Semi-Empirical Data Compression for Heliophysics Space Mission Data

Center for HelioAnalytics

Daniel da Silva[123], Christopher Bard[1], Barbara Thompson[1], John Dorelli[1], Brian Thomas[1], Alex Barrie[14]

1 – NASA Goddard Space Flight Center; 2 – Laboratory for Atmospheric and Space Physics, University of Colorado Boulder; 3 – University of Maryland, Baltimore County; 4 – Aurora Engineering

1. Introduction

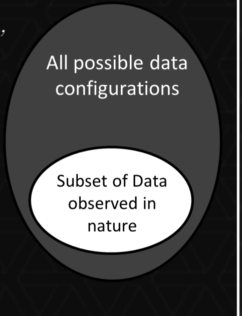
The ability for Heliophysics sensors to measure higher and higher resolution data has outpaced the ability to transmit the data. A mission's limited telemetry budget has become a bottleneck, with high-resolution measurements now being discarded simply because there is not enough bandwidth to transmit them. We present a new algorithm, SEPC (Semi-Empirical Plasma Compression) which implements data compression for ion velocity distribution functions in units of counts, validated through preservation of the derived plasma moments.

The algorithm utilizes a block-oriented transform method via a neural network auto-encoder to associate to-be-compressed measurements with previous measurements to reduce the dimensionality. The dimensionality reduction from the auto-encoder is followed by quantization of the floating-point coefficients and lossless entropy coding to produce a final compressed result. Applications for other type of Heliophysics space mission data such as solar imagery are expected to follow.

2. Semi-Empirical?

The term semi-empirical can be defined as "*involving assumptions, approximations, or generalizations designed to simplify calculation or to yield a result in accord with observation.*"

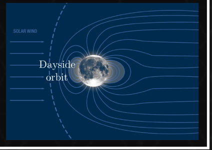
By utilizing auto-encoding technology, we can design compression algorithms for various types of data (e.g., multi-spectral imagery, in-situ velocity distribution functions) which use the traditional transform-method compression paradigm but with a transform method based on training data, therefore becoming empirical in nature. On top of this, the auto-encoder enforces its own mathematical structure onto the training data.



This approach contrasts general-purpose transform methods such as the Discrete Cosine Transform (used in JPEG) or Wavelet Transform (used in JPEG2000, DWT/BPE).

3. Our Demonstration (In-situ Ion Data)

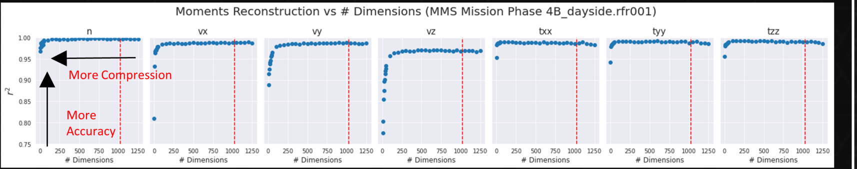
In this poster, we demonstrate the semi-empirical compression concept with an algorithm designed for ion velocity distribution function measurements, trained on data from the MMS FPI instrument on the day-side orbit.



4. Dimensionality Reduction

A neural network auto-encoder is used on patches of the 3D velocity structure. The architecture of the auto-encoder used was Multi-Layer-Perceptron with a single hidden layer. Experimentation with different latent sizes is done to determine the performance vs latent size curve. Heuristics are applied to force the average count to be the same before/after. It is found that a reasonable "knee" in the performance curve occurs around N=100, which corresponds to a dimensionality reduction of ~10X (= 16 · 32 · 2 / 100).

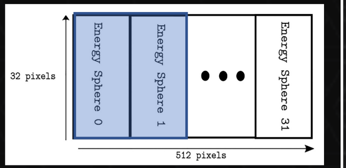
Future experimentation with different auto-encoder architectures is expected, including convolutional auto-encoders and GAN networks.



Above Figure: r^2 of each moment is calculated from the test set at different latent sizes (smaller ➔ more dimensionality reduction). The red line corresponds to where latent size equals the input size (no dimensionality reduction). The r^2 metric is chosen to prioritize the absolute scale over what might be acceptable noise, with the belief that a metric such as average relative error might unfairly penalize acceptable noise.

5. Blocking/Patching Methodology

The neural network auto-encoder is applied to individual *blocks* or *patches* of the 3D structure of each skymap. This approach is based on the block encoding approach utilized by JPEG, WEBP, and MPEG-4. It allows the transform encoder to be constrained to "local" velocity-space information.



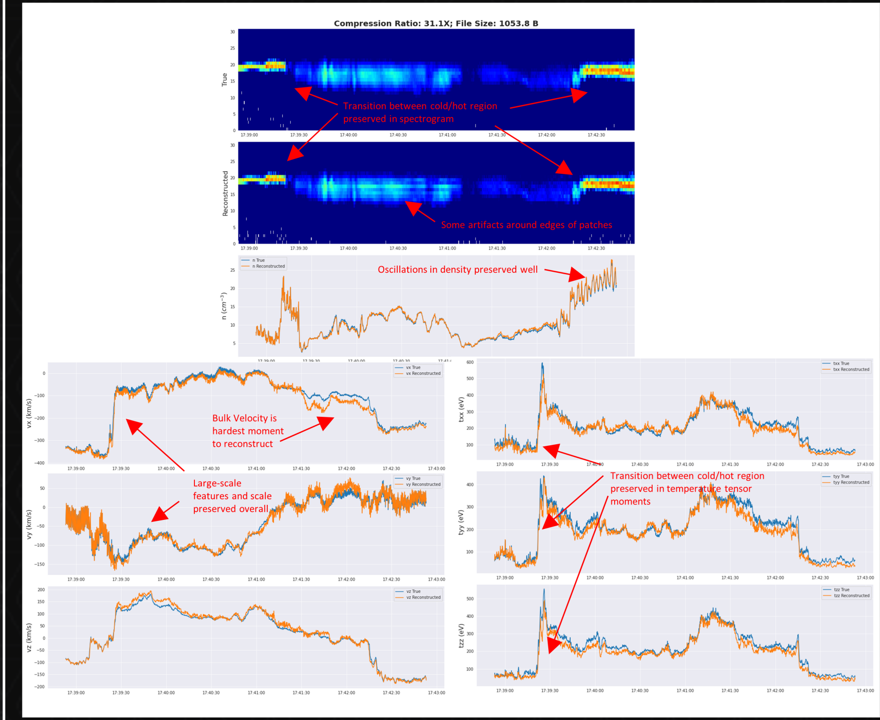
6. Latent Quantization and Entropy Coding

After the dimensionality is reduced using an auto-encoder, the latent vector coefficients are quantized. Quantization includes converting a FLOAT32 coefficient into a small digitized value to further reduce the bits per pixel. In the validation to the right, the quantization used is conversion to a FLOAT16 and the truncation of the decimal part for the number to a 10-bit "FLOAT10". This corresponds to size reduction from quantization of 1.6X.

Following this, the quantized values are passed through a lossless entropy coder, specifically the DEFALTE algorithm (commonly associated with ZIP and GZIP). This produces a size reduction from entropy coding of about 1.5-2.5X.

7. Validation of Spectrogram and Plasma Moments

The end-to-end compression system is tested on intervals from the test set. The total compression ratio is around 30X (≈ 10X · 1.6X · 2X) using the mentioned quantization and lossless entropy coding. This compares to a ratio of **17X on MMS/FPI Phase 1A's Fast Survey data** with the DWT/BPE algorithm (Barrie et al., 2018). In our validation, we look for what physics is preserved in the interval. The most important characteristics of a successful compression algorithm is that is does not lead to false conclusions, which is investigated in an example below.



8. Closing Remarks

A paper is being submitted to the Special Issue of *Frontiers in Astronomy and Space Sciences* associated with ML-Helio. We thank the Center for HelioAnalytics for the very important seed funding and community support, as well as the AWS Helioweb resources provided by Brian Thomas via the NASA/GSFC Heliophysics Sciences Division.