Submission code: VQ7LZ6K7DB

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Effect of 3D ZFP lossy compression of background wavefield in LSRTM

Átila Soares (University of Alberta; Department of Physics), Mauricio Sacchi (University of Alberta; Department of Physics)

Effect of 3D ZFP lossy compression of background wavefield in LSRTM

Abstract Summary

Least-squares reverse time migration gained attention in the last decade, given its ability to produce high-quality subsurface images with few iterations. However, this technique requires storing the entire background wavefield before calculating the misfit gradient. The size of this wavefield when working with 3D models can easily be several terabytes. Despite the plethora of techniques to trade storage by floating point arithmetic computation, they all still require much memory to operate. In this study, we employ the ZFP lossy compression algorithm to compress the entire background wavefield using a 3D model and study its effect on convergence, accuracy, and image quality. The present work is an extension of a previous work in which we examined the 2D model case.

Introduction

The main bottleneck of techniques like least-squares reverse time migration (LSRTM) involves storing the background wavefield computed in the forward modeling pass, as it is used in calculating the gradient while marching the adjoint wavefield backward in time. The reason for such a storage bottleneck is because it is an array of spatial and time dimensions, totaling several terabytes while using 3D models. Techniques such as checkpointing (Griewank, 1992). decimated or interpolated reconstructions (Yang et al., 2016), and effective-boundary schemes (Dussaud et al., 2008) all trade storage for extra arithmetic, yet none scale gracefully to modern 3D problems. The simples solution is to store in disk, but that can be slow. However, with the advances in solid state drive technology, new NVMe drives can reach upwards of 5 GB/s throughput, alleviating this bottleneck.

This work proposes applying the ZFP floating-point lossy compression algorithm (Lindstrom, 2014) to store the background wavefield permanently. The premise is that if a modest, quantifiable loss in numerical fidelity preserves the LSRTM final image quality, then the storage footprint (and the wear on solid-state drives) can fall by an order of magnitude without disturbing overall convergence. This work is a 3D extension from Soares and Sacchi (2024), which shows similar results, although with a different model. Although lossy compression has been tested in FWI before (Kukreja et al., 2022) to our knowledge, this is the first work on the 3D LSRTM case.

This paper starts with an explanation of the methods used in the experiments, followed by the results and conclusions. The results show that the final migrated image has minimal migration artifacts with essentially unaffected convergence.

Method

We implement LSRTM in the linearized acoustic (Born) setting, treating $m = \delta(1/c^2)$ as the model update. Background and scattered fields are propagated with a finite-difference solver; adjoint correctness is confirmed via dot-product tests. The quadratic misfit

$$J(m) \;=\; \frac{1}{2}\sum_{r=1}^{N_r}\sum_{s=1}^{N_s}\int_0^T \left(S_{s,r}\left\{\delta u_s(x,y,z,t)\right\}(t) - d_{s,r}(t)\right)^2 \mathrm{d}t, \tag{1}$$

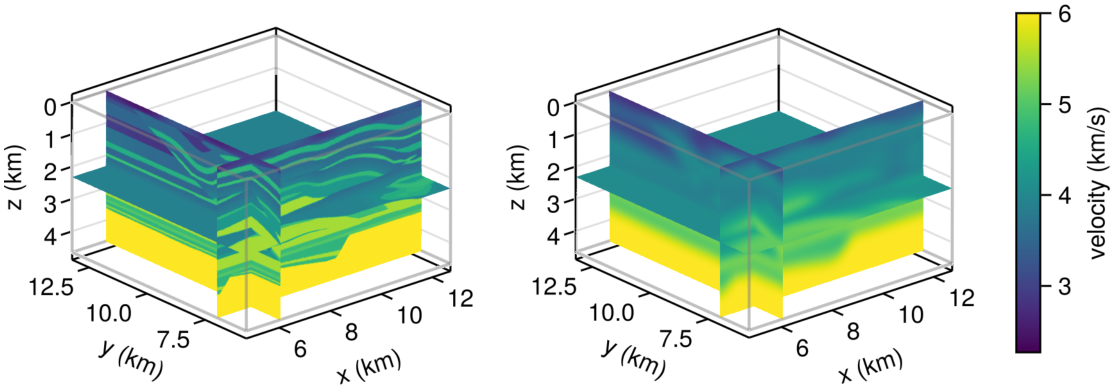

Figure 1: Sliced portion of the Overthrust model (Aminzadeh, 1996) in the numerical examples section. The resulting model has dimensions $\{nz, ny, nx\} = \{187, 301, 321\}$, and grid spacing of $\{dz, dy, dx\} = \{25, 25, 25\}\ m$. The model in the left was used to create the synthetic observed data, while the smoothed one in the right was used in the migration.

is minimized with ten conjugate-gradient least-squares (CGLS) (Hestenes and Stiefel, 1952) iterations. The gradient at each step is the zero-lag correlation of the time-reversed residual with ∂_t^2 of the background field. Here, N_r and N_s denote the numbers of receivers and sources, respectively. $S_{s,r}$ samples the scattered field δu at receiver r, and $d_{s,r}$ is the recorded trace for each source–receiver pair.

Storing the full 3-D wavefield is prohibitive, so we compress it on-the-fly with ZFP (fixed-rate or max-tolerance modes) via the `SequentialZfpCompression.jl` package. The latter slices each snapshot along the slowest axis, with each slice being compressed in parallel. ZFP then partitions the slice into 4^d blocks, transforms each block to a decorrelated bit-plane representation, and discards trailing planes according to a user-specified bit budget per digit in a block (fixed-rate mode) or a maximum tolerance of the difference between the lossless and lossy values in the block (maximum tolerance mode). In a nutshell, using a smaller rate or a higher tolerance makes the compression more lossy.

Results

We ran two tests to quantify compression effects. **(i) Three-layer model:** LSRTM on a 50^3 grid (10 m spacing, single shot) shows that lossless compression slows runtime by $\sim 20\%$. A rate = 4 or tolerance 10^{-2} restores baseline speed for both RTM and LSRTM. Convergence remains unchanged until iteration 7, after which only the 4 bpd run drifts slightly (Figs. 2–3). **(ii) Overthrust slice:** Migrating 120 shots (24.3 TB wavefield) using the models shown on Fig 1 with rate = 4 yields an 84 % space saving; tolerance 10^{-2} saves 75 %. Both settings preserve image quality (Fig. 4).

Conclusions

In summary, using ZFP compression yielded a significant space savings requirements. For instance, using rate=4 one can expect to save around 84 % in storage while maintaining good convergence, accuracy and similar image quality. There is also the potential to be coupled with other methods like

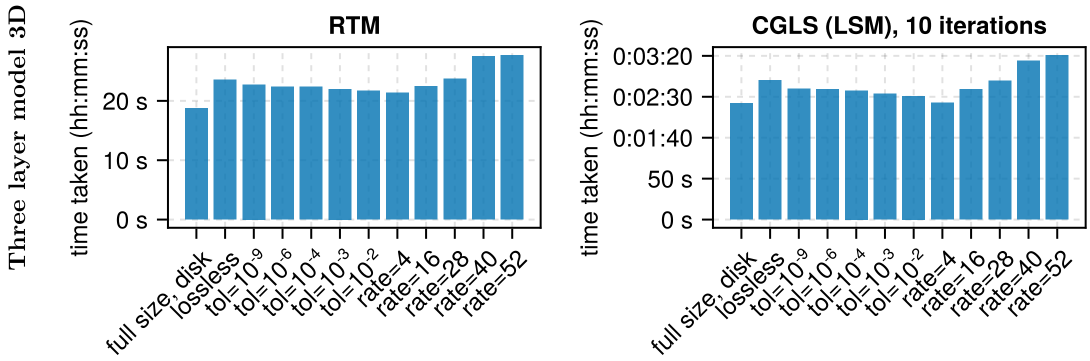

Figure 2: Execution time for RTM and CGLS (conjugate gradient least squares) migration for a three-layer model for different combination of lossy compression parameters. Only one shot was used in the experiment.

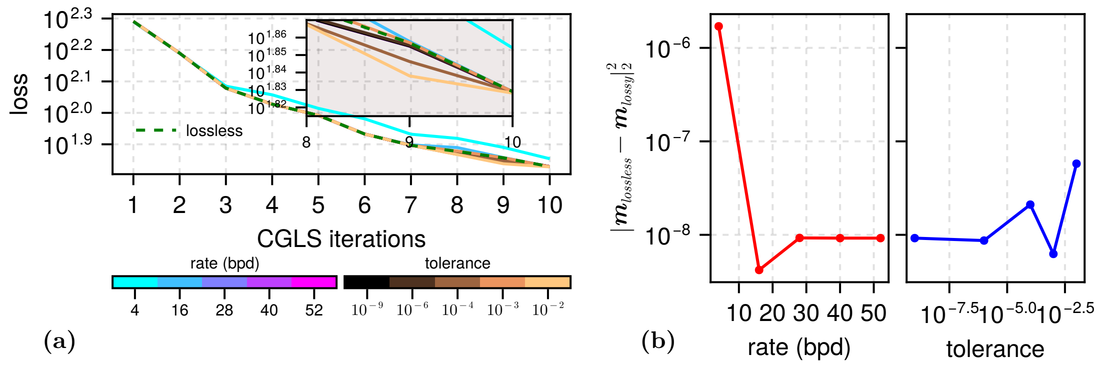

Figure 3: (a) CGLS convergence for 10 iterations using the three-layer model, and (b) L2-norm of the difference between the model obtained using lossy compression on the background wavefield, for different rates and tolerances.

checkpointing and effective boundary, for even greater savings. This would potentially allow hybrid disk and memory solutions.

Acknowledgments

We want to thank the Signal Analysis and Imaging Group sponsors at the University of Alberta for supporting the stimulating research environment that allowed the preparation of this work.

References

Aminzadeh, F., 1996, 3-D Salt and Overthrust Seismic Models.

Dussaud, E., W. W. Symes, P. Williamson, L. Lemaistre, P. Singer, B. Denel, and A. Cherrett, 2008, Computational strategies for reverse-time migration: SEG Technical Program Expanded Abstracts 2008, Society of Exploration Geophysicists, 2267–2271.

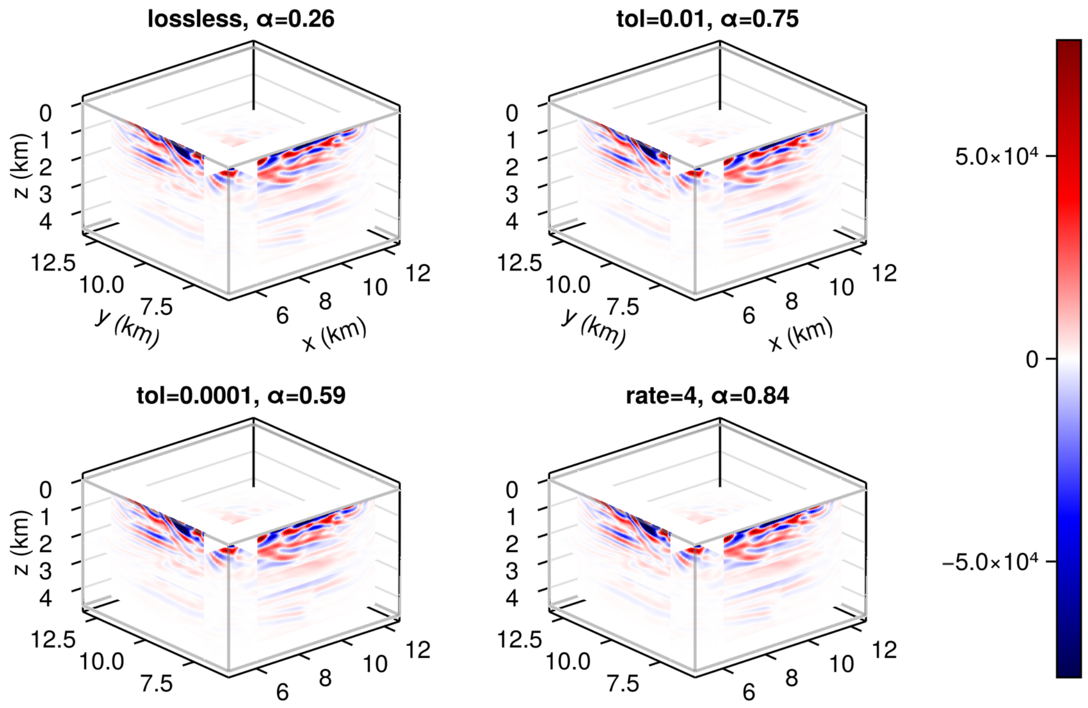

Figure 4: RTM image for different compression tolerances using the smoothed overthrust model (Fig. 1). The α term is the space-saving relative to the full size of the background wavefield (size=24.28 TB). This experiment used 120 shots. The final figure uses a Laplacian filter and a mute in the first 300 meters of depth.

Griewank, A., 1992, Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation: Optimization Methods and Software, **1**, 35–54.

Hestenes, M., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: Journal of Research of the National Bureau of Standards, **49**, 409.

Kukreja, N., J. Hückelheim, M. Louboutin, J. Washbourne, P. H. J. Kelly, and G. J. Gorman, 2022, Lossy checkpoint compression in full waveform inversion: a case study with ZFPv0.5.5 and the overthrust model: Geoscientific Model Development, **15**, 3815–3829.

Lindstrom, P., 2014, Fixed-Rate Compressed Floating-Point Arrays: IEEE Transactions on Visualization and Computer Graphics, **20**, 2674–2683.

Soares, Á. S. Q., and M. Sacchi, 2024, A study on lossy compression for background wavefield storage in LSM: Fourth International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists and American Association of Petroleum Geologists, 2077–2081.

Yang, P., R. Brossier, and J. Virieux, 2016, Wavefield reconstruction by interpolating significantly decimated boundaries: GEOPHYSICS, **81**, T197–T209.