Accelerating CERN and LIGO Big Data Processing Using an Optical Correlator

Tiwari, Aarushi (School: Adlai E. Stevenson High School)

The ever increasing demand for Big Data processing, ubiquitous in almost all fields of science, has become exceedingly resource intensive for endeavors such as CERN and LIGO. CERN alone produces more than 100 petabytes of data per year and it is expected to continue to grow. Both CERN and LIGO use electronic-based pre-filter algorithms to filter out insignificant data. However, optical-based systems can process up to 10^3 times faster than electronic systems. Therefore, this project investigates the use of high-speed Hybrid Opto-electronic Correlator (HOC) as a potential pre-filter system for data streams. Novel visualization methods needed to be designed as the HOC requires geometric representations of data as input. Similar to LIGO gravitational waves (GW), it was observed that typical CERN event histogram distributions of particle energy, momentum, transverse momentum, and mass change in a periodic-like fashion. Therefore, spectrogram images of LIGO and CERN datasets were generated. These spectrograms were tested for three simulated GWs from LIGO and four particle decay datasets from CERN using original code written for a simulation of the HOC on MATLAB. The simulation calculated the cross correlations and autocorrelations between each kind of image for both CERN and LIGO. Moreover, to make tests more realistic, white Gaussian noise was added to the spectrograms. The correlation coefficients for spectrogram images, with and without noise added, showed a greater average contrast for both CERN and LIGO data, compared to other iterations, which showed no meaningful contrast. This project successfully showed the potential of optical image correlation as a high-speed pre-filter system for Big Data processing using spectrogram images.