I am a 19 year old student in my 4th year of undergraduate studies in CS and Math at the University of Toronto. About 3 months ago, I took a break from university to work full-time at NVIDIA as a research intern. While I have not had any published previous work in astronomy, I am very interested in the idea of learning to do some kind of pattern matching in FRBs or photometric data.

My academic interests mainly lie in research around learning algorithms and vision. To this extent, most of my past work has been around this area. Furthermore, I also enjoy working on large-scale software engineering problems. For the rest of this cover letter, I share some of my most recent works and my interests. I have worked on other software projects and non-first-author works as well.

Currently my primary focus is developing novel methods for 4D neural simulation. Most notably, my first research project at NVIDIA (under submission) contributed significantly to NVIDIA's Project Newton, featured at NVIDIA's GTC Keynote, where I designed neural models for physically accurate, high-fidelity simulations enabling realistic deformations and contact dynamics. This research resulted in unprecedented simulation scalability using neural approaches, and we are currently exploring methods to entirely remove handcrafted components from such simulations.

Most recently at UofT, I finished working on a paper (under submission) with Prof. Nandita Vijaykumar on developing compressed latent space for 3D generators. Previously at UofT, I worked on SEE-2-SOUND (ICML Workshop 2024), the first zero shot model that generates 3D spatial audio from visual data, and NeRF-US (PMLR 2024), to train NeRFs on sound fields, HSI-Diffusion (SSC Oral 2024), to build diffusion models for hyperspectral imaging, DiffuseRAW (arXiv), to build diffusion models that could handle raw images, and Astroformer (ICLRW 2023), a transformer variant for low-data regimes.

Previously, I was a research intern at Qualcomm AI Research. I first-authored two papers. The first introduced a benchmark (AirLetters, ECCVW 2024) evaluating articulated motion understanding in video-language models. The second was QIVD (under submission, available on arXiv), a large multimodal model (video-audio-language) designed for real-time situated reasoning tasks, which addresses the limitations of existing models in real-world interactive scenarios. For both papers, I independently developed and executed large-scale experiments across extensive GPU clusters, writing highly performant experiment code.

I also believe I have a good understanding of software for machine learning. In the past, I have also worked very closely with CUDA and Triton, and I have also had one first-author and one second-author (oral) paper at the PyTorch Conference: one of which was about building specialized backprop kernels for NeRF training to potentially train NeRFs on small devices and the other was about building capabilities to AOT-compile NeRFs to WebAssembly code. These also required us to write new backprop kernels in CUDA and Triton. Furthermore, I am also one of the maintainers of Kubernetes, one of the most popular orchestrators for software at scale where I have contributed to building new capabilities in the project, most recently contributing to building "Dynamic Resource Allocation" capabilities which are popularly used by most large ML clusters.