Tracking Stellar Streams Using FAISS

Alec Hewitt[1], Xiangyang Ju[2], Maurice Garcia-Sciveres[2]
[1]University of Utah, [2]Lawrence Berkeley National Laboratory

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY
Office of Science

Abstract

A stellar stream was once a globular cluster or a dwarf galaxy that has been distorted and stretched out due to the tidal forces of a larger galaxy. In particle physics it is useful to know how to trace the trajectories of these interactions. In this project we apply methods of particle physics to locate and track the trajectories of stellar streams. The methods used in this project utilize the machine-learning package “FAISS” (Facebook AI Similarity Search). The algorithm scans the galactic plane for significant clusters, scanning one small dataset at a time. Once these datasets are identified, the algorithm tracks these clusters to determine whether they could be apart of a stellar stream. We found that this can be a promising method for locating and storing information for different stellar streams throughout our galaxy. Information of stellar streams could be helpful for providing evidence on our galaxy’s past as well as information on our galaxy’s gravitational field.

Introduction

- Stellar streams are structures that have been stretched out and distorted by our galaxy’s tidal forces
- These structures generally started as dwarf galaxies or globular clusters
- Stellar streams contain valuable information about our galaxies gravitational pull
- Our galaxy is dominated by dark matter at large radii, so streams could help physicists put constraints on dark matter theories
- This project develops an algorithm to detect and track stellar streams
- These streams follow trajectories much like a particle
- We use techniques similar to techniques used to follow particle trajectories
- We use Gaia EDR3, the largest database

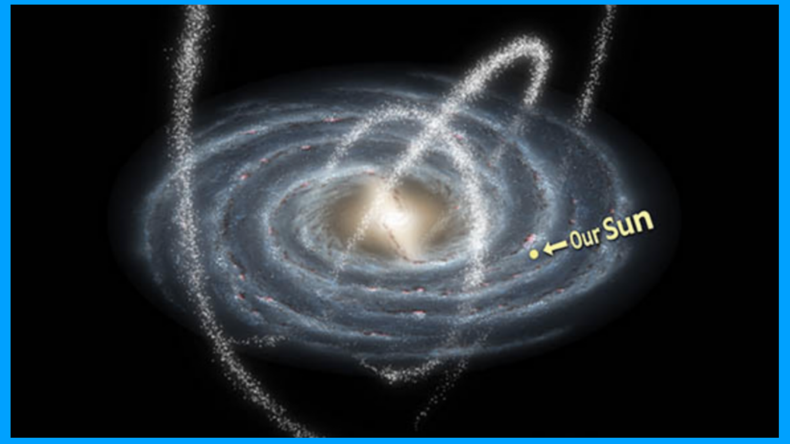


Figure 1: Stellar streams in the Milky Way
Credit: NASA/JPL-Caltech/R. Hurt (SSC/Caltech)

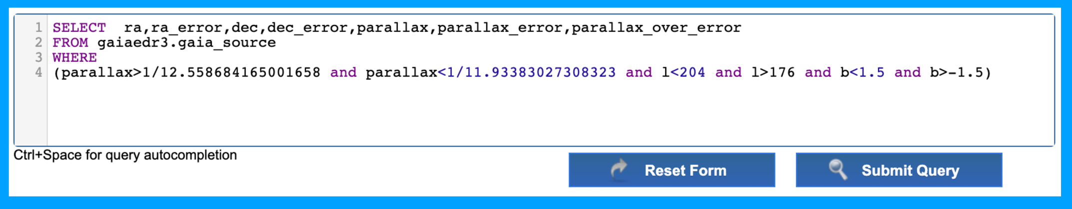
Research Question

Can we develop an algorithm that can locate and follow stellar streams?

Methods/Results

Obtaining the Datasets

- Datasets were obtained from the Gaia archive using ADQL
- More specifically, we used Gaia EDR3 (early data-release 3)



```
1 SELECT  ra,ra_error,dec,dec_error,parallax,parallax_error,parallax_over_error
2 FROM gaiaedr3.gaia_source
3 WHERE
4 (parallax>1/12.558684165001658 and parallax<1/11.93383027308323 and l<204 and l>176 and b<1.5 and b>-1.5)
```

Ctrl+Space for query autocompletion

Reset Form | Submit Query

Figure 2: Example of ADQL search query; allows searching and retrieval of data

Identifying Clusters

- Stellar streams do not form tight clusters in space
- Stellar. Streams form tight clusters in proper motion space (feature space)
- We use FAISS (Facebook AI Similarity Search to obtain clusters
- A plot of number of stars vs. cluster determines whether there is a cluster or not

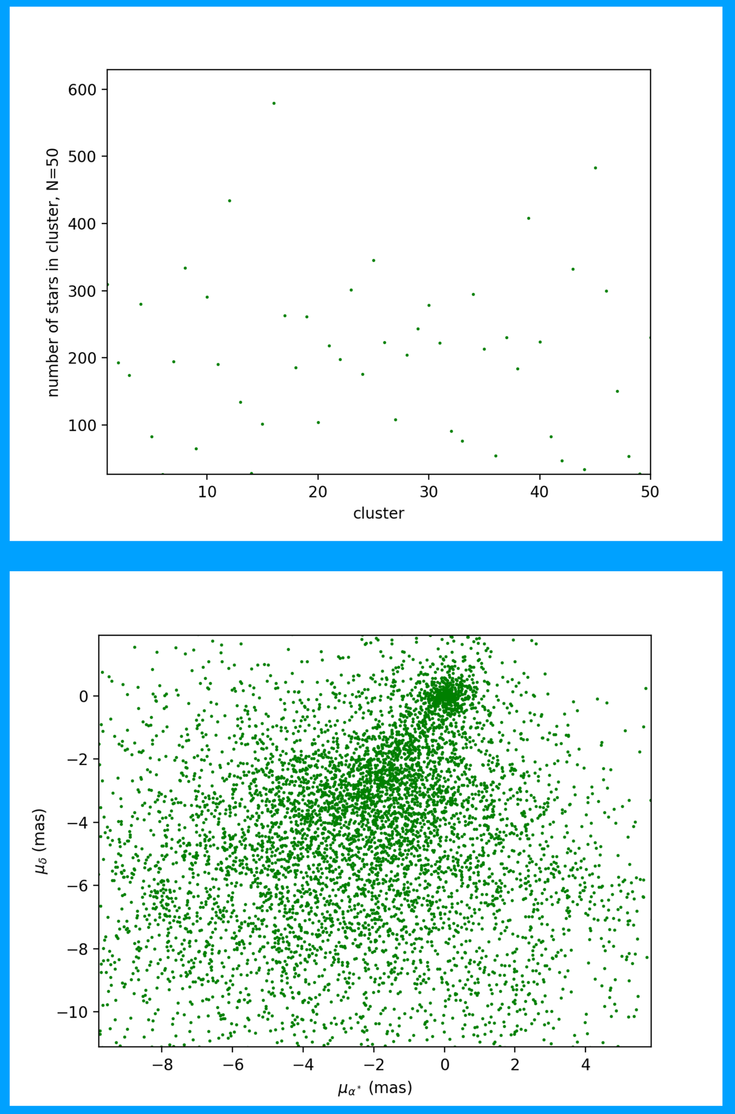


Figure 3: top image: plot of number of stars vs cluster spike signifies significant cluster. Bottom: plot of proper motion, showing a clear cluster

Automation

- Start with a disk-like lattice
- Each one represents the center of a dataset

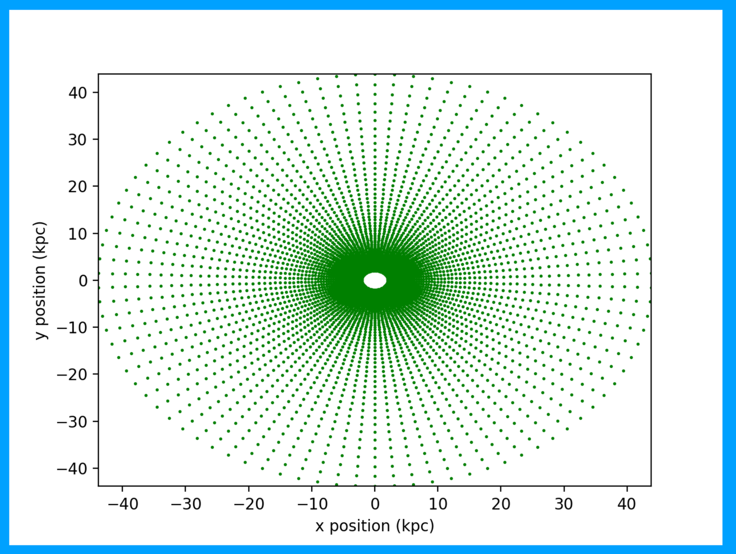


Figure 4: A lattice of points in galactic plane, each represents a dataset

- The algorithm randomly chooses a point and removes it, accesses the database and extracts the corresponding dataset to check for stellar streams

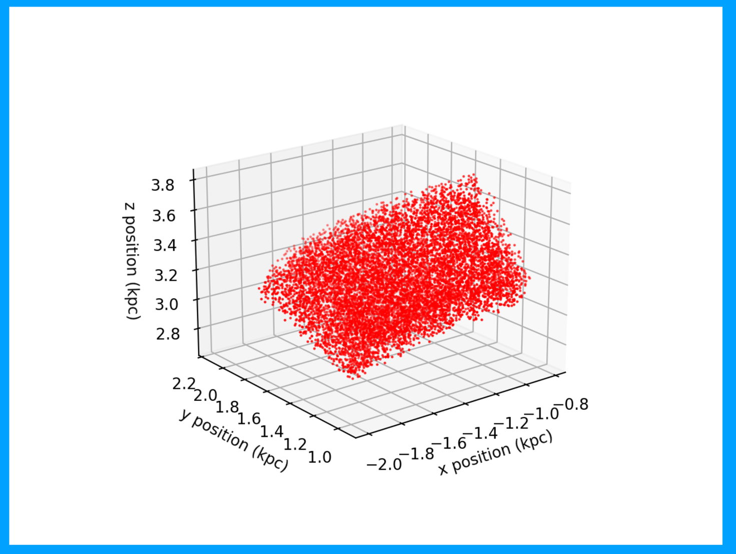


Figure 5: a typical lattice point after being transformed into a dataset

Connecting Clusters

- If a dataset contains a significant cluster, the algorithm checks all neighboring datasets for significant clusters
- This process repeats until no more clusters are found

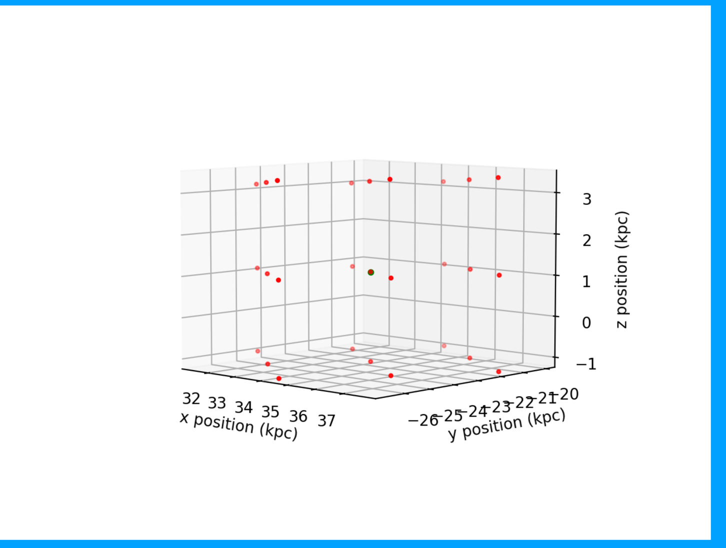


Figure 6: Once a stellar stream is found, the algorithm checks all neighboring sets (in the spherical lattice) for streams

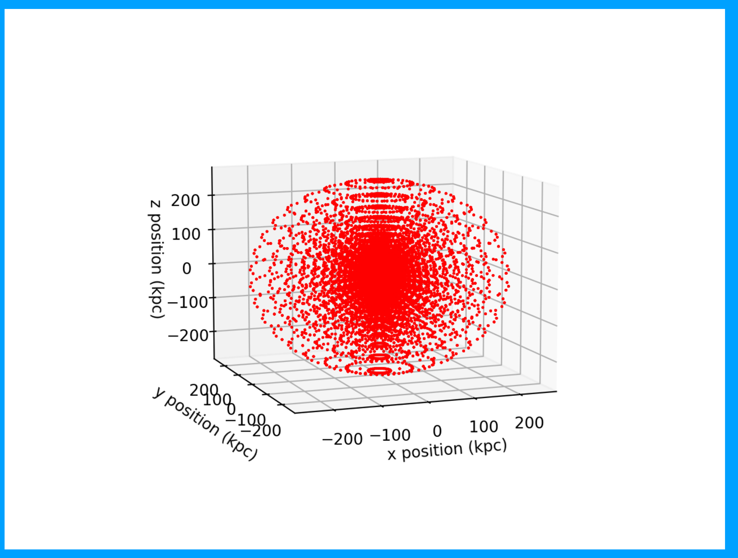


Figure 7: A spherical Lattice

Discussion

- This algorithm can successfully locate and track a stellar stream
- Algorithm needs much more testing
- Definition of stellar stream needs to be refined
- Algorithm is slow but can be drastically improved with multiprocessing
- Data extraction needs modified to ensure datasets are equal sizes
- Doesn’t work well near the sun

Conclusion

- FAISS provides a quick tool for finding clusters
- This algorithm scans galactic plane one dataset at a time for stellar streams
- If one is found it can follow it out of galactic plane
- Works especially well with multiprocessing
- Still needs work to reach full potential

Acknowledgements

This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program.. I would like to thank Maurice Garcia-Sciveres and Xiangyang Ju for useful advice in guiding this project

Poster Summary

Alec Hewitt
SULI Spring 2021

Our group designed an algorithm that can identify and follow stellar streams using Gaia EDR3. Stellar streams are a group of stars (galaxy or cluster) that have been stretched and strung out due to the tidal forces of our galaxy; many of which follow complex trajectories around and within our galaxy. Streams may contain valuable information pertaining to the gravitational field and the stars contained within so it is helpful to develop an algorithm to extract information about these structures. This algorithm starts with a lattice of points along the galactic plane. The algorithm will randomly choose and remove a point from this lattice and turn it into a dataset containing all stars near this point. The algorithm then determines whether that dataset has a stellar stream, if it does, it follows or tracks the stream, if it does not then it removes that dataset from consideration. The end result is hoped to track large portions of stellar streams. Certain properties of these stellar streams could help us nail down which dark matter theory is correct. For example, theories of dark matter predict that there are clumps of dark matter in our galaxy and depending on the theory, the size and number of these clumps vary. These clumps can collide with stellar streams and depending on the frequency and size of resulting break in the stellar stream could agree with certain theories more than others.