Exploring Optimizations in the COSI High-Level Data Analysis Library

THE COLLEGE OF WOOSTER

Augustus Thomas [1, 2], Laura DeGroot [1], & John Musgrave[2]

[1] Department of Physics
[2] Department of Mathematical and Computational Sciences

Abstract

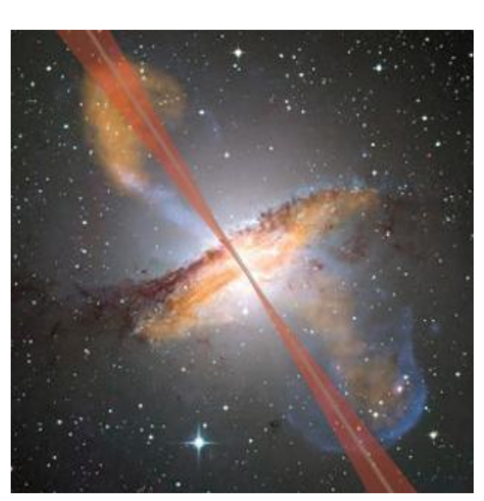
CosiPy is a custom-written data analysis Python package for the upcoming Compton Spectrometer and Imager (COSI) - a gamma-ray space telescope set for satellite launch in the near-future. The CosiPy data processing library will provide data retrieval and analysis functionality to astrophysicists, allowing them to make observations of nucleosynthesis, black hole accretion, and other gamma-ray band events using the COSI instrument's observational dataset.

This Independent Study represents a theoretical investigation into imaging algorithms, libraries and physical theory for Compton space telescopes in general and the COSI in particular. Theoretical work is done to explain de-convolution algorithms and the principle of image recovery from raw image datasets. Experimental work is done with the Compton experiment, to contextualize the raw COSI detector dataset and understand the Compton data space. Engineering work is done to accelerate the Cosipy library and meet performance requirements for the astrophysics community; in particular, I contribute memory and compute time optimizations to two different CosiPy modules.

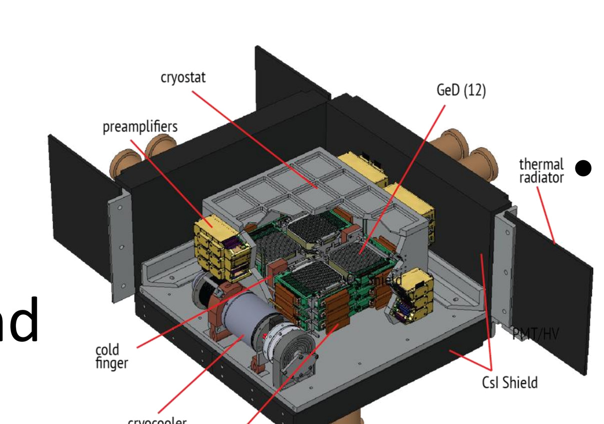
Introduction

COSI

- Satellite space telescope
- Low-energy gamma-ray band
- Gamma-ray bursts (GRBs) to study binary neutron star mergers, black hole accretion, supernovae, and active galactic nuclei



Theoretical GRB from AGN Centaurus A



Detector schematic, from [1]

CosiPy

- High level data analysis library for COSI (Python)
- Release of observations to astrophysics research community
- Methods include image deconvolution, data input/output, spectral fitting

Deconvolution: caching

- Recovering an original image after filtering from detector [4]
- Function spends 80% of its time retrieving value of a large array
- Cache immutable array

Image deconvolution method became 10x faster

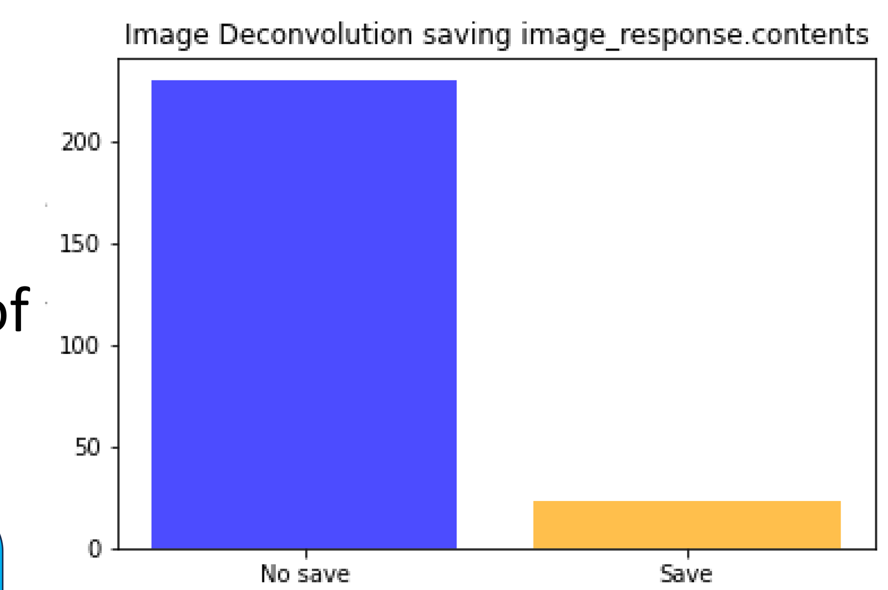


Figure 1: Performance Improvement from Caching Response Matrix (right) versus benchmark (left)

Deconvolution: JIT compiling

- Likelihood: estimated "correctness" of deconvolution
- Values in each cell in the image are independent
- Parallelize: multi-thread calculation

Calculating likelihood step became ~1.8x faster

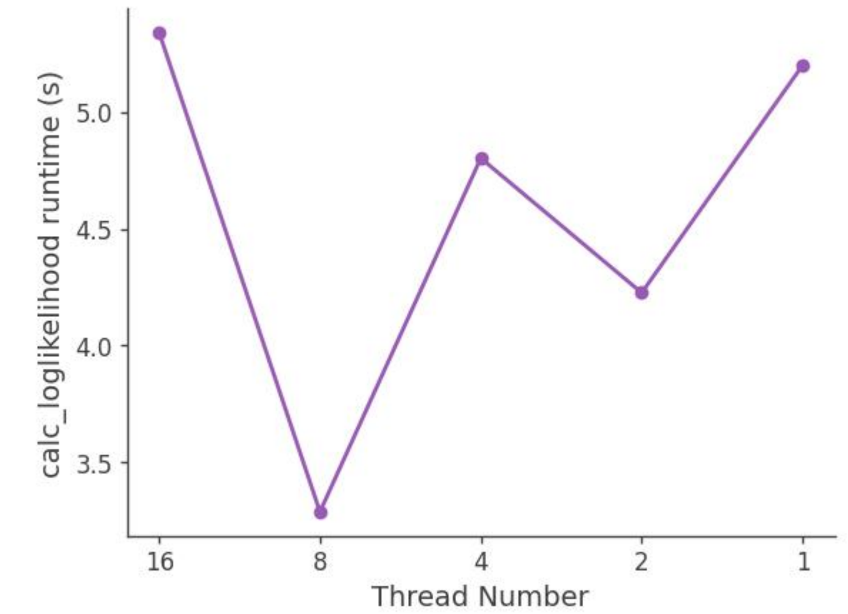


Figure 2: Execution latency for calculating loglikelihood in MREM algorithm as a function of Numba threads

dataIO: memory reduction

- Combine multiple images in a dataset into a single file
- Multiple copies of the same image are made
- Use headers to read images in dataset

Combining images requires ~1GB less RAM

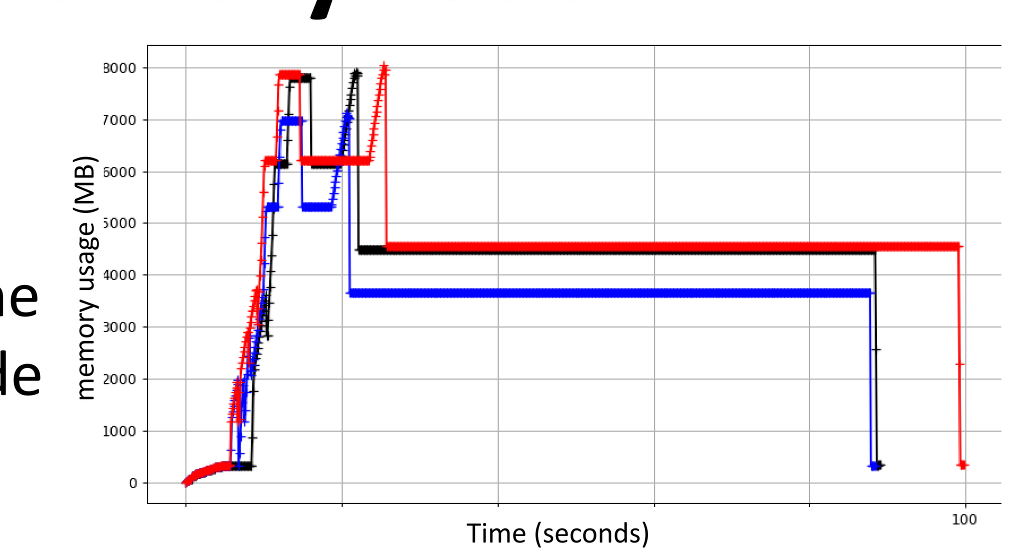


Figure 3: Memory profiles of dataIO.combine_unbinned method; original (red), speed optimization (black), memory optimization (blue)

Summary

- **Theoretical exploration:**
  - Image Data structures for computer science
  - Image Deconvolution algorithms for satellites
  - Physics of Compton scattering in detectors
- **Experimental results:**
  - Measuring the fundamental physics required for these detectors (rest energy of the electron)
- **Software Engineering contributions:**
  - Optimizing expensive components in cosipy (time, RAM)

Future Work

- **Techniques**
  - Automate runtime profiles with PyTests
  - Image processing on GPUs with CUDA
  - Other Python accelerators (Numba, JAXopt)
- **Library methods**
  - GRB localization module
- **Long-term goals**
  - On-device processing for multi-messenger astrophysics

References

[1]
A. Zoglauer *et al.*, "COSI: From Calibrations and Observations to All-sky Images," Feb. 25, 2021, *arXiv*: arXiv:2102.13158. doi: 10.48550/arXiv.2102.13158.

[2]
J. A. Tomsick et al, "The Compton Spectrometer and Imager," ," in *Proceedings of 38th International Cosmic Ray Conference — PoS(ICRC2023)*, Aug. 2023,, p. 785, doi: 10.22323/1.444.0745.

[3]
I. Martinez-Castellanos *et al.*, "The cosipy library: COSI's high-level analysis software," in *Proceedings of 38th International Cosmic Ray Conference — PoS(ICRC2023)*, Aug. 2023, p. 858. doi: 10.22323/1.444.0858.

[4]
W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration," *Journal of the Optical Society of America (1917-1983)*, vol. 62, p. 55, Jan. 1972.

Acknowledgements

Thank you Drs. Jeremy Buhler and Marion Sudvarg at the Washington University in Saint Louis for your outstanding outside collaboration and advisement; I would not have learned software optimization without you. Thank you Israel Martinez, Carolyn Kierans, John Tomsick, Andreas Zoglauer and the rest of the COSI team for inviting me to present my work. Thank you Dr. DeGroot, Dr. Manz and Dr. Musgrave for proof-reading my I.S. Thank you Dr. Heather Guarnera for helping me in my senior year. Thank you to the entire Physics faculty for patiently teaching me to be the scientist I am today.