PIRT- Parallel Iterative Reconstruction Tomography, with correction for center of rotation errors

Sajid Ali[1], Matthew Otten[2] & Wendy Di[3]

[1]Applied Physics
Northwestern University
[2]HRL Laboratories
[3]Math. & Comp. Sci. Divison
Argonne National Lab

International HPC Summer School 2021 July 2021

Outline

Outline

Basics

- Radon transform : Real $\rightleftarrows$ Sinogram space.
- $Rf(\tau,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(\tau - x\cos(\theta) - y\sin(\theta))dxdy$

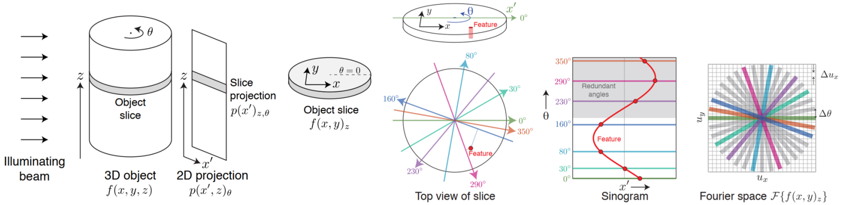


Figure: Spinning the object to obatin "sinograms", reconstruct each slice independently. Figure taken from Jacobsen [2019]

Center of rotation drifts

- $P_\theta = x^*_\theta(1 - cos(\theta) + y^*_\theta sin(\theta))$
- $Rf(\tau, \theta, 0, 0) = Rf(\tau - P_\theta, \theta, x^*_\theta, y^*_\theta)$
- Translation of sinogram by P_θ achieved by convolution with Gaussian.
- Recover P_θ to obtain accurate reconstruction[1].

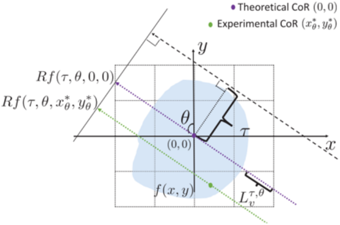


Figure: Center of rotation drift causes us to measure the shifted sinograms, figure from Austin et al. [2019]

[1] Austin et al. [2019]

Tomography software

- Efficient distributed memory parallel tomography software available for CPU and GPU [2]
- Error correcting capabilities available in popular tomography packages like TomoPy [3]
- Need a software package that has error correction capabilities and is distributed memory parallel

[2] Bicer et al. [2017]; Chen et al. [2019]; Hidayetoğlu et al. [2019]; Marchesini et al. [2020]; Palenstijn et al. [2016]

[3] Gürsoy et al. [2017]

Outline

Tomography
Algorithm & Implementation
3D solver

Optimization formulation

Discretize & Vectorize

- $\mathcal{W}$: object vector
- $\mathcal{L}$: discretized Radon transform
- $\mathcal{D}$: measured sinogram

Least squares cost function

- Assuming no shifts, we need $\underset{\mathcal{W}\geq 0}{min}\frac{1}{2}||\mathcal{L}\mathcal{W}-\mathcal{D}||$
- To recover both shifts and object : $\underset{\mathcal{W}\geq 0, P_\theta}{min}\phi(\mathcal{W}, P_\theta) = \frac{1}{2}||\mathcal{L}\mathcal{W}-g(\mathcal{D}, P_\theta)||$
- First order derivatives analytically computable : $\nabla\phi(\mathcal{W}, P_\theta) = [\mathcal{L}^T, \nabla_{P_\theta}\phi(\mathcal{W}, P_\theta)]^T(\mathcal{L}\mathcal{W}-g(\mathcal{D}, P_\theta))$

Implementation

- Implemented in C++ using :
  - PETSc/TAO (optimization routines, data management and parallel I/O)
  - Boost (geometry routines)
  - FFTW (fourier space convolution)

Joint

- Combine shifts and sample into one vector and optimize for both together

Alternating

- Alternate between optimizing with respect to sample and with respect to shifts

Demonstration

- Quality of reconstructed image and corrected sinograms, with and without error correction.

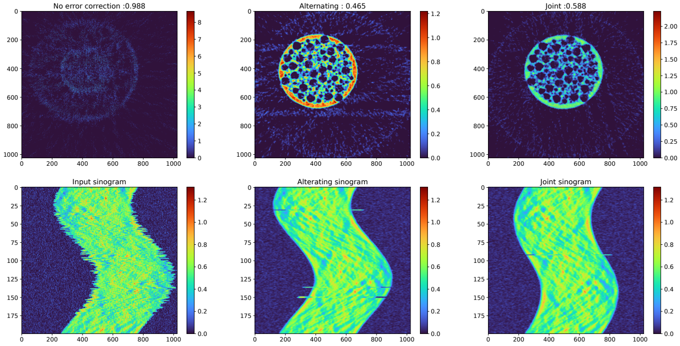


Figure: Demonstration of PIRT solvers : (Left to Right) with no error correction, alternating solver, joint solver. Test object is 1024^2 with 200 projections, with 10% center of rotation drifts and 10 % added noise.

Initial Results

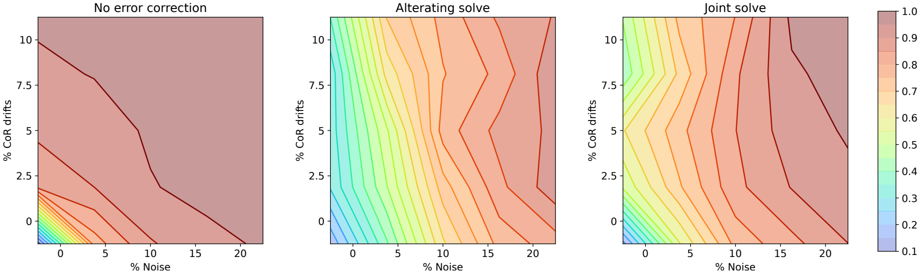


Figure: Quality of reconstruction as a function of center of rotation drifts and added noise for an object with dimensions 1024^2 and 200 projections, metric $\rightarrow$ translation invariant normalized root mean square error, no regularization.

Data decomposition

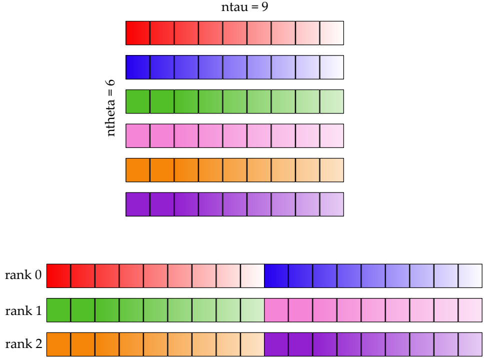


Figure: Naive data distribution over MPI ranks using PETSc heuristics, with the only constraint being that we don't distribute a single projection over multiple MPI ranks. *ntheta* $\rightarrow$ number of projections, *ntau* $\rightarrow$ size of each projection.

2D solve scaling

- Tests conducted on LCRC-bebop, dual-socket Xeon Broadwell, Intel Omni Patch interconnect.

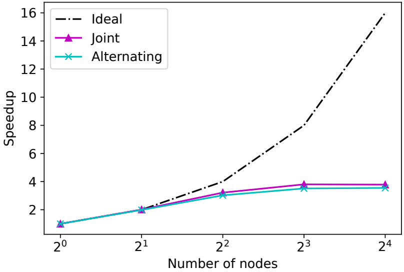


Figure: Strong scaling → object size 2896^2, 800 projections.

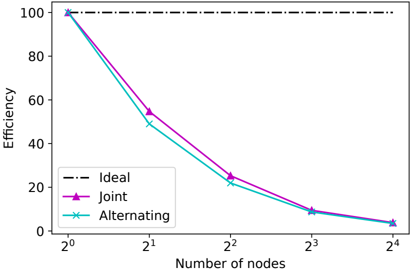


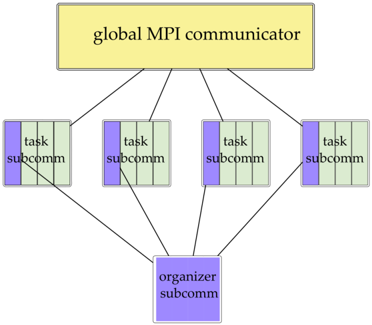
Figure: Weak scaling object size → 1024^2, projections per node 100

Outline

Research choices

- Profiling showed some improvement post enhancements but strong and weak scaling efficieny falls off beyond 4 nodes for a single 2D slice solve.
- Reason being the naive data distribution leading to load balancing issues, ≈ 80 of time being spent in matrix-vector and transposed matrix-vector multiplications.
- However, experimental data is always a 3D data-set! Focus on throughput instead of scalability of a single 2D slice solve.
- Introduced MPI-subcommunicator based 2-level MPI parallelism, with multiple concurrent instances of the solver.

2 level MPI parallelism



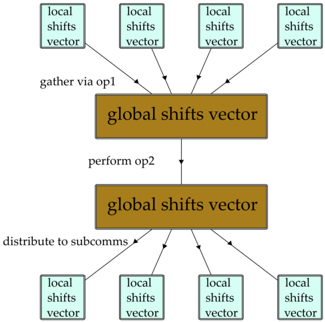


Figure: (Left) Architecture of the 2-level parallelism. Each solver instance solves for a "batch" of 2D slices (Right) Idea on accelerating convergence of 3D solver by combining results of 2D solves. To combine the shifts from subcomms, run multiple "overall sweeps" with checkpoint-restart.

Throughput Results

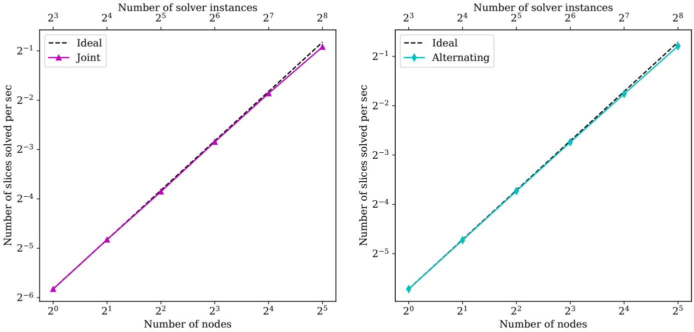


Figure: Number of slices solved per unit time, 256 slices of size 1024^2 and 200 projection angles. Tests conducted on LCRC-bebop, dual-socket Xeon Broadwell, Intel Omni Patch interconnect.

Acknowledgments

- Wendy Di MCS/APS-ANL, internship & thesis advisor
- Matthew Otten CNM-ANL, internship co-advisor
- Chris Jacobsen NU & XSD-ANL, thesis advisor
- Alp Dener, TAO-developer, MCS/ANL
- Hong Zhang/Barry Smith, GPU hackaton mentors/PETSc developers, MCS/ANL
- PETSc developers on the mailing lists
- bebop-LCRC ANL & theta-GPU ALCF for resources
- NIMH & ANL-MCS for funding

References I

[Austin et al. 2019] AUSTIN, Anthony P. ; DI, Zichao ; LEYFFER, Sven ; WILD, Stefan M.: Simultaneous Sensing Error Recovery and Tomographic Inversion Using an Optimization-Based Approach. In: *SIAM Journal on Scientific Computing* 41 (2019), Nr. 3, S. B497–B521

[Bicer et al. 2017] BICER, Tekin ; GÜRSOY, Doğa ; ANDRADE, Vincent D. ; KETTIMUTHU, Rajkumar ; SCULLIN, William ; CARLO, Francesco D. ; FOSTER, Ian T.: Trace: a high-throughput tomographic reconstruction engine for large-scale datasets. In: *Advanced Structural and Chemical Imaging* 3 (2017), Jan, Nr. 1, S. 6

[Chen et al. 2019] CHEN, Peng ; WAHIB, Mohamed ; TAKIZAWA, Shinichiro ; TAKANO, Ryousei ; MATSUOKA, Satoshi: IFDK: A Scalable Framework for Instant High-Resolution Image Reconstruction. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. New York, NY, USA : Association for Computing Machinery, 2019 (SC '19). – URL `https://doi.org/10.1145/3295500.3356163`. – ISBN 9781450362290

References II

[Gürsoy et al. 2017] GÜRSOY, Doğa ; HONG, Young P. ; HE, Kuan ; HUJSAK, Karl ; YOO, Seunghwan ; CHEN, Si ; LI, Yue ; GE, Mingyuan ; MILLER, Lisa M. ; CHU, Yong S. ; DE ANDRADE, Vincent ; HE, Kai ; COSSAIRT, Oliver ; KATSAGGELOS, Aggelos K. ; JACOBSEN, Chris: Rapid alignment of nanotomography data using joint iterative reconstruction and reprojection. In: *Scientific Reports* 7 (2017), Sep, Nr. 1, S. 11818

[Hidayetoğlu et al. 2019] HIDAYETOĞLU, Mert ; BIÇER, Tekin ; GONZALO, Simon G. de ; REN, Bin ; GÜRSOY, Doğa ; KETTIMUTHU, Rajkumar ; FOSTER, Ian T. ; HWU, Wen-mei W.: MemXCT: Memory-Centric X-Ray CT Reconstruction with Massive Parallelization. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. New York, NY, USA : Association for Computing Machinery, 2019 (SC '19). – ISBN 9781450362290

[Jacobsen 2019] JACOBSEN, Chris: *X-Ray Microscopy*. Cambridge University Press, 2019 (Advances in Microscopy and Microanalysis)

References III

[Marchesini et al. 2020] MARCHESINI, Stefano ; TRIVEDI, Anuradha ; ENFEDAQUE, Pablo ; PERCIANO, Talita ; PARKINSON, editor="Krzhizhanovskaya Valeria V. ; ZÁVODSZKY, Gábor ; LEES, Michael H. ; DONGARRA, Jack J. ; SLOOT, Peter M. A. ; BRISSOS, Sérgio ; TEIXEIRA, João": Sparse Matrix-Based HPC Tomography. In: *Computational Science – ICCS 2020*. Cham : Springer International Publishing", 2020, S. 248–261

[Palenstijn et al. 2016] PALENSTIJN, Willem J. ; BÉDORF, Jeroen ; SIJBERS, Jan ; BATENBURG, K. J.: A distributed ASTRA toolbox. In: *Advanced Structural and Chemical Imaging* 2 (2016), Dec, Nr. 1, S. 19