Architecture of ASTHROS' Spectral Data Collection and Onboard Anomaly Detection Pipeline

Paul Horton*[1], Christian Thompson[1], Christopher Groppi[1],
Jose V. Siles[2], and Jonathan H. Kawamura[2]

***Abstract*— Balloon borne missions have unique requirements concerning data collection, storage, and analysis. Due to the limited bandwidth between the instrument and ground team, all data collected must be stored on the instrument, requiring specialized storage solutions in cases with high data volume. Additionally, in-flight data assessment proves challenging as data typically cannot be analyzed until after the flight due to low bandwidth.**

ASTHROS (short for Astrophysics Stratospheric Telescope for High Spectral Resolution Observations at Submillimeter-wavelengths) is a high-altitude balloon mission utilizing multiple spectrometers to produce on-the-fly scans of gas and dust clouds in order to measure the density and dynamics of the specific elements in the clouds. For the spectrometers, we will be pairing four spectrometers to a Raspberry Pi Compute Module 4 (CM4) for readout. These readout computers will stream their data to a Network Attached Storage (NAS) device that will synchronize and safely store all the data collected on the flight. Collecting all of the data in a redundant and centralized location makes the data easier to analyze by other systems and makes the system easier to scale with more spectrometers.

Once all of the data is centralized to the NAS, a novelty detection pipeline will be utilized to identify bad calibration spectra. Using a variational autoencoder, we are able to create a system that can identify sudden changes in the science data quality by comparing the latent space representation of a given spectrum with a history of previous spectra. This will allow us to quickly assess changes in the spectrometer's behavior and downlink specific spectra to identify problems with the system. The variational autoencoder used here is a low power solution that can easily be run on-board on either the NAS or an additional Field Programmable Gate Array (FPGA). This solution allows us to identify quality issues with the spectrometer without having to blindly downlink data, thus helping in detecting problems as soon as they occur instead of after the flight.

Keywords—spectroscopy, system architecture, novelty detection, machine learning, operations, balloon-borne, stratospheric

References

[1] Růžička, Vít, et al., "Unsupervised Change Detection of Extreme Events Using ML On-Board", *Conference on Nerual Information Processing Systems (NeurIPS)*, Nov. 2021.

[2] Kerner, Hannah R., et al., "Domain-agnostic Outlier Ranking Algorithms (DORA): A Configurable Pipeline for Outlier Detection in Scientific Datasets", *American Geophysical Union (AGU) Fall Meeting,* Dec. 2021.

[3] Siles, Jose V., et al., "ASTHROS-Astrophysics stratospheric telescope for high-spectral resolution observation at submillimeter-waves: Mission overview and development status", *IEEE International Symposium on Space Terahertz Technology (ISSTT),* Mar. 2020

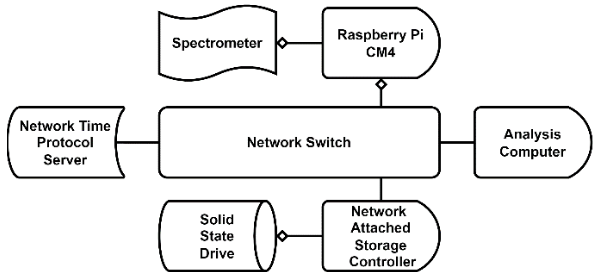

Fig. 1. System diagram for the spectrometer readout on ASTHROS. Up to 4 **Spectrometers** can be paired with a single **CM4** and multiple **CM4s** can be connected to the **Network Switch**. Each **CM4** will write data to **Solid State Drives** through the **Network Attached Storage Controller.** An **Analysis Computer** will be able to access the data and perform anomaly detection on the spectra. Everything will be synced with the time on the gondola using the gondola's **Network Time Protocol server.**

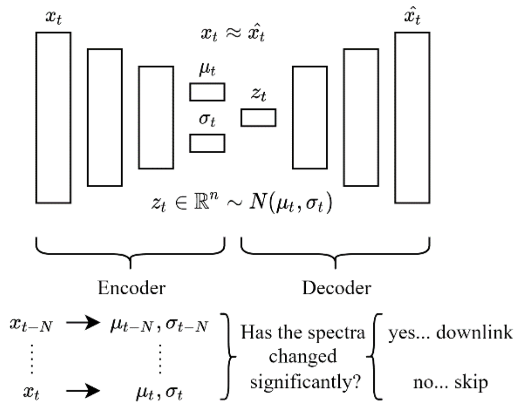

Fig. 2. Variational Autoencoder (VAE) diagram for anomaly detection on ASTHROS. New samples are encoded using the autoencoder and compared with previous samples to identify signifigant changes in calibration spectra.

[1]Arizona State University, School of Earth and Space Exploration, Tempe, Arizona, 85281, USA; [2]NASA Jet Propulsion Laboratory, Pasadena, California, 91109, USA

NOTES: