Adaptive Error Quantization for Lossy Compression of Scientific Datasets

Samuel Quan
Clemson University
Clemson, South Carolina, USA
quansamuelh@gmail.com

Jon C. Calhoun (Advisor)
Clemson University
Clemson, South Carolina, USA
jonccal@clemson.edu

ACM Reference Format:
Samuel Quan and Jon C. Calhoun (Advisor). 2026. Adaptive Error Quantization for Lossy Compression of Scientific Datasets. In *Proceedings of The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'26).* ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Background

Large-scale scientific instruments and applications generate massive amounts of data. The cost to transfer this data for analysis is a major bottleneck to workflow efficiency and scientific throughput. Lossy compression is one way to address this storage challenge. Error-bounded lossy compression allows a user to bound the error introduced during the lossy compression process by some user-defined metric. On floating point data, lossy compression achieves significantly higher compression ratios than lossless compression at a slight cost to data fidelity. Certain unpredictable datasets make it difficult to obtain large compression ratios with prediction based compressors without introducing a large error. To improve the compression quality for hard-to-compress data while minimizing compression overhead, this work introduces a new approach to lossy compression: adaptive error quantization.

Adaptive error quantization further refines the quantization of the error between the true value and the reconstructed value. When the original quantized error is too large or can be improved in a meaningful way, our method dynamically additionally refines the quantized error. Creating additional layers of quantization on top of the existing compressed value from a prediction based compressor increases the accuracy of the reconstructed value while adding minimal overhead. Adaptive error quantization is a powerful tool to aid scientists in fine-tuning error introduced in lossy-compressed scientific data sets while minimally impacting the compression ratio.

2 Methodology

Adaptive error quantization is a further refinement of the quantization of the error between the true value and the reconstructed value. Our method leverages the fact that the prediction mechanism within a lossy compressor is often *close enough*. When the prediction is not accurate enough, adaptive error quantization refines the prediction back to within the user-specified range. As our baseline lossy compressor, we use the prediction-based SZ3 compressor [1]. Our implementation takes advantage of SZ3's modular framework, acting as a customized quantizer in the full compression pipeline.

Our implementation of adaptive error quantization involves continuous bisection of the desired quantization bin. Given k quantization levels, the single quantization bin from SZ3 is bisected k times, using one-hot encoding to indicate whether to bisect to the left or right, determining whether the true value is less than (0) or greater than (1) the predicted value. For example, if $k = 2$ and the true value is less than the predicted value through two sequential bisections, the resulting k-bit adaptive quantization code is 00. We further apply variable-length encoding as lossless compression for these adaptive quantization codes. To indicate adaptive quantization occurred, an additional quantization code must be reserved, leaving one less bin available to map an error to.

Using adaptive quantization requires storing adaptive quantization bits in addition to the original quantization array. The method used to store this info can significantly affect the final compression ratio. This work developed and tested several of these methods, including appending adaptive codes to the end of the single quantization codes, "packing" away unused bins, and performing Huffman encoding on them.

3 Results

We evaluate adaptive quantization against single quantization based on peak signal to noise ratio (PSNR), given by $20 \cdot log_{10} \frac{R_{max}}{RMSE}$ where $RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (e_{abs})^2}$ is the root mean squared error. We also compare the compression ratio obtained with each method to estimate the overhead incurred from the additional bits required to store adaptive quantization codes.

Our implementation was evaluated on the HACC particle simulation data set [2]. This data set contains snapshots of particle data spread across 6 different fields, stored in a single-precision floating-point format. We fix the number of single quantization bins to 2^{10} bins ($m = 10$), and vary the number of adaptive quantization levels from 1, ..., 4 as well as the error bound eb from 1×10^{-1}, ..., 1×10^{-5}. HACC fields represent either a velocity or position component of a particle in a three-dimensional coordinate system. For clarity, we present our adaptive compression pipeline performance on an example from each "type" of field in figures 1, 4, and **??** due to their similar performances.

Figure 1 demonstrates the achieved compression ratio vs the user set error bound of our compression pipelines on fields *vx*, representing velocity fields, and *xx*, representing position fields.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SC'26, Chicago, IL
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

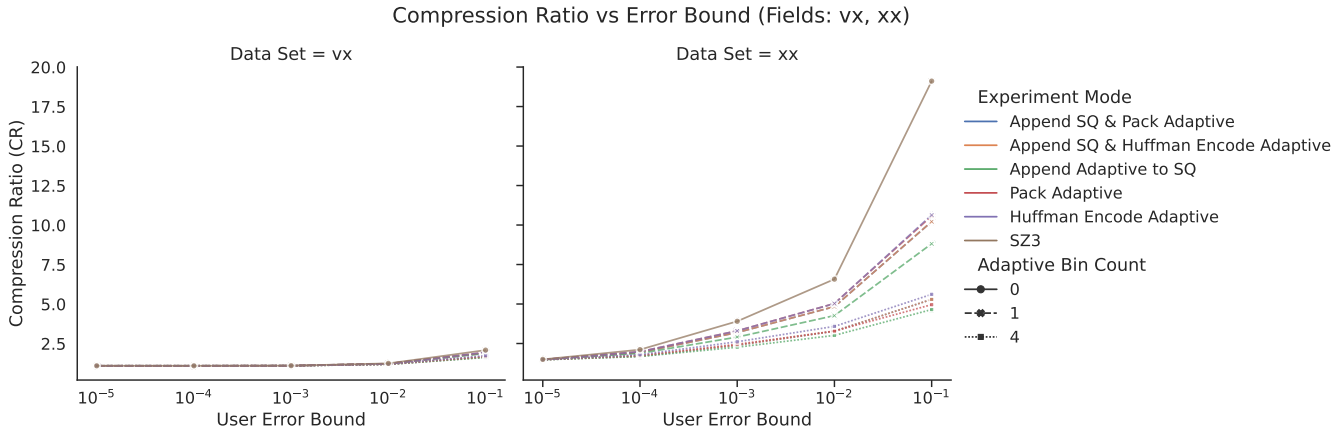


Figure 1: Compression Ratio vs Error Bound performance for compression pipelines on the *vx* and *xx* HACC fields

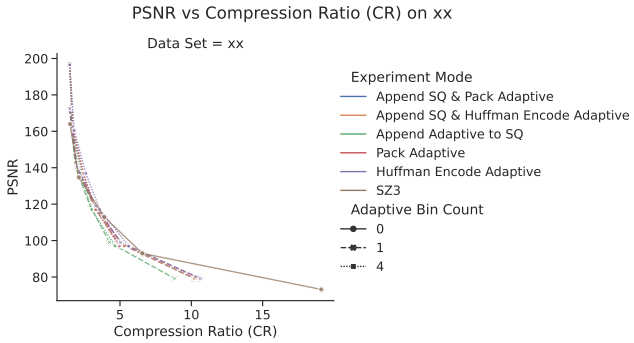


Figure 2: PSNR vs Compression Ratio performance for compression pipelines on the *xx* HACC field

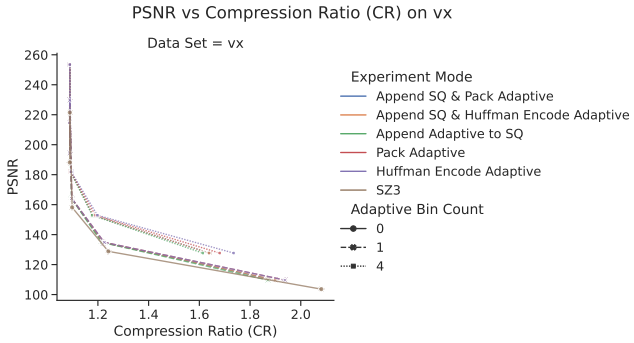


Figure 3: PSNR vs Compression Ratio performance for compression pipelines on the *vx* HACC field

While our method incurs a trade-off in compression ratio over SZ3, this trade-off is significantly smaller on the harder to compress velocity fields as compared to the more predictable position fields.

Velocity is a harder to compress field than position, with our methods incurring a significantly lower trade-off in compression ratio over the baseline SZ3. As we expect, increasing the adaptive bin count results in a lower compression ratio due to the extra overhead in storing additional adaptive bins for every data point.

Figure 4 shows how PSNR scales over the error bound across our methods. Our adaptive methods achieve improved PSNR values over SZ3 in both the Velocity and Position fields, achieving higher PSNR with an increasing adaptive bin count. In particular, our methods achieve similar PSNR improvements while incurring a much smaller compression cost on the velocity field compared to the position field.

Figures 3 and 2 show the achieved PSNR at a given CR for a compression run on the vx and xx fields respectively. A compression run that achieves a higher PSNR for a fixed CR, or vice versa, represents a demonstrable compression advantage over other methods. Figure 3 shows that adaptive quantization maintains an advantage over baseline SZ3 for any chosen compression ratio on the hard to compress velocity field. On the other hand, Figure 2 shows adaptive methods are advantageous only for lower desired compression ratios.

4 Conclusions

In this work, we introduced adaptive error quantization, a novel extension to the prediction-based lossy compressors SZ3 designed to improve data quality for the lossy compression of scientific datasets. Dynamic refinement of the error quantization bins allows scientists to lossily compress scientific data at higher fidelity without sacrificing compression ratio.

Our evaluation on the HACC particle simulation dataset demonstrates that adaptive quantization is a highly effective tool for unpredictable data. For high-entropy velocity fields, our method maintains a strict data fidelity advantage over the baseline SZ3 compressor across all evaluated compression ratios, significantly improving the PSNR with minimal overhead. While the method also improves PSNR for more predictable position fields, the storage overhead trade-off limits its advantage to lower desired compression ratios. Adaptive error quantization has potential to be a valuable tool for fine-tuning compression ratios with acceptable error bounds for scientists working on hard-to-compress datasets.

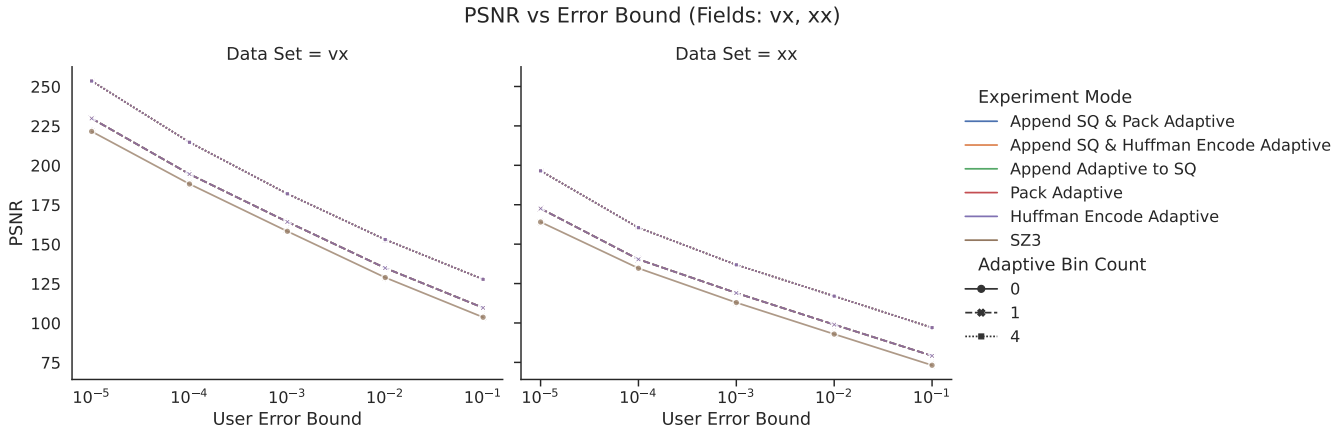


Figure 4: PSNR vs Error Bound performance for compression pipelines on the *vx* and *xx* HACC fields

References

[1] Xin Liang, Kai Zhao, Sheng Di, Sihuan Li, Robert Underwood, Ali M. Gok, Jiannan Tian, Junjing Deng, Jon C. Calhoun, Dingwen Tao, Zizhong Chen, and Franck Cappello. 2023. SZ3: A Modular Framework for Composing Prediction-Based Error-Bounded Lossy Compressors . *IEEE Transactions on Big Data* 9, 02 (April 2023), 485–498. doi:10.1109/TBDATA.2022.3201176

[2] Kai Zhao, Sheng Di, Xin Liang, Sihuan Li, Dingwen Tao, Julie Bessac, Zizhong Chen, and Franck Cappello. 2021. SDRBench: Scientific Data Reduction Benchmark for Lossy Compressors. *CoRR* abs/2101.03201 (2021). arXiv:2101.03201 https://arxiv.org/abs/2101.03201