Milestone 3 Report

Applicant:
Gayatri Aniruddha
Student ID: 20875110

Supervisor:
Dr. Marcin Sokolowski

Curtin Institute for Radio Astronomy
Masters of Philosophy (MPhil)

Thesis: High-Time Resolution GPU Imager for Low-Frequency Radio Telescopes

October, 2023

International Centre for
Radio Astronomy Research

Curtin University

1 Research overview

Fast Radio Bursts (FRBs) are very bright milli-second radio pulses from a distant Universe. Since the first FRB was discovered in 2007, this field has rapidly advanced to a state where real-time detections are routinely made by multiple telescopes worldwide. The initial detections were made at 1-2 GHz, and despite extensive searches over the past several years, only a few FRBs have been detected below 350 MHz. Typically, FRB searches are performed using beamformed data. However, this method applied to wide-field data from low-frequency radio telescopes can be computationally expensive. Therefore, image-based searches may be a more efficient alternative. Though existing imagers such as CASA, MIRIAD and WSCLEAN are suitable for FRB searches, they involve an additional cost of converting the data (hence I/O operations) as they require input data in specific formats. Hence, in my Master's Research, I have developed a new imager, executed nearly entirely on Graphical Processing Units (GPUs), which will become a part of a larger processing pipeline for FRB searches. The primary target instruments for the imager are the low-frequency telescopes in Western Australia, the Engineering Development Array 2 (EDA2) and the Murchison Widefield Array (MWA).

2 Work to date

I completed the following tasks until Milestone 2:

1. **Understanding of interferometric imaging and development of test Python code.**
   - To understand every stage of interferometric imaging, I implemented a standalone Python version of an imager from scratch. The imager was based on the equations in radio astronomy textbooks and applied to visibility data from the Engineering Development Array 2 (EDA2) from a single time step and frequency channel.
   - I validated this Python imager by comparing the images from EDA2 data with those produced with the existing C++ imager (implemented by my supervisor).
   - The Python code can be found here: https://github.com/gayatri-projects97/Python-Imager.
2. **Development of C++ version of the imager for GPUs:**
   - Based on the Python (my implementation) and C++ versions (my supervisor's implementation) of the imager, I developed a GPU version for a single time step and frequency channel.
   - I developed the following two test versions for GPUs:
     - Single time-stamp, single-channel GPU imager (using function cufftPlan2D() from cuFFT library)

    - Single time-stamp, multi-channel GPU imager
      (using function cufftPlanMany() from cuFFT library)
  - The main difference between the CPU and GPU versions was that the GPU code has gridding implemented as a CUDA kernel (i.e., it is executed in parallel on multiple CUDA cores) and used the cuFFT library instead of the fftw library.
  - I tested and verified the GPU imager by comparing the resulting sky image with those produced by the CPU imager. The images produced by the CPU imager were already validated by my supervisor with those produced by the standard radio astronomy software such as CASA, MIRIAD or WSCLEAN on the same real and simulated data and using compatible options.
  - I benchmarked and compared the performance of GPU and CPU versions of the imager on Topaz and Setonix Supercomputers at PAWSEY.
  - I also benchmarked and compared the performance of different versions of parallel gridding implementations (using CUDA streams and 2D-grid of CUDA blocks) on Topaz and Setonix supercomputers.

3. **Written work:**
   - I started planning and writing sections of my proposed paper and completed writing 4200+ words. A paper draft and the Turnitin report were submitted during Milestone 2.
   - I also started planning the sections of my thesis.

Since Milestone 2, I have completed the following tasks:

1. **Understanding of brute-force de-dispersion algorithm:** To understand brute force de-dispersion, I reviewed an existing code written by my supervisor's Honours student (Jia-Ming Toh). I tested and implemented my own version based on this code in Python and also tested it on simulated test FRB data.

2. **Application of the imager to EDA2 data:**
   Data given to me:
   - My supervisor recorded EDA2 visibilities in 100 ms time resolution and bandwidth of about 0.94 MHz fine channelised into 32 frequency channels (centre frequency of 230 MHz). These visibilities were imaged into 180x180 all-sky images.
   - Each observation was 1 minute long, and the observations were recorded on:
     - 6 June 2023: 38 observations ($\approx$ 38 minutes duration)
     - 9 July 2023: 108 observations ($\approx$ 108 minutes duration)

- Earlier developed software (described in Anderson et al., 2021) was used to create dynamic spectra for every direction in the sky and then de-disperse in DM range 0 to 900 pc cm^{-3} with DM steps of 150 pc cm^{-3}. My supervisor then provided me with two consolidated .txt files (one for each 2023-06-01 and 2023-07-09 data set) to analyse and search for potential FRB candidates. These .txt files contained data of those candidates from the de-dispersed time series (3D image in FITS file format with the X-axis being pulse arrival time, Y-axis trial DM and Z-axis mean flux density along the dispersion sweep) of both the June & July data, with S/N > 5.

Data processing done by me:

- I read these .txt files into a pandas DataFrame, and all further data analysis was done in Python.
- I filtered out candidates that had S/N > 10 to inspect them. The summary of the final candidates, obtained from June and July data, can be found in Table 1.

Table 1: The final candidates obtained from June and July data, including the number of near-horizon candidates and RFI. Some images had multiple (20+) candidates, which indicated that something could be wrong at that time step as RFI could have affected these images. These rubbish/false positive candidates have been marked as RFI and excluded from this analysis. This left us with 8 candidates from June and 5 from July data for further inspection.

Data	Number of candidates with S/N > 10	RFI	Remaining candidates	Number of candidates near-horizon	Potential candidates
2023-06-01	272	259	13	5	8
2023-07-09	1829	1821	8	3	5

- For each potential candidate (i.e. 8 from the June data, 5 from the July data), I first inspected images at the corresponding time steps to see if these candidates were visible (or faintly visible). If the candidates were visible, I further checked to see if they were on the horizon, which meant they were due to RFI from one of the towns surrounding the MRO. For the non-horizon candidates, I checked if their RA and DEC matched those of known pulsars from the Australian Telescope National Facility (ATNF)[1] pulsar catalogue. For the pulsar candidates, I performed the following two checks:
  - We know the bandwidth we used to record this data (i.e. 1 MHz). Based on the dynamic spectrum, we could see that the maximum dispersive delay observed was less than 100 ms. So, the maximum DM any of these candidates could have should not exceed 150 pc cm^{-3}. This check eliminated pulsars above this DM limit.

[1] https://www.atnf.csiro.au/research/pulsar/psrcat/

    - In these images, the sensitivity (sigma) was around 75 Jy. So, to get detected in a 10 σ detection, a pulsar should have a flux density of at least 750 Jy. All the pulsar candidates we got had their flux densities of a few mJy at GHz frequencies. So, after extrapolating these values to get their flux densities at 230 MHz (using the mean spectral index value of -1.6), they were of the order of milli Jansky. This essentially eliminated the possibility of any of these candidates being pulsars.

  - I also checked whether the non-horizon candidates' RA and DEC matched the positions of known satellites at the times of the observations of known nearby satellites. I looked for satellites that could present in a 5-degree and 3-degree radius. For the satellite candidates, I performed the following checks:
    - I eliminated any satellites that were either debris or rocket-bodies.
    - I checked whether or not the suspected satellite candidates were active using this website: http://n2yo.com/.
    - I checked if the down-link frequencies of any of these suspected satellites were around 230 MHz, mainly using the following website: https://db.satnogs.org/transmitters.

    There were no satellites that generally transmitted signals in this downlink frequency range. I am currently working with a PhD student from the Space Surveillance group (Dylan) to further see if any of these signals were possibly from satellites, as these could be unintentional transmissions from these satellites.

  - In my M3 talk, I will present these results and analyses.

3. **Thesis:** I have been writing my thesis, and currently, I have completed around 75 pages. Here's a break-down of the chapters and my current progress:

  - **Chapter 1: Introduction and Background**
    Current state: **Completed**, iterations underway.
  - **Chapter 2: Methodology**
    Current state: **Completed**, iterations underway.
  - **Chapter 3: GPU Imager for low-frequency radio telescopes**
    Current state: **Completed**, iterations underway.
  - **Chapter 3: Pilot FRB searches on ED2A data**
    Current state: **On-going**, iterations underway.
  - **Chapter 4: Summary**
    Current state: **Not started**.
  - **Appendix**
    Current state: **Not started**.

4. Conferences:

- International High-Performance Computing Summer School and Conference (IPHCSS)[2] in Atlanta from 9-14 July 2023. Here, I got the opportunity to present the GPU imager as an electronic poster presentation. My poster presentation can be found here: Poster-Presentation.
- MWA project meeting, 25-28 July 2023.
- PaCER Conference (PCon), 21-22 September 2023.

3 Future and completion plan

I will focus on the following tasks after my Milestone 3:

1. Thesis:
   - I will continue working on my thesis and polish up the currently written chapters when I get further feedback/iterations from my supervisors.
   - I plan to submit my thesis on 8 December, and I plan to have around 80 - 90 pages written by then.
2. Conferences:
   - I plan to attend the Australian Institute of Physics (AIP) Student Conference on 17 November 2023 and give a talk (if my abstract gets accepted).
   - I also plan to attend and give a talk on ICRAR Student Day, which will most likely happen in the first week of December.
3. Paper: The remaining sections of the paper corresponding to imager validation and demonstration of real, simulated data need to be drafted and updated. After Milestone 2, I mainly worked on my thesis and did not update my paper much.

[2] https://ss23.ihpcss.org/