**Automatic and Robust Data processing for Global Adjoint Tomography**

Wenjie Lei, Youyi Ruan, Jame Smith, Matthieu Lefebvre, Ebru Bozdag, Daniel Peter, and Jeroen Tromp

Princeton University

lei@princeton.edu

**Abstract**

Global adjoint tomography involves a complex workflow. The seismic wave propagation solver, used both for forward and adjoint runs, is the most computationally intensive part. Porting it to GPUs has helped us to simulate seismic waves at increased resolution while reducing the time to solution. Another important part of the workflow is data processing. While not as computationally intensive, its sensitivity to human mistakes requires careful handling and safeguards. To better accommodate the growing volume of seismic data, a new data format, called the Adaptable Seismic Data Format (ASDF), is designed based on modern data container standards. It greatly enhances the reproducibility and accountability of our research, but also reduces the risk of errors and failures potentially contaminating the scientific result. Accordingly, a whole set of data processing tools has been re-developed based upon the new data format using Python, with internal parallel I/O support and better exception handling. This set of Python software is carefully designed so their integration in workflow management systems will be seamless. Future work will strive to improve data quality detection and assimilation by adding data-intelligent algorithms able to automatically tune parameters for the data processing chain.