

# CTA200 Project

Aishani Chaudhuri

## 1 Introduction

The purpose of this project is to investigate binary candidates in the early data release of the DESI multi-epoch survey. This is achieved by applying quality cuts to the DESI survey data, crossmatching with the Gaia eclipsing binary catalog, performing statistical analyses, and using the orbit-fitting program, the Joker, to generate possible orbital solutions for a selected binary candidate.

## 2 Preparing the Data

The quality cuts applied to the DESI multi-epoch data resulted in a catalog of 4,067,737 unique sources, of which 251,664 had at least 3 independent observations, or epochs, and 42,663 had at least 10. This is summarised in Fig. 1 below. The sources that appear in this catalog were sources that:

- have GAIA source IDs
- have stellar spectral types
- have successful RVS reductions with no warnings

- have signal-to-noise  $> 2$
- have velocity uncertainties  $< 20$  km/s

A summary dataframe containing information on the number of epochs acquired for each source, the number of days between the first and last epochs for each source, and the difference between the maximum and minimum measured radial velocities was generated. This information is useful because it is crucial to characterising observed radial velocity variations for a given source. Multiple observations are required to detect any RV variations, and more observations mean more data points to determine RV variations; a longer period of observations with more epochs means periodicity might be better constrained, and the difference between maximum and minimum RV values is a direct measure of the variability of the data. This data is presented in the forms of histograms in Fig. 2 and 3.

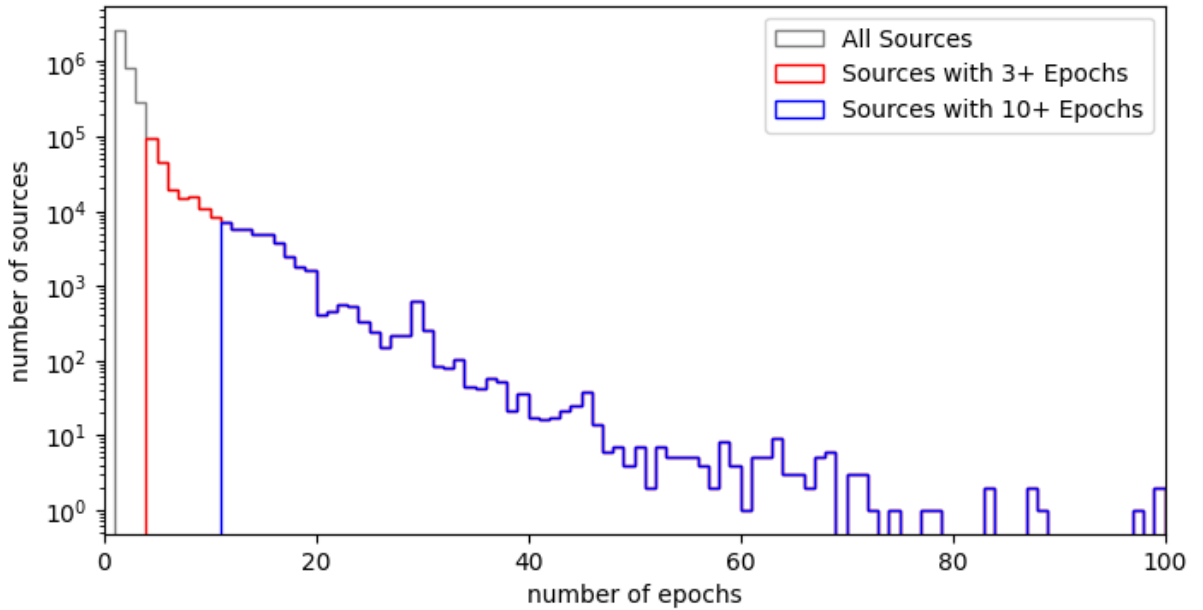


Figure 1: Histogram of the number of sources vs. number of epochs. Sources with at least 3 epochs are shown in blue, and sources with at least 10 epochs are shown in red.

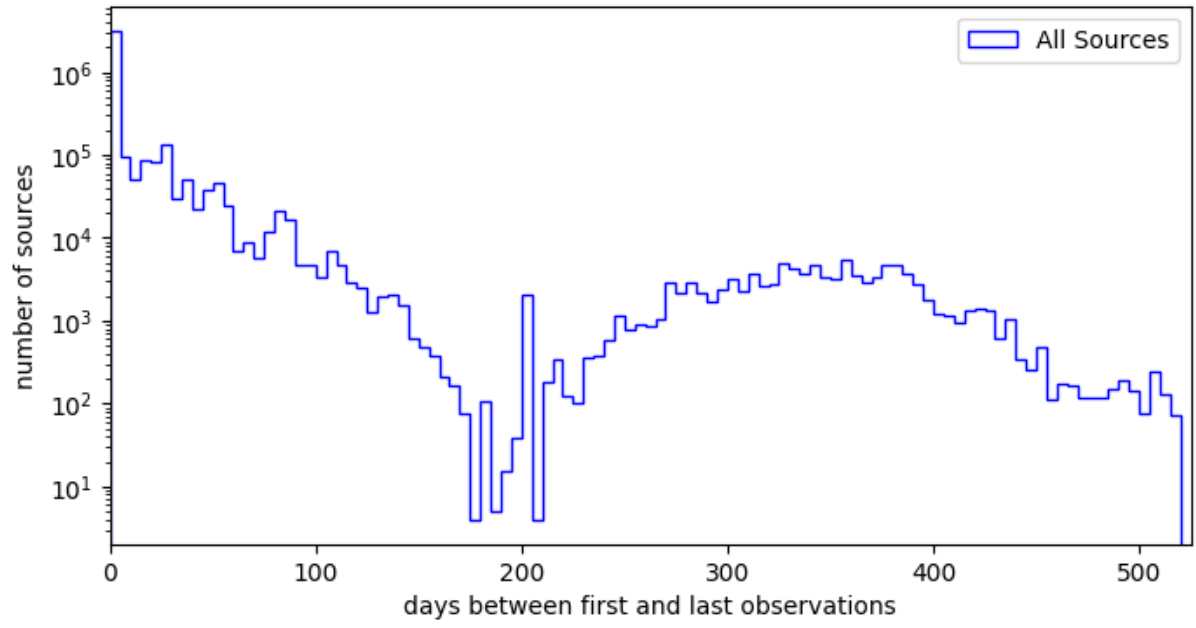


Figure 2: Histogram of the number of sources vs. number of days between first and last observations.

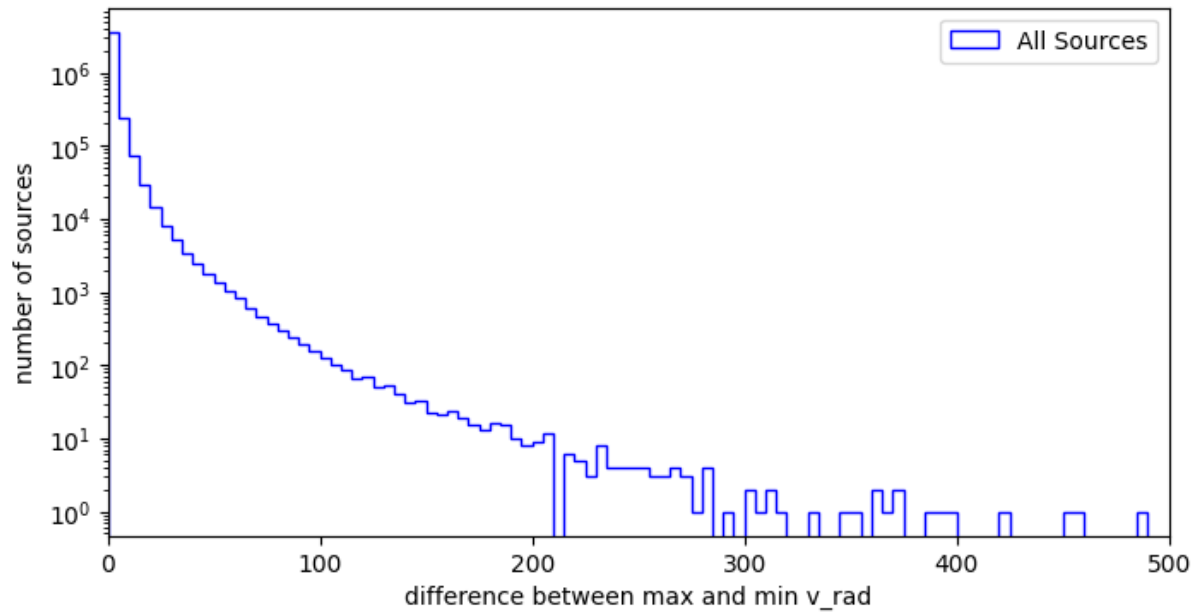


Figure 3: Histogram of the number of sources vs. difference between max. and min. RV measurements.

### 3 Choosing a Binary System to Investigate

Through crossmatching with the GAIA eclipsing binary catalog, it was found that 628 GAIA eclipsing binaries are in the DESI catalog, and 32 have more than 10 epochs. Of these 32 sources, source with GAIA ID 1434118068653836928 was chosen for orbit-fitting. The RV data for this source is shown in Fig. 4.

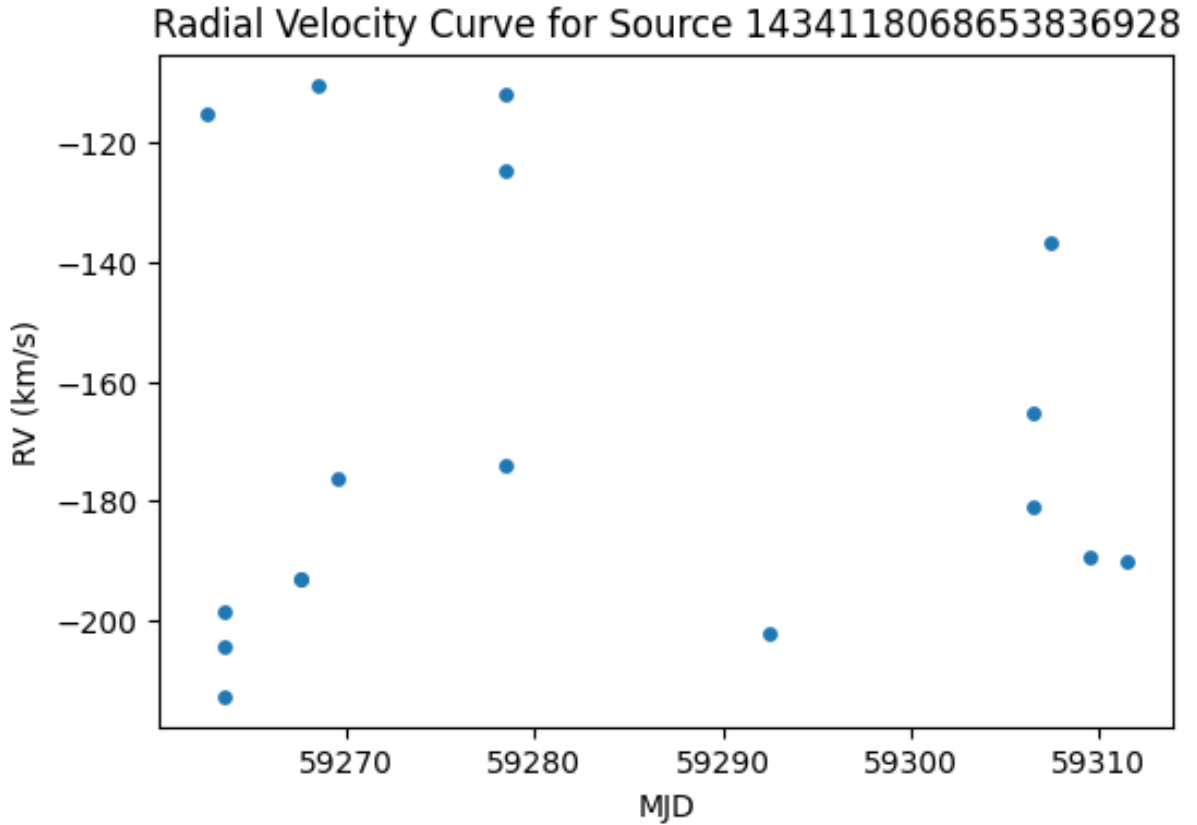


Figure 4: DESI radial velocity curve for source 1434118068653836928.

Statistical analyses were run on the data for the source to detect RV variability, and the following values were obtained:

- chi-squared value: 48.61
- p-value:  $3.81 \times 10^{-5}$

- F2 statistic: 3.92
- weighted mean velocity:  $-167.71$  km/s

The weighted mean velocity is an estimate of the true velocity of the source, to compare other measurements to so chi-squared values, p-values and F2 statistics can be calculated. The chi-squared value represents the degree to which measurements differ from the mean value when, as compared to the error of the measurements; higher values indicate higher probability that there is radial variability. The p-value is a measurement of how likely it is for the measurements to have the values/distribution they do given the null hypothesis (that there is no radial variability); lower values indicate higher probability that there is radial variability. The F2 statistic has a similar meaning to the chi-squared value, modified to be independent of the degrees of freedom so it can be generalised to a population; it is calculated using Eq. 1 shown below.

$$F2 = \sqrt{\frac{9\nu}{2}} \left[ \left( \frac{\chi^2}{\nu} \right)^{1/3} + \frac{2}{9\nu} - 1 \right]. \quad (1)$$

It has been shown that DESI radial velocity measurements have an error floor of  $1$  km/s. Adding this in quadrature to the error values reported in the DESI catalog, the calculations of the values above were repeated:

- chi-squared value: 44.62
- p-value:  $1.54 \times 10^{-4}$
- F2 statistic: 22.41
- weighted mean velocity:  $-168.04$  km/s

Since the p-value of this source is less than 0.01 and the F2 statistic is greater than 3.1, source 1434118068653836928 shows strong evidence of radial variability, and is an ideal choice for orbit fitting.

## 4 Fitting an Orbit with the Joker

The Joker is an orbit-fitting package designed for binary candidates. When applied to source 1434118068653836928, it returned one unimodal orbit, indicating that the RV data constrained the orbital parameters very well. The fitted orbit is shown plotted over the DESI RV curve (black dots) in Fig. 5.

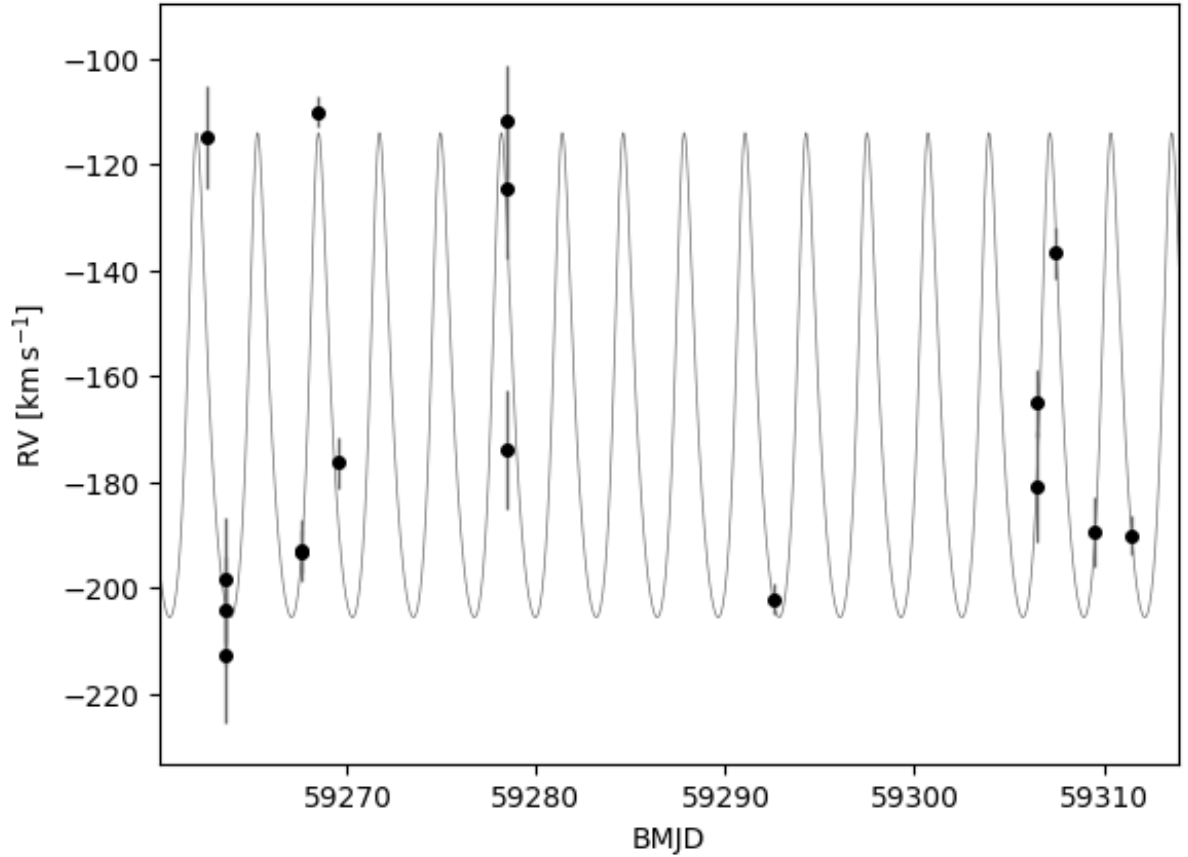


Figure 5: Orbit generated by the Joker for source 1434118068653836928.

Further, the orbit was also fit using No U-Turn Sampling (NUTS), which is a Hamiltonian Monte Carlo Method. The posterior distribution for this is depicted in Fig. 6 as orbits for the source, and in Fig. 7 as a corner plot of orbital parameters.

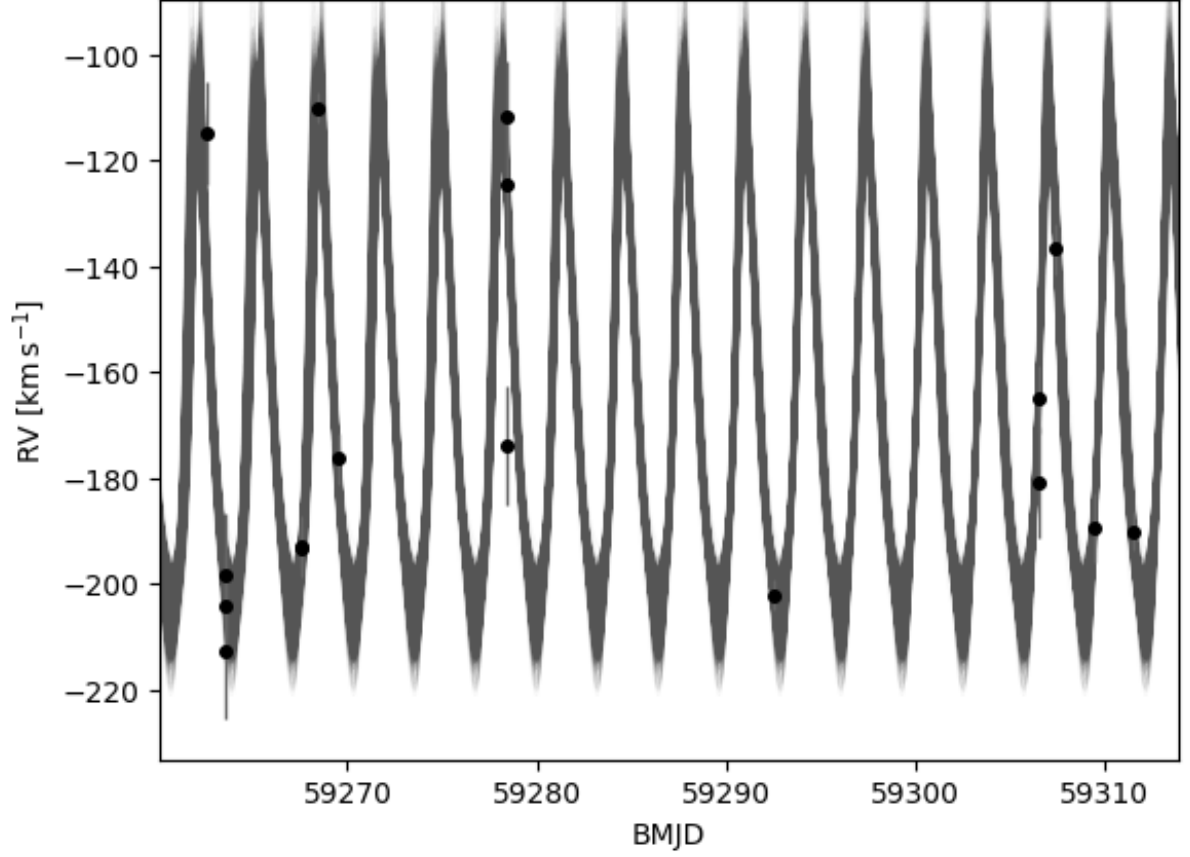


Figure 6: Posterior distribution of orbits generated by the NUTS method using the Joker for source 1434118068653836928.

As can be seen from these plots, the fitted orbits for this source are strongly constrained by the data.

The phase-folded RV curve and residual plot for the Joker orbit are also presented below in Fig. 8. The phase here refers to the point in a period, or the fraction of the orbit traversed. It can be seen from the residual plot that, though the orbit is a good fit for the data, there are a few potential outliers in the measurements.

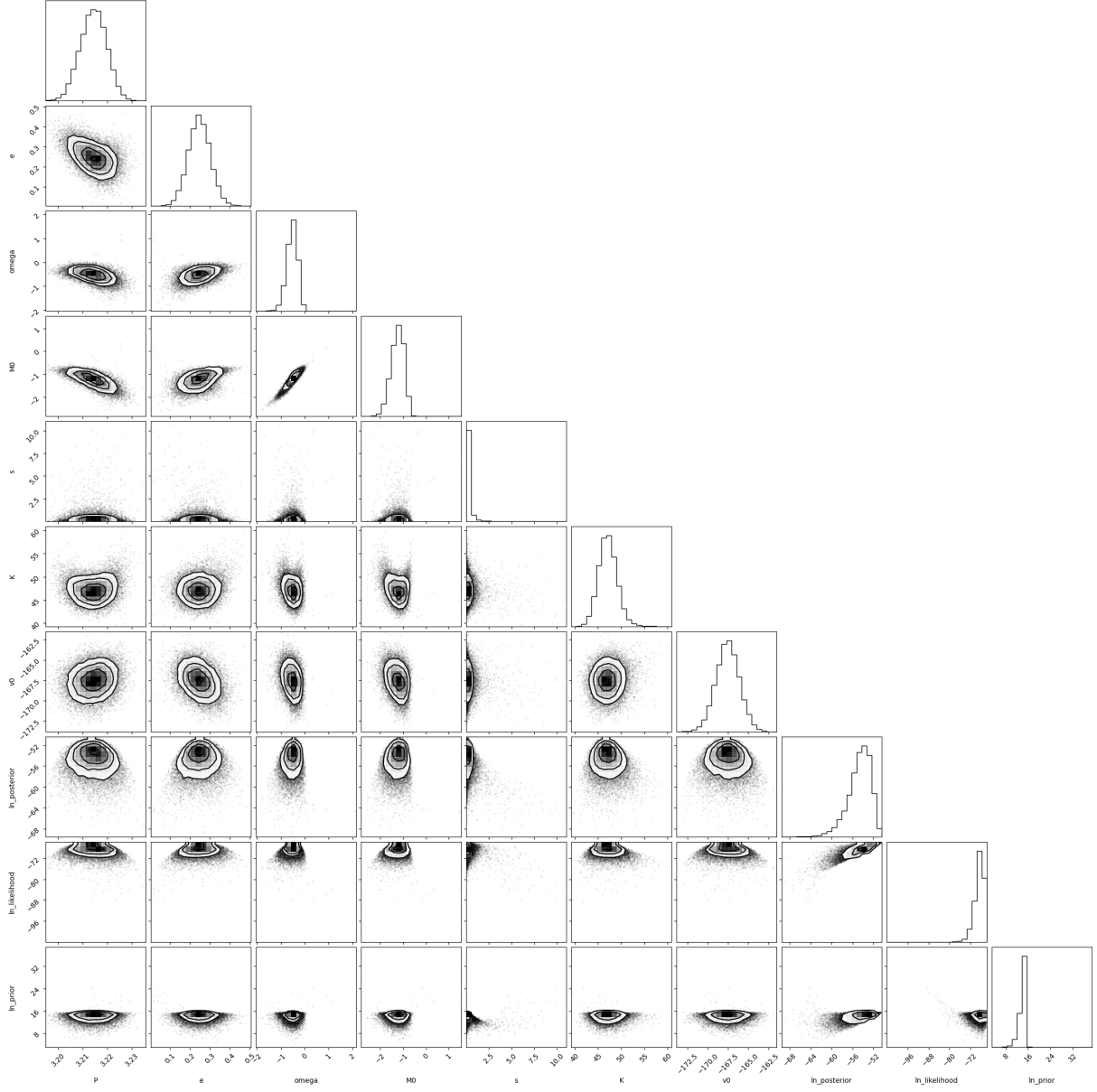


Figure 7: Corner plot of the posterior distribution for the orbital parameters for source 1434118068653836928.



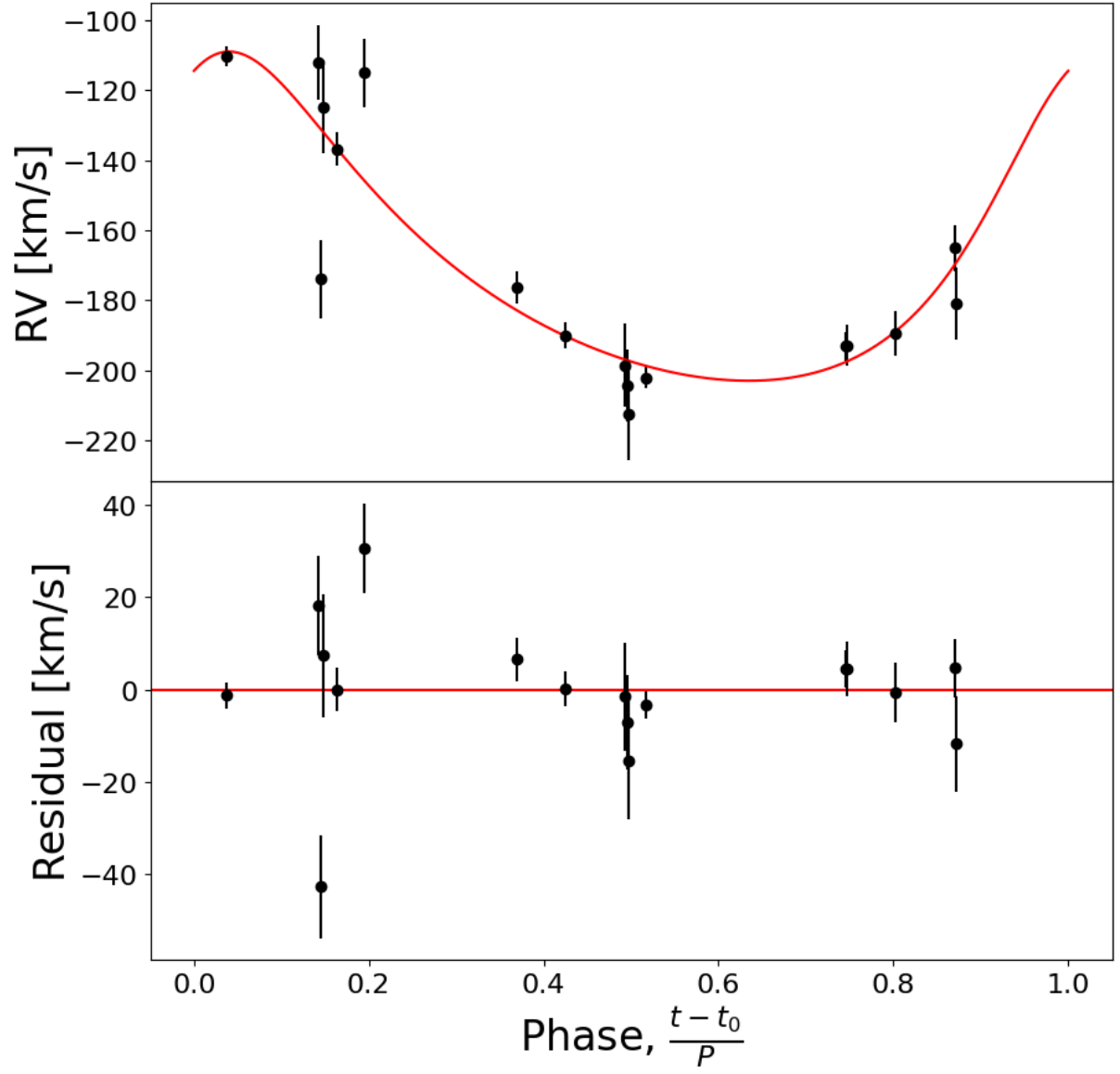


Figure 8: Phase-folded RV curve and residual plot for the Joker orbit for source 1434118068653836928.

## 5 Validating Results

The orbital frequency of source 1434118068653836928 is reported in the GAIA eclipsing binary catalog, and the period calculated as its inverse is 0.397 days. However, the period of the orbit generated using the Joker is 3.22 days, which is a little under 10 times the GAIA period. It is possible that the obtained period using DESI data is much longer than the GAIA period due to too long intervals between epochs, resulting in something akin to aliasing, where measurements are not frequent enough to detect all features of an oscillation accurately. The potential outliers noted in the residual plot in Fig. 8 may also have contributed to this discrepancy.