Accelerating CHIME/FRB Baseband Localization

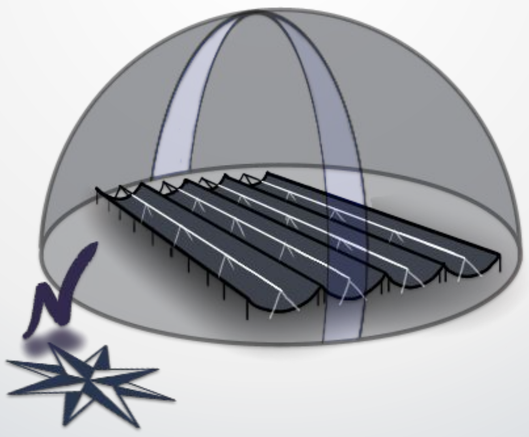
Diana Korotun, Keith Vanderlinde
Dunlap Institute for Astronomy & Astrophysics, University of Toronto

Abstract

My research focuses on enhancing the efficiency of the localization stage within the baseband analysis pipeline of the CHIME telescope. The primary objective is to reduce runtime, which is achieved by optimizing the MCMC (Markov Chain Monte Carlo) segment of the pipeline. Through the implementation of parallelized code using the Multiprocessing package, the runtime of the MCMC algorithm was notably reduced from 41.09 minutes to just 7.139 minutes when running with 8 cores. This marks a 82.63% reduction in the execution time of the localization stage while maintaining data integrity.

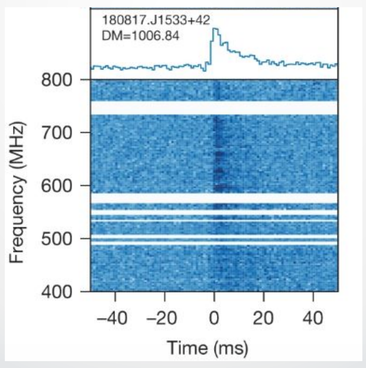
Introduction
CHIME Telescope:

The CHIME telescope in British Columbia started in September 2017. It consists of four 20x100m parabolic cylinders which detects radio signals. Its design creates daily 400-800MHz signal maps by scanning the sky utilizing the Earth's rotation, consequently, resulting in the observation of numerous FRBs, surpassing earlier efforts.[1]



Introduction - FRBs:

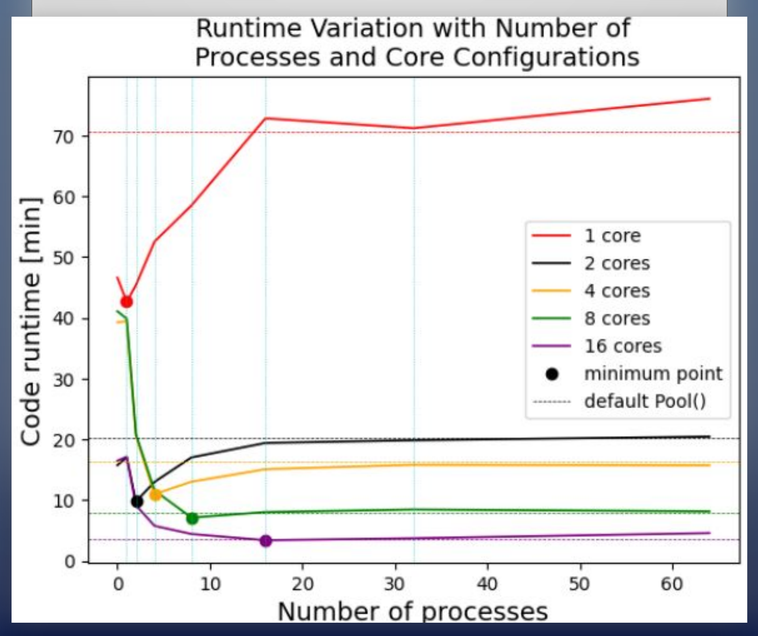
FRBs are short and intense radio signals, with uncertain origins. CHIME's unique design allows for an effective FRB detection despite their sporadic nature.[2] [1]



Results:

After implementing parallelization in July 2023, I found that optimal results occurred when the process count matched the number of cores. Balancing this is critical since too few processes increase core workload, while excessive processes can lead to code stalls due to cumulative time consumption.

#cores	1	2	4	8	16
#Processes	1	2	4	8	16
Original Time [min]	46.63	15.76	39.28	41.09	16.50
MCMC Time [min]	42.66	9.791	10.99	7.139	3.419



Introduction baseband analysis:

The process of CHIME detecting signals and analyzing them offline is called Baseband analysis. It consists of:

- Detecting radio signals and Identifying strong and sudden signals
- Mapping their strength on the sky and fitting with a 2d Gaussian function for approximating location.
- Once the location is found, maximize telescope sensitivity in that direction.
- providing data for further related research

Initially, this process took hours, with MCMC alone needing an hour. This highlighted code inefficiencies and caused delayed data analysis. Thus, I my work was to speed up the pipeline.

Method:

To speed up processing, I added parallelization to the MCMC algorithm. This meant I made the data to process concurrently instead of element by element sequentially.

CHIME's pipeline uses the MCMC function from the "emcee" package and its EnsembleSampler() method. This method takes the "pool" parameter, using the Pool function from the "multiprocessing" package. Where Pool is a function that enables simultaneous task execution across processes, This setup allows for a customizable number of processes during multiprocessing. In my research, I tested different combinations of "number of processes" and "number of cores" for the best outcomes.[3]

Discussion

With, this refinement of the pipeline, the baseband analysis is able to perform faster, thus facilitating more efficient program testing to produce better quality data for further research with FRB, in related fields such as polarization analysis, signal morphology , and many more.

Moreover, faster running pipeline, means faster detection of new FRBs. Which could allow for potential collaboration with other radio telescopes observing the sky, and the opportunity to follow up on the detected interesting events from new perspectives.

This in term would enable a better understanding of FRB and, maybe, their exact cause and origin.

Bibliography

[1] CHIME Experiment. (n.d.). Home. CHIME Experiment. https://chime-experiment.ca/en

[2] Foreman-Mackey, D. (n.d.). Parallelization — emcee v3.1.4 documentation. emcee. https://emcee.readthedocs.io/en/v3.1.4/tutorials/parallel/

[3] Petroff, E., Hessels, J. W. T., & Lorimer, D. R. (2021). Fast radio bursts at the dawn of the 2020s. arXiv preprint arXiv:2107.10113.

I would like to express my gratitude to the Dunlap Institute for giving me the opportunity to participate in the SURP program. Additionally, I extend my thanks to Keith Vanderlinde, Daniele Michilli, Kaitlyn Shin, Shiny Brar, and the CHIME collaboration for their guidance and assistance throughout my journey so far.