Classifying Variable Stars with Gaia Color-Magnitude Diagrams

Xander Hall1, Adam A. Miller2,3, Nicholas Easton4, and Aaron M. Geller2,3

I. The Need For Classification

As astronomical data sets grow beyond what can be explored via visual examination, there is a paramount need for automated classification techniques. This is especially true for the Large Synoptic Survey Telescope (LSST), an 8.4 m telescope in Chile, that will repeatedly observe ~37 billion stars and galaxies. Of these, more than 20 million stars will exhibit significant brightness variations over the project’s 10 year duration. Efficient algorithms are needed to classify these stars to better understand their nature and the formation history of the Milky Way galaxy.

II. Classifying Variable Stars

To improve the automated classification of variable stars, we have started a Zooniverse project, Stellar Sleuths. The Zooniverse is a public portal to citizen science. Zooniverse volunteers can inspect images produced by researchers in a variety of fields and provide feedback on the quality or content in those images. We will use classifications from the Zooniverse to train machine learning algorithms to classify variable stars. Currently, Stellar Sleuths only provides light curves to Zooniverse users (see III). However, there is additional information (such as the temperature of a star) that can be used to aid the classification process. In this project we have experimented with the addition of supplemental information to add to the light curves for classification. We find that color-magnitude diagrams (see V) are useful for classifying variable stars.

III. Light Curves

Traditionally, astronomers use light curves, plots of the brightness of a star as a function of time, to classify variable stars. Light curves can be used to determine if a star shows periodic modulations (see the 3 plots below for examples). Stellar Sleuths provides light curves from the Asteroid Terrestrial-impact Last Alert System (ATLAS) to citizen scientists for classification. ATLAS has found ~4.7 million variable stars, which we use to test our classification methods in advance of LSST.

Can you tell the difference between the 3 light curves below?

IV. Distances

For a majority of stars within the second data release of Gaia, reliable distances cannot be attained by simply taking the inverse of the parallax. A correct inference procedure must instead be used to account for the nonlinearity of the transformation and the asymmetry of the resulting probability distribution. By using an unnormalized posterior provided by Bailer-Jones (2018) we are able to find more accurate estimates for distance, which allows us to more accurately plot stars on the CMD.

V. Color-magnitude Diagram

Color-magnitude diagrams (CMDs) can be used to isolate stars in different stages of stellar evolution. With Gaia, we can place hundreds of millions of stars on the CMD to classify their nature. Gaia is a space-based telescope with an unprecedented ability to measure precise distances to hundreds of millions of stars. Prior to April 2018, this information was completely unknown. Gaia distance measurements allow us to determine the intrinsic brightness of the stars that it observes.

VI. CMDs and Classification

We find that CMDs significantly improve our ability to classify different variable stars. For example, RR Lyrae stars and Cepheid variables have very similar light curves (see the 1st and 3rd light curves, above), yet their positions in the CMD are completely different.