/2/1229.

Brennan, Pat. Ed. by Kristen Walbolt. URL: https://exoplanets.nasa.gov/ (visited on 04/13/2022).

Butler, R. P. et al. "Catalog of Nearby Exoplanets". In: 646.1 (July 2006), pp. 505–522. DOI: 10.1086/504701. arXiv: astro-ph/0607493 [astro-ph].

Carroll, Bradley W. and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. 2nd ed. Cambridge University Press, 2017.

Landau, L.D. and E.M Lifshitz. *Mechanics*. 3rd ed. Elsevier, 1976.

NASA Exoplanet Archive. URL: https://exoplanetarchive.ipac.caltech.edu/bulk_data_download/#TSD (visited on 03/28/2022).

Sanna, A. et al. "SWIFT J1756.92508: spectral and timing properties of its 2018 outburst". In: *Monthly Notices of the Royal Astronomical Society* 481 (Dec. 2018), pp. 1658–1666. URL: https://doi.org/10.1093/mnras/sty2316.

Sharma, Sanjib. "Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy". In: *Annual Review of Astronomy and Astrophysics* 55 (2017), pp. 213–259. DOI: 10.1146.

Zhang, Zhoujian et al. "The Second Discovery from the COol Companions ON Ultrawide orbiTS (COCONUTS) Program: A Cold Wide-Orbit Exoplanet around a Young Field M Dwarf at 10.9 pc". In: (July 2021).

A Corner Plots

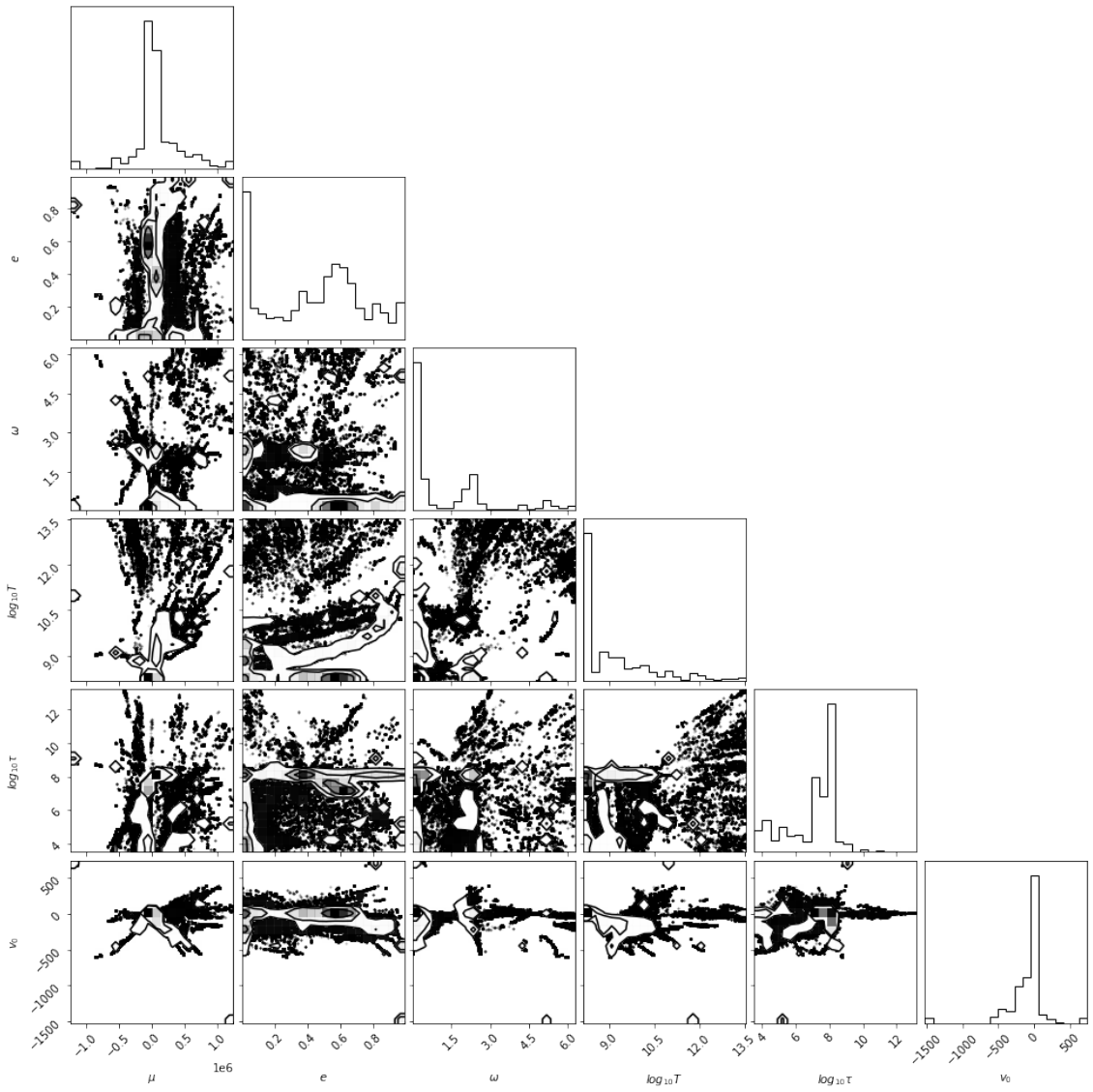

Figure 4: Corner plot for the binary system HD 129445. The actual values for the parameters of this system are discussed in Section 3.1. The first 500 steps for each walker were discarded as burn-in time.

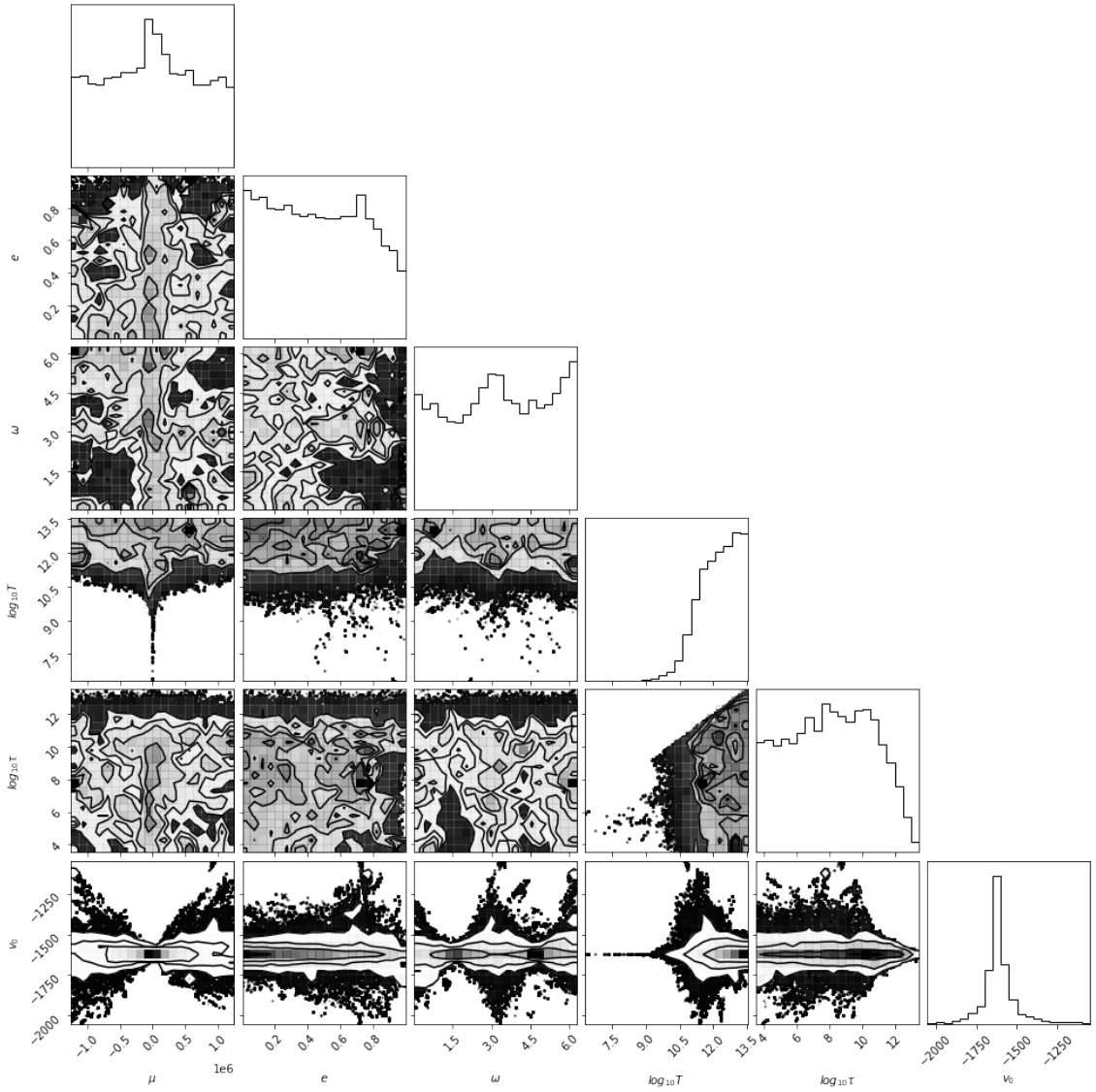

Figure 5: Corner plot for the edge case where $\mu = 0$ described in Section 3.2. The values used to the generate the data are μ=0, $e = 0.6115$, $\omega = 1.257$, $\log T = 12.37$, $\log \tau = 9.962$, $v_0 = -1,621$. The first 500 steps for each walker were discarded as burn-in time.