Development and Implementation of Full 3-Dimensional Waveform Tomography
Report for SCEC Award 18121, May 6, 2019

Principal Investigator: Thomas H. Jordan, University of Southern California
Graduate Student: Alan Juarez, University of Southern California

We proposed to assemble software elements for automating F3DT on high-performance computers using the Hercules Toolchain. The F3DT toolkit includes subroutines for automated seismic source inversion, data extraction and an iterative tomography algorithm. We report on progress in developing several toolkits, and we raise one issue regarding how our work can best be coordinated with SCEC CME objectives.

(1) Database administration. The toolkit earthquake_database.m creates a database with all the earthquakes selected for the tomographic inversion. The entries in the database include the source locations, the moment-tensor elements obtained from the Global-CMT catalogue (GCMT; Ekström et al., 2012), and the waveforms at all the receivers within the simulation domain. The toolkit reads waveforms from the standard SAC format (Goldstein & Snoke, 2005; Goldstain et al., 2003), performs the instrument correction, and computes the signal-to-noise ratio to select waveforms suitable for the inversion. The selected seismograms and the station information are attached to the database. The toolkit also includes a utility to analyse Green functions computed from ambient noise crosscorrelation. Finally, earthquake_database.m creates all the necessary inputs for the forward simulation with the Hercules Toolchain (Tu et al., 2006; Taborda & Bielak, 2011). A simulation is prepared for each earthquake in the database. The velocity model in Hercules format and the mesh to store the wavefields are constructed independently. However, we have developed subroutines to operate them (hercules_grid.m & model_files.m). These subroutines can handle transformation between geographic and Cartesian coordinates and rotated grids.

(2) Seismic source inversion. We developed the toolkit loc_cmt.m that estimates the centroid-moment tensor (CMT) parameters of the seismic source using three-dimensional (3D) Green functions. The CMT is calculated by minimizing the squared misfit between three-component synthetic and observed waveforms. The optimization problem is expressed as a filtered version of $\mathbf{u} = \mathbf{G}\mathbf{m}$, where $\mathbf{m}$ is a vector with the six-independent CMT components, $\mathbf{G}$ is a matrix constructed with the synthetic Green functions, and $\mathbf{u}$ is a vector with the observed seismograms. The observed data are expressed as linear combinations of the Green functions and the elements of the CMT.

Synthetic receiver-side Strain Green Tensors (SGT; Zhao et al., 2006) for each station, which have been calculated for a 3D model using the Hercules Toolchain, are stored in a grid of fixed (typically 2-km) resolution. loc_cmt.m uses a preliminary location taken from an earthquake database and searches the closest points in the grid, loads the Green functions and estimates preliminary CMT's. The algorithm initializes the CMT by minimizing the squared differences between data and synthetics, and the locations are iterated to convergence holding the 3D model fixed. The storage expense can be reduced by reducing the grid resolution. The development of the toolkit is mature, but it needs further testing both as an isolated code and as a component of the tomography workflow.

(3) Toolkit for extracting seismic data for F3DT. We developed the toolkit adjoint_input.m which prepares the inputs for the Hercules adjoint simulation needed in the F3DT workflow (e.g., Tromp, 2015). Furthermore, this toolkit administrates several subroutines for seismogram decomposition and recombination:

(a) atom_dcmp.m systematically isolates seismic phases in a synthetic seismogram using the time-frequency spectrum computed with the S-transform. Figure 1 shows an example of this approach. The outputs from this subroutine are isolated waveforms that are later used as isolation filters. The isolation filter is the synthetic seismogram with the information we want to invert. Examples of measurements include the linear difference between waveforms, travel times, and frequency-dependent amplitude and phase. These options are included in the adjoint_input.m. During the past year, we have focused our efforts on developing a new technique for optimizing data functionals (Juarez & Jordan 2018a,b). Our results are included in gsdf.m.

(b) gsdf.m utilizes the output of atom_dcmp.m as isolation filter and estimates the differences between modelled and observed waveforms as generalized seismological data functionals (GSDF; Gee & Jordan, 1992), and the interference effect of non-isolated waveforms is modelled by the sum-of-wavelets representation. This subroutine also estimates the adjoint source time functions.

The toolkit adjoint_input.m allows the user to decide which formulation of F3DT is going to be performed; adjoint wavefield or scattering integral (Chen at al., 2007). Depending on this selection prepares the input files for the earthquake simulations in Hercules and stores the GSDF data and adjoint source time functions. Figure 2 shows an example of the GSDF calculation for the seismogram in Figure 1, and Figure 3 shows their interference factors.

With the completed set of tools, we have gathered most of the software required for performing an iteration of F3DT. However, the toolkit for administrating all the task in the workflow is still under development.

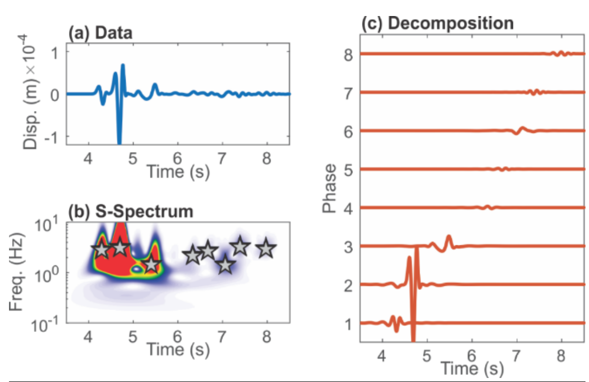

Figure 1. Seismogram decomposition using the S-transform. (a) The synthetic seismogram. (b) Frequency-time representation of the seismogram in (a) computed using the S-transform; grey stars show local maxima of the spectrum. (c) The decomposition algorithm identifies eight phases in the spectrum. Orange seismograms are the isolated phases in relative amplitude.

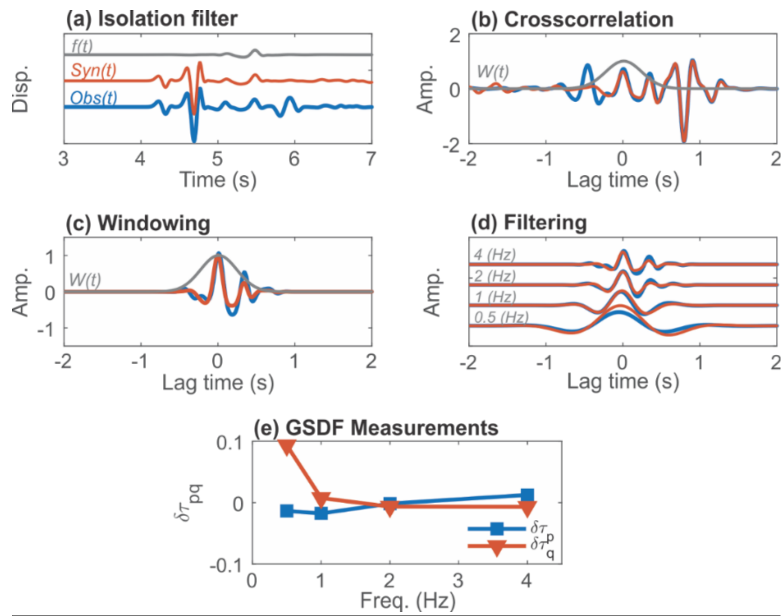

Figure 2. GSDF data processing: (a) Synthetic and observed seismograms and isolation filter. The isolation filter was obtained from the seismogram decomposition in Fig. 1. (b) Synthetic and data cross-correlograms. The cross-correlation function has many pulses at different lag-times. (c) Windowed cross-correlograms. (d) Filtered windowed cross-correlograms. These signals can be approximated with a five-parameter Gaussian wavelet. (e) GSDF measure the difference between one phase on both seismograms.

Key issue regarding project coordination within SCEC. The PC review of our proposal included the following comment: "One panelist noted that the version of Hercules mentioned in the proposal (for Task 2) is no longer being supported. A supported version of Hercules should be incorporated into the workflow." We agree that we need to migrate these F3DT toolkits into SCEC's Community Modeling Environment as part of a curated Hercules Toolchain, and we would readily do so if a SCEC-supported version of Hercules were available. The PI has discussed the issue with Phil Maechling. We both see an advantage in SCEC's curation of multiple anelastic wave propagation codes, but we recognize that this will require a strategic commitment by the SCEC leadership. The PI therefore requests further guidance regarding how our Hercules-based developments can best contribute to the SCEC CME objectives.

References

Chen, P., Jordan, T.H. & Zhao, L. (2007). Full three-dimensional tomography: a comparison between the scattering-integral and the adjoint-wavefield methods. *Geophys. J. Int.*, **170**, 175-181.

Ekström, G., Nettles, M. and Dziewoński, A.M., 2012. The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200, pp.1-9.

Gee, L.S. & Jordan, T.H. (1992). Generalized seismological data functional. *Geophys. J. Int.*, **111**, 363-390.

Goldstein, P., A. Snoke, (2005), "SAC Availability for the IRIS Community", *Incorporated Institutions for Seismology Data Management Center Electronic Newsletter.*

Goldstein, P., D. Dodge, M. Firpo, Lee Minner (2003) "SAC2000: Signal processing and analysis tools for seismologists and engineers, Invited contribution to "The IASPEI International Handbook of Earthquake and Engineering Seismology", Edited by WHK Lee, H. Kanamori, P.C. Jennings, and C. *Kisslinger, Academic Press*, London.

Juarez, A., & Jordan, T. H. (2018a). Optimization of data functionals for full-3D tomography by seismogram decomposition and resynthesis, poster presentation at 2018 SSA Annual Meeting. Miami, Florida.

Juarez, A., & Jordan, T. H. (2018b). Optimization of data functionals for full-3D tomography, abstract S51A-07, presented at 2018 Fall Meeting, AGU, Washington, D.C., Dec. 2019.

Juarez, A., & Jordan, T. H. (2019). Representation of the 2016 Mw 7.8 Kaikoura earthquake by orthogonal moment tensor fields, oral presentation at 2019 SSA Annual Meeting, Seattle, Washington, Apr. 2019.

Taborda R. and Bielak J. (2011). Large-scale earthquake simulation – Computational seismology and complex engineering systems, *Computing in Science and Engineering*, **13**, 14-26.

Tape, C., Q. Liu, A. Maggi, and J. Tromp (2010). Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.* **180**, 433-462.

Tromp, J., Tape, C. & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.*, **160**, 195-216.

Tu, T. et al. (2006). From Mesh Generation to Scientific Visualization: An End-To-End Approach to Parallel Supercomputing. *SC 2006 Conference, Proceedings of the ACM/IEEE,* pp.1-14.

Zhao, L., Chen, P. and Jordan, T.H., 2006. Strain Green's tensors, reciprocity, and their applications to seismic source and structure studies. *Bulletin of the Seismological Society of America,* 96(5), pp.1753-1763.