Miscellaneous notes on multitaper measurements for adjoint tomography

Carl Tape

October 26, 2009

`This document: /home/carltape/latex/notes/tomo/multitaper_notes.pdf`

`Vala's notes: /home/carltape/latex/notes/tomo/multitaper_vala.pdf`

Contents

List of Figures

List of Tables

1 Introduction

A multitaper measurements uses a set of optimal tapers design to extract frequency-dependent measurements of traveltime and amplitude differences. Theory and examples of multitaper measurements can be found in *Zhou et al.* (2004, 2005); *Ekström et al.* (1997); *Laske and Masters* (1996); *Thomson* (1982), and also *Percival and Walden* (1993, p. 333–347). The objective of this Appendix is to state the multitaper misfit functions (both traveltime and amplitude), and then to derive the corresponding adjoint source, ala *Tromp et al.* (2005).

2 Misfit functions, measurements, and adjoint sources

The conventions for the measurement, the Fourier transform, and the transfer function used for multitaper measurements are related. First we specify the measurement convention (Figure 1), then we specify the Fourier convention, and these determine the convention for the transfer function (Section 3.2).

2.1 The misfit function and measurement convention

Our misfit functions consists of a set of discrete time windows covering the seismic dataset of S earthquake sources, with N_s time windows for source index s. The total number of measurements is

$$N = \sum_{s=1}^{S} N_s \ . \tag{1}$$

For traveltime and amplitude tomography we seek to minimize the objective functions

$$F_{\rm P}(\mathbf{m}) = \tfrac{1}{2} \sum_{s=1}^{S} \sum_{p=1}^{N_s} \frac{1}{F_{sp}} \int_{-\infty}^{\infty} W_{sp}(\omega) \left[\frac{\tau_{sp}^{\rm obs}(\omega) - \tau_{sp}(\omega, \mathbf{m})}{\sigma_{{\rm P}sp}(\omega)} \right]^2 {\rm d}\omega \ , \tag{2}$$

$$F_{\rm Q}(\mathbf{m}) = \tfrac{1}{2} \sum_{s=1}^{S} \sum_{p=1}^{N_s} \frac{1}{F_{sp}} \int_{-\infty}^{\infty} W_{sp}(\omega) \left[\frac{\ln A_{sp}^{\rm obs}(\omega) - \ln A_{sp}(\omega, \mathbf{m})}{\sigma_{{\rm Q}sp}(\omega)} \right]^2 {\rm d}\omega \ , \tag{3}$$

where p is the index for measurement window "pick," and P and Q are labels denoting measures of traveltime and amplitude, respectively. For example, $\tau_{sp}^{\rm obs}(\omega) - \tau_{sp}(\omega, \mathbf{m})$ represents the frequency-dependent traveltime difference between synthetics and data for one time-windowed waveform on a single seismogram for source s. The frequency-dependent uncertainty associated with the traveltime measurement is estimated by $\sigma_{{\rm P}sp}(\omega)$. The function $W_{sp}(\omega)$ denotes a windowing filter whose width corresponds to the frequency range over which the measurements are assumed

reliable. A normalization factor G_{sp} is defined as[1]

$$G_{sp} = \int_{-\infty}^{\infty} W_{sp}(\omega)\,\mathrm{d}\omega. \tag{4}$$

We can incorporate the measurement uncertainty and normalization factor into the frequency filter by defining

$$W_{\mathrm{P}sp}(\omega) \equiv \frac{W_{sp}(\omega)}{G_{sp}\,\sigma^2_{\mathrm{P}sp}(\omega)}\,, \tag{5}$$

$$W_{\mathrm{Q}sp}(\omega) \equiv \frac{W_{sp}(\omega)}{G_{sp}\,\sigma^2_{\mathrm{Q}sp}(\omega)}\,, \tag{6}$$

and then the misfit functions (Eqs. 2 and 3) become

$$F_{\mathrm{P}}(\mathbf{m}) = \tfrac{1}{2}\sum_{s=1}^{S}\sum_{p=1}^{N_s}\int_{-\infty}^{\infty} W_{\mathrm{P}sp}(\omega)\left[\tau^{\mathrm{obs}}_{sp}(\omega) - \tau_{sp}(\omega,\mathbf{m})\right]^2 \mathrm{d}\omega\,, \tag{7}$$

$$F_{\mathrm{Q}}(\mathbf{m}) = \tfrac{1}{2}\sum_{s=1}^{S}\sum_{p=1}^{N_s}\int_{-\infty}^{\infty} W_{\mathrm{Q}sp}(\omega)\left[\ln A^{\mathrm{obs}}_{sp}(\omega) - \ln A_{sp}(\omega,\mathbf{m})\right]^2 \mathrm{d}\omega\,. \tag{8}$$

The variations of Equations (7) and (8) are given by

$$\delta F_{\mathrm{P}}(\mathbf{m}) = -\tfrac{1}{2}\sum_{s=1}^{S}\sum_{p=1}^{N_s}\int_{-\infty}^{\infty} W_{\mathrm{P}sp}(\omega)\,\Delta\tau_{sp}(\omega,\mathbf{m})\,\delta\tau_{sp}(\omega,\mathbf{m})\,\mathrm{d}\omega\,, \tag{9}$$

$$\delta F_{\mathrm{Q}}(\mathbf{m}) = -\tfrac{1}{2}\sum_{s=1}^{S}\sum_{p=1}^{N_s}\int_{-\infty}^{\infty} W_{\mathrm{Q}sp}(\omega)\,\Delta\ln A_{sp}(\omega,\mathbf{m})\,\delta\ln A_{sp}(\omega,\mathbf{m})\,\mathrm{d}\omega\,, \tag{10}$$

where

$$\Delta\tau_{sp}(\omega,\mathbf{m}) \equiv \tau^{\mathrm{obs}}_{sp}(\omega) - \tau_{sp}(\omega,\mathbf{m})\,, \tag{11}$$

$$\Delta\ln A_{sp}(\omega,\mathbf{m}) \equiv \ln A^{\mathrm{obs}}_{sp}(\omega) - \ln A_{sp}(\omega,\mathbf{m})\,, \tag{12}$$

are the measured traveltime difference and amplitude difference between data and synthetics, and $\delta\tau(\omega,\mathbf{m})$ and $\delta\ln A(\omega,\mathbf{m})$ are the traveltime and amplitude perturbations with respect to changes in the model parameters.

[1]In practice we define our measurement window according to positive angular frequencies only. Thus we can write Equation (4) as

$$G_{sp} = \int_{-\infty}^{\infty} W_{sp}(\omega)\,\mathrm{d}\omega = 2\int_{0}^{\infty} W_{sp}(\omega)\mathrm{d}\omega,$$

where $W_{sp}(\omega)$ is defined over positive frequencies only.

Sign convention for ΔT and $\Delta \ln A$

The conventions in Equations (11) and (12) are such that $\Delta\tau(\omega) > 0$ corresponds to a delay in the data, i.e., the data at frequency ω arrive late with respect to the synthetics at frequency ω (Figure 1). Similarly, $\Delta \ln A(\omega) > 0$ corresponds to an amplification of the data with respect to the synthetics, for frequency ω. These are the same conventions used in defining the cross-correlation measurements of (*Marquering et al.*, 1999; *Dahlen et al.*, 2000; *Dahlen and Baig*, 2002).[2]

2.1.1 Reduction for a single measurement ($N = 1$)

The notation is simpler if we consider a single event, a single receiver, a single component, and a single phase. In that case, the misfit functions are

$$F_{\rm P}(\mathbf{m}) = \tfrac{1}{2} \int_{-\infty}^{\infty} W_{\rm P}(\omega) \left[\tau^{\rm obs}(\omega) - \tau(\omega, \mathbf{m})\right]^2 {\rm d}\omega \,, \tag{13}$$

$$F_{\rm Q}(\mathbf{m}) = \tfrac{1}{2} \int_{-\infty}^{\infty} W_{\rm Q}(\omega) \left[\ln A^{\rm obs}(\omega) - \ln A(\omega, \mathbf{m})\right]^2 {\rm d}\omega \,, \tag{14}$$

where

$$W_{\rm P}(\omega) \equiv \frac{W(\omega)}{G\,\sigma_{\rm P}^2(\omega)} \,, \tag{15}$$

$$W_{\rm Q}(\omega) \equiv \frac{W(\omega)}{G\,\sigma_{\rm Q}^2(\omega)} \,. \tag{16}$$

The variations of Equations (13) and (14) are given by

$$\delta F_{\rm P}(\mathbf{m}) = -\int_{-\infty}^{\infty} W_{\rm P}(\omega)\, \Delta\tau(\omega, \mathbf{m})\, \delta\tau(\omega, \mathbf{m})\, {\rm d}\omega \,, \tag{17}$$

$$\delta F_{\rm Q}(\mathbf{m}) = -\int_{-\infty}^{\infty} W_{\rm Q}(\omega)\, \Delta \ln A(\omega, \mathbf{m})\, \delta \ln A(\omega, \mathbf{m})\, {\rm d}\omega \,, \tag{18}$$

where

$$\Delta\tau(\omega, \mathbf{m}) \equiv \tau^{\rm obs}(\omega) - \tau(\omega, \mathbf{m}) \,, \tag{19}$$

$$\Delta \ln A(\omega, \mathbf{m}) \equiv \ln A^{\rm obs}(\omega) - \ln A(\omega, \mathbf{m}) \,. \tag{20}$$

[2] *Tromp et al.* (2005) listed a sign convention for ΔT (their Eq. 39) that was opposite from the convention used in the formula for δT (their Eq. 41; directly from *Marquering et al.* (1999)). Of course, as long as the convention for δT is consistent with the convention for ΔT, everything is okay.

2.1.2 Reduction for frequency-independent measurements

For frequency-independent measurements (and uncertainties), we have

$$\Delta\tau(\omega) = \tau^{\rm obs} - \tau(\mathbf{m}) \,, \tag{21}$$

$$\Delta \ln A(\omega) = \ln A^{\rm obs} - \ln A(\mathbf{m}) \,, \tag{22}$$

$$\sigma_{\rm P}(\omega) = \sigma_{\rm P} \,, \tag{23}$$

$$\sigma_{\rm Q}(\omega) = \sigma_{\rm Q} \,. \tag{24}$$

The misfit functions (Eqs. 13 and 14) become

$$F_{\rm P}(\mathbf{m}) = \tfrac{1}{2}\Big[\tau^{\rm obs} - \tau(\mathbf{m})\Big]^2 \int_{-\infty}^{\infty} W_{\rm P}(\omega)\,{\rm d}\omega = \tfrac{1}{2}\left[\frac{\tau^{\rm obs} - \tau(\mathbf{m})}{\sigma_{\rm P}}\right]^2 \tag{25}$$

$$F_{\rm Q}(\mathbf{m}) = \tfrac{1}{2}\Big[\ln A^{\rm obs} - \ln A(\mathbf{m})\Big]^2 \int_{-\infty}^{\infty} W_{\rm Q}(\omega)\,{\rm d}\omega = \tfrac{1}{2}\left[\frac{\ln A^{\rm obs} - \ln A(\mathbf{m})}{\sigma_{\rm Q}}\right]^2 \tag{26}$$

which are the traveltime and amplitude cross-correlation misfit functions, $F_{\rm T}(\mathbf{m})$ and $F_{\rm A}(\mathbf{m})$, shown in *Tromp et al.* (2005), but also including measurement uncertainties.

The variations in Equations (25) and (26) are then

$$\delta F_{\rm P}(\mathbf{m}) = \left[-\frac{\Delta\tau(\mathbf{m})}{\sigma_{\rm P}^2}\right]\delta\tau(\mathbf{m}) \,, \tag{27}$$

$$\delta F_{\rm Q}(\mathbf{m}) = \left[-\frac{\Delta \ln A(\mathbf{m})}{\sigma_{\rm Q}^2}\right]\delta \ln A(\mathbf{m}) \,. \tag{28}$$

2.1.3 Uncertainty estimate (σ) based on cross-correlation measurements

Suppose we measure the traveltime anomaly based on the cross-correlation

$$\Gamma(\tau) = \int s(t-\tau)\, d(t)\,{\rm d}t, \tag{29}$$

where d denotes the observed seismogram and s the synthetic. Let δT denote the cross-correlation traveltime anomaly and $\delta \ln A$ the amplitude anomaly (Eq. 22) defined in *Maggi et al.* (2009, Eq. 11)[3]:

$$\delta \ln A = \ln\left(\frac{A^{\rm obs}}{A^{\rm syn}}\right) = 0.5\,\ln\left[\frac{\int d^2(t)\,{\rm d}t}{\int s^2(t)\,{\rm d}t}\right] . \tag{30}$$

We seek to determine $\sigma_{\rm T}$ and $\sigma_{\rm A}$ associated with δT and $\delta \ln A$. Therefore we write[4]

$$d(t) = [\exp(\delta \ln A) \pm \sigma_{\rm A}]\, s(t - \delta T \pm \sigma_{\rm T}) \,. \tag{31}$$

[3]See first-order version listed in *Dahlen and Baig* (2002).

[4]To first order, $\exp(\delta \ln A) \approx 1 + \delta \ln A$, but there is no need to make the first-order approximation.

Expanding the second term to first order, we obtain

$$\begin{aligned} d(t) &\approx [\exp(\delta \ln A) \pm \sigma_{\mathrm{A}}]\ [s(t-\delta T) \pm \sigma_{\mathrm{T}}\, \dot{s}(t-\delta T)] \\ &= [\exp(\delta \ln A) \pm \sigma_{\mathrm{A}}]\, s(t-\delta T) \pm \sigma_{\mathrm{T}}\ [\exp(\delta \ln A) \pm \sigma_{\mathrm{A}}]\, \dot{s}(t-\delta T) \\ &= \exp(\delta \ln A)\, s(t-\delta T) \pm \sigma_{\mathrm{A}}\, s(t-\delta T) \pm \sigma_{\mathrm{T}}\, \exp(\delta \ln A)\, \dot{s}(t-\delta T) \pm \sigma_{\mathrm{T}}\, \sigma_{\mathrm{A}}\, \dot{s}(t-\delta T). \end{aligned}$$

Thus, to first order in σ_{T} and σ_{A}, this may be written as

$$d(t) - \exp(\delta \ln A)\, s(t-\delta T) = \pm\sigma_{\mathrm{T}} \exp(\delta \ln A)\, \dot{s}(t-\delta T) \pm \sigma_{\mathrm{A}}\, s(t-\delta T). \tag{32}$$

Squaring (32) and integrating over time, we obtain

$$\begin{aligned} \int [d(t) - \exp(\delta \ln A)\, s(t-\delta T)]^2\, \mathrm{d}t &= \sigma_{\mathrm{T}}^2 \int [\exp(\delta \ln A)\, \dot{s}(t-\delta T)]^2\, \mathrm{d}t \\ &+ \sigma_{\mathrm{A}}^2 \int [s(t-\delta T)]^2\, \mathrm{d}t \\ &\pm \sigma_{\mathrm{T}}\, \sigma_{\mathrm{A}} \int [\exp(\delta \ln A)\, \dot{s}(t-\delta T)\, s(t-\delta T)]\, \mathrm{d}t. \end{aligned} \tag{33}$$

If we assume the errors are uncorrelated, we find that

$$\sigma_{\mathrm{T}}^2 = \frac{\int [d(t) - \exp(\delta \ln A)\, s(t-\delta T)]^2\, \mathrm{d}t}{\int [\exp(\delta \ln A)\, \dot{s}(t-\delta T)]^2\, \mathrm{d}t}\,, \tag{34}$$

$$\sigma_{\mathrm{A}}^2 = \frac{\int [d(t) - \exp(\delta \ln A)\, s(t-\delta T)]^2\, \mathrm{d}t}{\int [s(t-\delta T)]^2\, \mathrm{d}t}\,. \tag{35}$$

Because σ_{T} and σ_{A} appear in the denominator of the adjoint source, one must specify a nonzero water-level value for each. Otherwise, for a perfect cross-correlation measurement, $\sigma_{\mathrm{T}} = \sigma_{\mathrm{A}} = 0$, and the adjoint source (and therefore event kernel) will blow up. The water-level is an input parameter in `mt_measure_adj.f90`.

2.2 Transfer function

Each windowed pulse on an individual seismogram is characterized by a (complex) transfer function from the modeled synthetics to the observed data:

$$T(\omega)\, s(\omega) = d(\omega). \tag{36}$$

Note that the transfer function is the same, whether the data and synthetics are in displacement, velocity, or acceleration, etc:

$$\begin{aligned} T(\omega)\, i\omega s(\omega) &= i\omega d(\omega)\ , \\ -T(\omega)\, \omega^2 s(\omega) &= -\omega^2 d(\omega)\ . \end{aligned}$$

Here, the convention $ds/dt \leftrightarrow i\omega s(\omega)$ is consistent with the Fourier convention in Section 3.2.

We write the transfer function in general form in terms of frequency-dependent amplitude and traveltime differences, $\Delta \ln A(\omega)$ and $\Delta\tau(\omega)$:

$$T(\omega) \;=\; \exp[-i\omega\, \Delta\tau(\omega) + \Delta \ln A(\omega)] \;=\; \exp[-i\omega\, \Delta\tau(\omega)]\, \exp[\Delta \ln A(\omega)]\ . \tag{37}$$

Solving for $\Delta \ln A(\omega)$ and $\Delta\tau(\omega)$, we obtain

$$\Delta\tau(\omega) \;=\; -\frac{1}{\omega} \tan^{-1}\left(\frac{\mathrm{Im}\,[T(\omega)]}{\mathrm{Re}\,[T(\omega)]}\right), \tag{38}$$

$$\Delta \ln A(\omega) \;=\; \ln |T(\omega)|\ . \tag{39}$$

The sign convention in (37) is consistent with (19)–(20) and with the Fourier convention $\partial_t \leftrightarrow i\omega$ (Sections 3.2 and 3.3).

2.3 Multitaper measurements

Consider a single record of synthetics and data, $s(t)$ and $d(t)$, windowed in time over a particular phase and *both preprocessed in the same way*, e.g., filtered over a particular frequency window. The tapered versions are given by

$$s_j(t) \;=\; s(t)\, h_j(t)\ , \tag{40}$$

$$d_j(t) \;=\; d(t)\, h_j(t)\ , \tag{41}$$

where $h_j(t)$ is the taper.

Following *Laske and Masters* (1996), we use the multitaper method of *Thomson* (1982), which uses prolate spheroidal eigentapers (*Slepian*, 1978) for the $h_j(t)$. The transfer function $T(\omega)$ between the data and synthetics is given by (see Section 3.1)

$$T(\omega) \;=\; \frac{\sum_j d_j(\omega)\, s_j^*(\omega)}{\sum_j s_j(\omega)\, s_j^*(\omega)}\ . \tag{42}$$

This function may be computed directly from the data and synthetics, and then represented in terms of the real functions, $\Delta\tau(\omega)$ and $\Delta \ln A(\omega)$, by using Equations (38) and (39).

2.4 Multitaper adjoint sources

The units on various quantities in this section are shown in Table 1.

Following *Tromp et al.* (2005), we must express the misfit function variations in (17)–(18) in terms of the perturbed seismograms δs. The tapered data, $d_j(t)$, can be expressed as

$$d_j = d\, h_j = (s + \delta s)\, h_j = s\, h_j + \delta s\, h_j = s_j + \delta s_j, \tag{43}$$

where we have defined the tapered, perturbed synthetics as

$$\delta s_j = \delta s\, h_j. \tag{44}$$

Substituting (43) into (42), we obtain

$$T(\omega) \;=\; \frac{\sum_j \left[s_j(\omega) + \delta s_j(\omega)\right]\, s_j^*(\omega)}{\sum_j s_j(\omega)\, s_j^*(\omega)} \;=\; 1 \;+\; \frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\;. \tag{45}$$

The first-order approximation of (37) is given by

$$T(\omega) \;\approx\; 1 - i\omega\, \delta\tau(\omega) + \delta \ln A(\omega)\;, \tag{46}$$

and thus, from (45) and (46),

$$T(\omega) - 1 \;=\; \frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*} \;\approx\; -i\omega\, \delta\tau(\omega) + \delta \ln A(\omega)\;. \tag{47}$$

Following the same approach used to obtain (38)–(39), we solve (47) for $\delta\tau(\omega)$ and $\delta \ln A(\omega)$:

$$\delta\tau(\omega) \;=\; -\frac{1}{\omega}\, \mathrm{Im}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right), \tag{48}$$

$$\delta \ln A(\omega) \;=\; \mathrm{Re}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right). \tag{49}$$

Using the identities in Appendix A, we obtain

$$\begin{aligned}
\delta\tau(\omega) &= -\frac{1}{\omega}\,\mathrm{Im}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \\
&= \mathrm{Im}\left(\frac{-1}{\omega}\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \\
&= \mathrm{Re}\left(\frac{i}{\omega}\,\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \\
&= \mathrm{Re}\left(-\frac{i}{\omega}\,\frac{\sum_j s_j\, \delta s_j^*}{\sum_j s_j\, s_j^*}\right)\,, \qquad (50)\\
\delta \ln A(\omega) &= \mathrm{Re}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \\
&= \mathrm{Re}\left(\frac{\sum_j s_j\, \delta s_j^*}{\sum_j s_j\, s_j^*}\right)\,. \qquad (51)
\end{aligned}$$

Inserting these expressions for the transfer function into (17)–(18), and omitting the explicit dependence on $\mathbf{m}$, we have

$$\begin{aligned}
\delta F_{\mathrm{P}} &= -\int_{-\infty}^{\infty} W_{\mathrm{P}}(\omega)\,\Delta\tau(\omega)\,\mathrm{Re}\left(\frac{-i}{\omega}\,\frac{\sum_j s_j\,\delta s_j^*}{\sum_j s_j\, s_j^*}\right)\mathrm{d}\omega \\
&= \mathrm{Re}\left[\int_{-\infty}^{\infty} W_{\mathrm{P}}(\omega)\,\Delta\tau(\omega)\sum_j\left(\frac{i}{\omega}\frac{s_j}{\sum_k s_k s_k^*}\right)\delta s_j^*(\omega)\,\mathrm{d}\omega\right] \\
&= \mathrm{Re}\left[\int_{-\infty}^{\infty} W_{\mathrm{P}}(\omega)\,\Delta\tau(\omega)\sum_j p_j(\omega)\,\delta s_j^*(\omega)\,\mathrm{d}\omega\right], \qquad (52)
\end{aligned}$$

$$\begin{aligned}
\delta F_{\mathrm{Q}} &= -\int_{-\infty}^{\infty} W_{\mathrm{Q}}(\omega)\,\Delta\ln A(\omega)\,\mathrm{Re}\left(\frac{\sum_j s_j\,\delta s_j^*}{\sum_j s_j\, s_j^*}\right)\mathrm{d}\omega \\
&= \mathrm{Re}\left[\int_{-\infty}^{\infty} W_{\mathrm{Q}}(\omega)\,\Delta\ln A(\omega)\sum_j\left(\frac{-s_j}{\sum_k s_k s_k^*}\right)\delta s_j^*(\omega)\,\mathrm{d}\omega\right] \\
&= \mathrm{Re}\left[\int_{-\infty}^{\infty} W_{\mathrm{Q}}(\omega)\,\Delta\ln A(\omega)\sum_j q_j(\omega)\,\delta s_j^*(\omega)\,\mathrm{d}\omega\right]\,, \qquad (53)
\end{aligned}$$

where

$$\begin{aligned}
p_j(\omega) &\equiv \frac{i}{\omega}\frac{s_j}{\sum_k s_k s_k^*} = \frac{i\omega\, s_j}{\sum_k (i\omega s_k)(-i\omega s_k^*)} = \frac{i\omega\, s_j}{\sum_k (i\omega s_k)(i\omega s_k)^*}\,, \qquad (54)\\
q_j(\omega) &\equiv \frac{-s_j}{\sum_k s_k s_k^*} = i\omega\, p_j(\omega), \qquad (55)
\end{aligned}$$

where in (54) we have used the property (Eq. 108) $-iz^* = (iz)^*$. Note that p_j and q_j are based

on the (tapered) synthetics alone, and that, based on our Fourier convention in Section 3.2,

$$q_j(t) = \dot{p}_j(t). \tag{56}$$

Furthermore, note that the time-domain terms in (54) are all $\dot{s}_j(t)$, the derivative of the tapered synthetics.

We now use Plancherel's theorem (Section 3.5), one version of which is

$$\int_{-\infty}^{\infty} f(\omega)\, g^*(\omega)\, \mathrm{d}\omega = 2\pi \int_{-\infty}^{\infty} f(t)\, g^*(t)\, \mathrm{d}t, \tag{57}$$

to convert (52)–(53) into the time domain:

$$\begin{aligned}
\delta F_{\mathrm{P}} &= \mathrm{Re}\left[\int_{-\infty}^{\infty} W_{\mathrm{P}}(\omega)\, \Delta\tau(\omega) \sum_j p_j(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \\
&= \mathrm{Re}\left[\sum_j \int_{-\infty}^{\infty} \left[W_{\mathrm{P}}(\omega)\, \Delta\tau(\omega)\, p_j(\omega)\right] \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \\
&= \sum_j \int_{-\infty}^{\infty} 2\pi \left[W_{\mathrm{P}}(t) * \Delta\tau(t) * p_j(t)\right] \delta s_j(t)\, \mathrm{d}t \\
&= \sum_j \int_{-\infty}^{\infty} 2\pi \left[W_{\mathrm{P}}(t) * \Delta\tau(t) * p_j(t)\right] h_j(t)\, \delta s(t)\, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} \left\{2\pi \sum_j h_j(t) \left[W_{\mathrm{P}}(t) * \Delta\tau(t) * p_j(t)\right]\right\} \delta s(t)\, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} f_{\mathrm{P}}^{\dagger}(t)\ \delta s(t)\, \mathrm{d}t\ ,
\end{aligned} \tag{58}$$

$$\begin{aligned}
\delta F_{\mathrm{Q}} &= \mathrm{Re}\left[\int_{-\infty}^{\infty} W_{\mathrm{Q}}(\omega)\, \Delta \ln A(\omega) \sum_j q_j(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \\
&= \mathrm{Re}\left[\sum_j \int_{-\infty}^{\infty} \left[W_{\mathrm{Q}}(\omega)\, \Delta \ln A(\omega)\, q_j(\omega)\right] \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \\
&= \sum_j \int_{-\infty}^{\infty} 2\pi \left[W_{\mathrm{Q}}(t) * \Delta \ln A(t) * q_j(t)\right] \delta s_j(t)\, \mathrm{d}t \\
&= \sum_j \int_{-\infty}^{\infty} 2\pi \left[W_{\mathrm{Q}}(t) * \Delta \ln A(t) * q_j(t)\right] h_j(t)\, \delta s(t)\, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} \left\{2\pi \sum_j h_j(t) \left[W_{\mathrm{Q}}(t) * \Delta \ln A(t) * q_j(t)\right]\right\} \delta s(t)\, \mathrm{d}t \\
&= \int_{-\infty}^{\infty} f_{\mathrm{Q}}^{\dagger}(t)\ \delta s(t)\, \mathrm{d}t\ ,
\end{aligned} \tag{59}$$

where $\delta s_j(t) = h_j(t)\,\delta s(t)$ is the tapered, perturbed time series, $\Delta\tau(t)$ and $\Delta \ln A(t)$ are the time domain versions of (38)–(39), and we have defined our adjoint sources for multitaper traveltime measurements (P) and multitaper amplitude measurements (Q) as[5]

$$f^\dagger_{\rm P}(t) \equiv \sum_j h_j(t)\,[2\pi\ W_{\rm P}(t) * \Delta\tau(t) * p_j(t)]\ , \tag{60}$$

$$f^\dagger_{\rm Q}(t) \equiv \sum_j h_j(t)\,[2\pi\ W_{\rm Q}(t) * \Delta \ln A(t) * q_j(t)]\ . \tag{61}$$

The frequency domain versions of (60)–(61) are

$$f^\dagger_{\rm P}(\omega) = \sum_j h_j(\omega) * [2\pi\ W_{\rm P}(\omega)\,\Delta\tau(\omega)\,p_j(\omega)]\ , \tag{62}$$

$$f^\dagger_{\rm Q}(\omega) = \sum_j h_j(\omega) * [2\pi\ W_{\rm Q}(\omega)\,\Delta \ln A(\omega)\,q_j(\omega)]\ . \tag{63}$$

We also define the following functions:

$$P_j(t) \equiv 2\pi\ W_{\rm P}(t) * \Delta\tau(t) * p_j(t)\ , \tag{64}$$

$$Q_j(t) \equiv 2\pi\ W_{\rm Q}(t) * \Delta \ln A(t) * q_j(t)\ , \tag{65}$$

$$P_j(\omega) \equiv 2\pi\ W_{\rm P}(\omega)\,\Delta\tau(\omega)\,p_j(\omega)\ , \tag{66}$$

$$Q_j(\omega) \equiv 2\pi\ W_{\rm Q}(\omega)\,\Delta \ln A(\omega)\,q_j(\omega)\ . \tag{67}$$

These lead to succint expressions for the adjoint sources:

$$f^\dagger_{\rm P}(t) \equiv \sum_j h_j(t) P_j(t)\ , \tag{68}$$

$$f^\dagger_{\rm Q}(t) \equiv \sum_j h_j(t) Q_j(t)\ , \tag{69}$$

$$f^\dagger_{\rm P}(\omega) = \sum_j h_j(\omega) * P_j(\omega)\ , \tag{70}$$

$$f^\dagger_{\rm Q}(\omega) = \sum_j h_j(\omega) * Q_j(\omega)\ . \tag{71}$$

[5]Note that we have not written the time-dependence using the time-rerversal convention, i.e., $T - t$.

Table 1: Units for adjoint quantities for multitaper measurements. The misfit functions $F_{\rm P}(\mathbf{m})$ and $F_{\rm Q}(\mathbf{m})$ (and $F_{\rm T}(\mathbf{m})$ and $F_{\rm A}(\mathbf{m})$) are unitless if we take into account the units for the σ terms. In practice, the adjoint sources have a m^{-3} or m^{-2} quantity as well, due to the 3D or 2D volume for the delta function, $\delta(\mathbf{x})$, that is applied at the source location. Bottom two rows are for adjoint sources based on cross-correlation measurements.

frequency domain $[\widetilde{h}(\omega)] = [h(t)]$ s		time domain $[h(t)] = [\widetilde{h}(\omega)]$ s^{-1}	
$s(\omega)$, $d(\omega)$, $\delta s(\omega)$	m s	$s(t)$, $d(t)$, $\delta s(t)$	m
$\dot{s}(\omega)$	m	$\dot{s}(t)$	m s^{-1}
$W_{\rm P}(\omega)$	s^{-1}	$W_{\rm P}(t)$	s^{-2}
$W_{\rm Q}(\omega)$	s	$W_{\rm Q}(t)$	none
$\Delta\tau(\omega)$	s	$\Delta\tau(t)$	none
$\Delta \ln A(\omega)$	none	$\Delta \ln A(t)$	s^{-1}
$p_j(\omega)$	m^{-1}	$p_j(t)$	m^{-1} s^{-1}
$q_j(\omega)$	m^{-1} s^{-1}	$q_j(t)$	m^{-1} s^{-2}
$f^\dagger_{\rm P}(\omega)$	m^{-1}	$f^\dagger_{\rm P}(t)$	m^{-1} s^{-1}
$f^\dagger_{\rm Q}(\omega)$	m^{-1}	$f^\dagger_{\rm Q}(t)$	m^{-1} s^{-1}
$f^\dagger_{\rm T}(\omega)$	m^{-1}	$f^\dagger_{\rm T}(t)$	m^{-1} s^{-1}
$f^\dagger_{\rm A}(\omega)$	m^{-1}	$f^\dagger_{\rm A}(t)$	m^{-1} s^{-1}

2.5 A comparison with adjoint sources derived from cross-correlation measurements

In Section 2.1.2 we saw how the misfit functions $F_{\rm P}(\mathbf{m})$ and $F_{\rm Q}(\mathbf{m})$ reduce in the case of frequency-independent measurements. Here we show how the corresponding multitaper adjoint sources reduce in the case

$$\Delta\tau(\omega) = \Delta T \tag{72}$$

$$\Delta \ln A(\omega) = \Delta \ln A \tag{73}$$

$$\sigma_{\rm P}(\omega) = 1 \tag{74}$$

$$\sigma_{\rm Q}(\omega) = 1. \tag{75}$$

The multitaper adjoint sources can be determined from (60)–(61):

$$f^\dagger_{\rm P}(t) = \sum_j h_j(t)\,[W_{\rm P}(t) * \Delta T\,\delta(t) * p_j(t)] = \Delta T \sum_j h_j(t)\,[W_{\rm P}(t) * p_j(t)] \tag{76}$$

$$f^\dagger_{\rm Q}(t) = \sum_j h_j(t)\,[W_{\rm Q}(t) * \Delta \ln A\,\delta(t) * q_j(t)] = \Delta \ln A \sum_j h_j(t)\,[W_{\rm Q}(t) * q_j(t)]\,, \tag{77}$$

The misfit adjoint sources for cross-correlation measurements of traveltime (T) and displacement-record amplitude (Ad) are given by *Tromp et al.* (2005, Equations 45, 57, 67, and 77):

$$f^\dagger_{\mathrm{T}}(t) \;=\; -\Delta T \left[-\frac{\dot{s}(T-t)}{-\int \ddot{s}(t')\, s(t')\, \mathrm{d}t'} \right] = \Delta T \left[\frac{\dot{s}(T-t)}{-\int \ddot{s}(t')\, s(t')\, \mathrm{d}t'} \right] \tag{78}$$

$$f^\dagger_{\mathrm{A}d}(t) \;=\; \Delta \ln A \left[\frac{s(T-t)}{\int s(t')\, s(t')\, \mathrm{d}t'} \right]. \tag{79}$$

The misfit adjoint source for velocity-record amplitude (Av) is

$$f^\dagger_{\mathrm{A}v}(t) \;=\; \Delta \ln A \left[\frac{-\ddot{s}(T-t)}{\int \ddot{s}(t')\, s(t')\, \mathrm{d}t'} \right]. \tag{80}$$

It is not obvious which cross-correlation amplitude adjoint source should correspond to the multitaper amplitude adjoint source. MORE HERE.

By comparison of (78)–(79) with (76)–(77), we expect that for the case of frequency-independent measurements, $f^\dagger_{\mathrm{T}}(t) \approx f^\dagger_{\mathrm{P}}(t)$ and $f^\dagger_{\mathrm{A}}(t) \approx f^\dagger_{\mathrm{Q}}(t)$, that is,

$$\frac{\dot{s}(t)}{-\int \ddot{s}(t')\, s(t')\, \mathrm{d}t'} \;\approx\; \sum_j^{N_j} h_j(t)\, [W_{\mathrm{P}}(t) * p_j(t)] \tag{81}$$

$$\frac{s(t)}{\int s(t')\, s(t')\, \mathrm{d}t'} \;\approx\; \sum_j^{N_j} h_j(t)\, [W_{\mathrm{Q}}(t) * q_j(t)]\,. \tag{82}$$

The multitaper traveltime adjoint source contains the time series $p_j(t)$, and the multitaper amplitude adjoint source contains $q_j(t)$. Writing the expressions for (54) and (55) in the time domain, we find:

$$p_j(t) \;=\; \mathcal{F}^{-1} \left[\frac{-i\omega\, s_j(\omega)}{\sum_k [-\omega^2 s_k(\omega)]\, s_k^*(\omega)} \right] \;=\; \dot{s}_j(t) * \mathcal{F}^{-1} \left[\frac{-1}{\sum_k [-\omega^2 s_k(\omega)]\, s_k^*(\omega)} \right] \tag{83}$$

$$q_j(t) \;=\; \mathcal{F}^{-1} \left[\frac{-s_j(\omega)}{\sum_k s_k(\omega)\, s_k^*(\omega)} \right] \;=\; s_j(t) * \mathcal{F}^{-1} \left[\frac{-1}{\sum_k s_k(\omega)\, s_k^*(\omega)} \right]. \tag{84}$$

These bear similarities to the cross-correlation adjoint sources in (78)–(79).

2.6 Implementation and examples

We compute (60) as follows:

$$\begin{aligned} p_j(\omega) &\rightarrow W_{\rm P}(\omega)\,\Delta\tau(\omega)\,p_j(\omega) \\ &\rightarrow W_{\rm P}(t) * \Delta\tau(t) * p_j(t) \\ &\rightarrow f^\dagger_{\rm P}(t) = \sum_j h_j(t)\,[W_{\rm P}(t) * \Delta\tau(t) * p_j(t)] \end{aligned}$$

In Figures 4–12 we show an example of the adjoint source construction for a single arrival pulse. The purpose of these figures is to show the time series that go into making an adjoint source, and also to compare with the adjoint sources derived from cross-correlation measurements (*Tromp et al.*, 2005).

Figure 4a shows the pre-processed[6] data and synthetic pulses for displacement. The data arrives early and is somewhat diminished in amplitude. Thus, using the measurement conventions in Equations (19) and (20) (Figure 1), we note that $\Delta\tau > 0$ and $\Delta \ln A > 0$. Finally, the green dashed curve is the time-domain version of $T(\omega)s(\omega)$; as expected (Section 3.1), this is very close to the data, $d(t)$.

Figure 5a is the same as Figure 4a. Figure 5b shows the displacement, velocity, and acceleration for the synthetic arrival. These are the waveforms that are used in constructing the cross-correlation adjoint sources, as shown in Figure 5c-d. The dashed curve in Figure 5c, $\bar{f}^\dagger_{\rm T}$, is the cross-correlation traveltime adjoint source for a banana-doughnut kernel, i.e., an adjoint source that is independent of the measurement. This signal is proportional to $-\dot{s}(t)/N$, where $N > 0$ is the normalization factor[7] and $\dot{s}(t)$ is the synthetic velocity in Figure 5b. Thus, the shape of $\bar{f}^\dagger_{\rm T}$ is flipped from $\dot{s}(t)$. The solid curve in Figure 5c, $f^\dagger_{\rm T}$, is the cross-correlation traveltime adjoint source for an event kernel, i.e., an adjoint source that is weighted by the measurement. This signal is given by

$$f^\dagger_{\rm T}(t) \propto -\Delta T\,\bar{f}^\dagger_{\rm T}(t) \propto -\Delta T \left(-\frac{\dot{s}(t)}{N}\right).$$

Because $\Delta T > 0$, $f^\dagger_{\rm T}$ is flipped compared to $\bar{f}^\dagger_{\rm T}$, and is the same shape as $\dot{s}(t)$.

The dashed curve in Figure 5d, $\bar{f}^\dagger_{\rm A}$, is the cross-correlation amplitude adjoint source for a banana-doughnut kernel. This signal is proportional to $s(t)/M$, where $M > 0$ is the normalization factor[8] and $s(t)$ is the synthetic displacement in Figure 5b. Thus, the shape of $\bar{f}^\dagger_{\rm A}$ is similar to $s(t)$. The solid curve in Figure 5d, $f^\dagger_{\rm A}$, is the cross-correlation amplitude adjoint source for an

[6] In practice, we apply a taper of the form $H(t) = 1 - \cos^{10} t$ to the adjoint sources in (60)–(61), to ensure that the endpoints are zero, prior to numerically computing the adjoint wavefield. This is not a crucial step in the process. Also, using real data, the observed and synthetic time series will be pre-filtered over a particular frequency band.

[7] This differs from the convention in *Tromp et al.* (2005), where $N < 0$.

[8] This is the same M as in *Tromp et al.* (2005).

event kernel, i.e., an adjoint source that is weighted by the measurement. This signal is given by

$$f_{\mathrm{A}}^{\dagger}(t) \propto -\Delta \ln A \bar{f}_{\mathrm{A}}^{\dagger}(t) \propto -\Delta \ln A \left(\frac{s(t)}{M} \right).$$

Because $\Delta \ln A > 0$, $f_{\mathrm{A}}^{\dagger}$ is flipped compared to $\bar{f}_{\mathrm{A}}^{\dagger}$ and $s(t)$.

Figure 5e-f shows the multitaper adjoint sources. Also plotted are the cross-correlation adjoint sources for comparison. The agreement is good.

Figure 6 show the real parts of transfer function between data and synthetics, $T(\omega)$ (Eq. 37). In particular, Figure 6a shows $\Delta\tau(\omega)$ (Eq. 38), and Figure 6e shows $\Delta \ln A(\omega)$ (Eq. 39). Note that both of these quantities are positive for these measurements, which is consistent with the convention discussed above (data are late and greater in amplitude). Other subplots in Figure 6 show the functions after applying a filter in the frequency domain.

Figures 7–10 show the constituent signals used in constructing the multitaper adjoint sources.

Figure 11 shows the final step in constructing the multitaper traveltime adjoint sources. The left column shows the time series $P_j(t)h_j(t)$ in the time domain, and the right column shows the frequency domain representation. The bottom left plot shows the adjoint source, which is constructing by a simple summation (Eq. 60).

Figure 12 is identical to Figure 11, only for the multitaper amplitude adjoint source.

3 Miscellaneous

3.1 Deriving the transfer function

The transfer function, $T(\omega) = T_r(\omega) + i\,T_i(\omega)$ between data, d, and synthetics, s, is found by minimizing the objective function

$$F\left[T(\omega)\right] \;=\; \tfrac{1}{2}\sum_j \left|d_j(\omega) - T(\omega)\,s_j(\omega)\right|^2 \;,$$

where the sum over j represents the multitapers of Section 2.3. Expanding this into real and imaginary parts, we have

$$\begin{aligned}
F(T) &= \tfrac{1}{2}\sum_j |d_j - T s_j|^2 \\
&= \tfrac{1}{2}\sum_j (d_j - T s_j)(d_j^* - T^* s_j^*) \\
&= \tfrac{1}{2}\sum_j \left[d_j - (T_r + i\,T_i)\,s_j\right]\left[d_j^* - (T_r - i\,T_i)\,s_j^*\right] \\
&= \tfrac{1}{2}\sum_j \left[d_j\,d_j^* - (T_r - i\,T_i)\,d_j\,s_j^* - (T_r + i\,T_i)\,d_j^*\,s_j + (T_r^2 + T_i^2)\,s_j\,s_j^*\right] \\
&= \tfrac{1}{2}\sum_j \left[d_j\,d_j^* - T_r\,d_j\,s_j^* + i\,T_i\,d_j\,s_j^* - T_r\,d_j^*\,s_j - i\,T_i\,d_j^*\,s_j + T_r^2\,s_j\,s_j^* + T_i^2\,s_j\,s_j^*\right] \\
&= \tfrac{1}{2}\sum_j \left[d_j\,d_j^* + T_r\left(-d_j\,s_j^* - d_j^*\,s_j\right) + T_i\left(i\,d_j\,s_j^* - i\,d_j^*\,s_j\right) + T_r^2\,s_j\,s_j^* + T_i^2\,s_j\,s_j^*\right] \;.
\end{aligned}$$

The derivatives with respect to the real and imaginary parts of the transfer function are given by

$$\begin{aligned}
\frac{\partial F}{\partial T_r} &= \tfrac{1}{2}\sum_j \left[-d_j\,s_j^* - d_j^*\,s_j + 2\,T_r\,s_j\,s_j^*\right] \;=\; -\tfrac{1}{2}\sum_j \left(d_j^*\,s_j + d_j\,s_j^* +\right) + T_r\sum_j s_j\,s_j^* \\
\frac{\partial F}{\partial T_i} &= \tfrac{1}{2}\sum_j \left[i\,d_j\,s_j^* - i\,d_j^*\,s_j + 2\,T_i\,s_j\,s_j^*\right] \;=\; -\frac{i}{2}\sum_j \left(d_j^*\,s_j - d_j\,s_j^*\right) + T_i\sum_j s_j\,s_j^* \;.
\end{aligned}$$

Setting each equation equal to zero and solving for T_r and T_i gives:

$$\begin{aligned}
T_r &= \frac{\tfrac{1}{2}\sum_j \left(d_j\,s_j^* + d_j^*\,s_j\right)}{\sum_j s_j\,s_j^*} \\
T_i &= \frac{\tfrac{i}{2}\sum_j \left(d_j^*\,s_j - d_j\,s_j^*\right)}{\sum_j s_j\,s_j^*} \;.
\end{aligned}$$

Thus, we obtain (42):

$$T = T_r + i\,T_i = \frac{\frac{1}{2}\sum_j \left(d_j\,s_j^* + d_j^*\,s_j\right)}{\sum_j s_j\,s_j^*} - \frac{\frac{1}{2}\sum_j \left(d_j^*\,s_j - d_j\,s_j^*\right)}{\sum_j s_j\,s_j^*} = \frac{\frac{1}{2}\sum_j \left(d_j\,s_j^* + d_j^*\,s_j - d_j^*\,s_j + d_j\,s_j^*\right)}{\sum_j s_j\,s_j^*}$$
$$= \frac{\sum_j d_j\,s_j^*}{\sum_j s_j\,s_j^*}\,.$$

3.2 Conventions for measurements, Fourier transform, and transfer function

We define our measurements according to the conventions in (19)–(20), such that a positive traveltime measurement, $\Delta\tau > 0$, corresponds to a delay in the data with respect to the synthetics.

We define forward and inverse Fourier transforms

$$\mathcal{F}\left[h(t)\right] = \widetilde{h}(\omega) = \int_{-\infty}^{\infty} h(t)\,e^{-i\omega t}\,\mathrm{d}t \tag{85}$$

$$\mathcal{F}^{-1}\left[\widetilde{h}(\omega)\right] = h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{h}(\omega)\,e^{i\omega t}\,\mathrm{d}\omega\,. \tag{86}$$

Note that this convention follows that of *Dahlen and Tromp* (1998, p. 109), and that the units conversion is

$$\left[\widetilde{h}(\omega)\right] = \left[h(t)\right]\ \mathrm{s},$$

which is reflected in Table 1.

The Fourier transform of $\dot{h}(t)$ can be determined using integration by parts,

$$\int u\,dv = [u\,v] - \int v\,du,$$

with $dv = \dot{h}(t)\,dt$, $u = e^{-i\omega t}$, $v = h(t)$, and $du = -i\omega\,e^{-i\omega t}dt$. Thus, we can write

$$\begin{aligned}\mathcal{F}\left[\dot{h}(t)\right] &= \int_{-\infty}^{\infty} \dot{h}(t)\,e^{-i\omega t}\,\mathrm{d}t \\ &= \left[h(t)\,e^{-i\omega t}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} h(t)\,(-i\omega)\,e^{-i\omega t}\,\mathrm{d}t \\ &= i\omega\int_{-\infty}^{\infty} h(t)\,e^{-i\omega t}\,\mathrm{d}t \\ &= i\omega\,\mathcal{F}\left[h(t)\right]\,.\end{aligned} \tag{87}$$

This process can be iterated for the nth derivative to yield

$$\mathcal{F}\left[h^{(n)}(t)\right] = (i\omega)^n\,\mathcal{F}\left[h(t)\right]\,. \tag{88}$$

Note that (87)–(88) will depend on the Fourier convention.

The transfer function is defined according to $T(\omega)\,s(\omega) = d(\omega)$ (Eq. 36). The convention for

the measurements and the Fourier convention imply that the transfer function is to be written as

$$T(\omega) = e^{-i\omega\Delta\tau},$$

where we have ignored the amplitude measurement and have assumed that $\Delta\tau$ is constant over all ω. Then, the data in the frequency and time domain are given by

$$d(\omega) \;=\; s(\omega)\, e^{-i\omega\Delta\tau} \tag{89}$$

$$d(t) \;=\; \mathcal{F}^{-1}\left[d(\omega)\right] \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega)\, e^{-i\omega\Delta\tau}\, e^{i\omega t}\,\mathrm{d}\omega \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega)\, e^{i\omega(t-\Delta\tau)}\,\mathrm{d}\omega \;=\; s(t-\Delta\tau). \tag{90}$$

For example, for data arriving early with $\Delta\tau = -3$, we have $d(t) = s(t+3)$, indicating that the data are advanced by 3 seconds with respect to the synthetics.

Fourier transform: figures and examples

In Figures 2 and 3 we illustrate the forward and inverse Fourier transform on the test function

$$f(t) \;=\; H(t)\, e^{-at}\, \sin\omega_0 t, \tag{91}$$

where $H(t)$ is the Heaviside function. The Fourier transform of (91) is given by (Harris and Stocker, 1998, p. 720)

$$\widetilde{h}(\omega) \;=\; \frac{\omega_0}{(i\omega + a)^2 + \omega_0^2} \qquad \text{for } \mathrm{Re}[a] > 0. \tag{92}$$

In Figure 2 we show the agreement between the analytical and numerical functions. In Figure 3 the agreement confirms that the multitaper code convention for differentiation is given by (87).

3.3 Implementation of conventions for measurement, Fourier, and transfer function

In practice, the original data are shifted by the cross-correlation measurement, ΔT, prior to making the multitaper measurement. In other words, $d(\omega) = d_0(\omega)\, e^{i\omega\,\Delta T}$, where d_0 are the original, unshifted data. This convention is checked as follows:

$$d(t) \;=\; \mathcal{F}^{-1}\left[d(\omega)\right] \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} d_0(\omega)\, e^{i\omega\Delta T}\, e^{i\omega t}\,\mathrm{d}\omega \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} d_0(\omega)\, e^{i\omega\,(t+\Delta T)}\,\mathrm{d}\omega \;=\; d_0(t+\Delta T)\;, \tag{93}$$

that is, $d_0(t) = d(t - \Delta T)$. For example, for (original) data arriving early with respect to the synthetics, with $\Delta T = -3$, then we have $d_0(t) = d(t+3)$, indicating that the original data are

advanced by 3 seconds with respect to the shifted data.

3.4 Applying ΔT and $\Delta \ln A$ prior to making multitaper measurements

In practice, prior to making the multitaper measurement, we apply the cross-correlation (CC) measurements ΔT and $\Delta \ln A$. The multitaper adjoint sources are derived from linearized versions of the transfer function, and the application of the CC measurements brings the data and synthetics closer together, diminishing the effect of the linearized approximation. We consider three operations:

$$T^{\rm cc}(\omega)\, s_0(\omega) = s(\omega), \tag{94}$$

$$T'(\omega)\, s(\omega) = d_0(\omega), \tag{95}$$

$$T(\omega)\, s_0(\omega) = T'(\omega)\, T^{\rm cc}(\omega)\, s_0(\omega) = d_0(\omega)\ . \tag{96}$$

Equation (94) applies CC measurements to the original synthetics. Equation (95) applies the MT transfer function to CC-reconstructed synthetics. Equation (96) combines Equations (94) and (95); $T(\omega)$ is the transfer function from the original synthetics to the original data.

Expressions for these transfer functions are

$$T^{\rm cc}(\omega) = \exp[\Delta \ln A]\, \exp[-i\omega \Delta T] \tag{97}$$

$$T'(\omega) = \exp[\Delta \ln A'(\omega)]\, \exp[-i\omega \Delta \tau'(\omega)] \tag{98}$$

$$\begin{aligned} T(\omega) &= T'(\omega)\, T^{\rm cc}(\omega) \\ &= \exp[\Delta \ln A + \Delta \ln A'(\omega)]\, \exp\{-i\omega[\Delta T + \Delta \tau'(\omega)]\} \end{aligned} \tag{99}$$

Thus the phase and amplitude anomalies between the synthetics and the original data are

$$\Delta \tau(\omega) = \Delta T + \Delta \tau'(\omega), \tag{100}$$

$$\Delta \ln A(\omega) = \Delta \ln A + \Delta \ln A'(\omega)\ . \tag{101}$$

Alternatively, one could apply the *negative* CC measurements to the original data, prior to making the MT measurements. In this case, the operations are

$$d_0(\omega)/T^{\rm cc}(\omega) = d(\omega), \tag{102}$$

$$T'(\omega)\, s_0(\omega) = d(\omega) \tag{103}$$

$$T(\omega)\, s_0(\omega) = T'(\omega)\, T^{\rm cc}(\omega)\, s_0(\omega) = d_0(\omega)\ . \tag{104}$$

Equation (102) deconstructs the data with the CC measurements. Equation (104) applies the transfer function between the original synthetics and the deconstructed data. Equation (104) combines Equations (102) and (104) and is identical to Equation (96). Note that the $T'(\omega)$ in Equation (95) transfers the reconstructed synthetics to the original data, whereas $T'(\omega)$ in

Equation (104) transfers the original synthetics into the deconstructed data.

We have presented two possibilities for applying the CC measurements, prior to making the MT measurements:

1. Apply ΔT and $\Delta \ln A$ to the synthetics (Eq. 94).

2. Apply $-\Delta T$ and $-\Delta \ln A$ to the data (Eq. 102).

The windows are picked using an automated procedure (*Maggi et al.*, 2009) that chooses window boundaries based on the synthetic seismograms. Therefore, in practice, we take the second option: keep the synthetics fixed and apply the CC measurements to the data.

Note that the notation in Section 2 is the "prime" notation, i.e., we assume that the transfer function is between data and synthetics after CC measurements have been applied: $T(\omega)\, s(\omega) = d(\omega)$.

3.5 Plancherel's theorem

Parseval's theorem applied to Fourier series is known as *Rayleigh's theorem* or *Plancherel's theorem*. The following derivation of Plancherel's Theorem is adapted from the Mathworld website.

The exact representation of the theorem depends on the Fourier convention. Using the Fourier conventions in Equations (85) and (86), we have

$$\begin{aligned}
f(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(\omega)\, e^{i\omega t}\, \mathrm{d}\omega \\
f^*(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}^*(\omega)\, e^{-i\omega t}\, \mathrm{d}\omega \\
\widetilde{f}(\omega) &= \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t}\, \mathrm{d}t \\
\widetilde{f}^*(\omega) &= \int_{-\infty}^{\infty} f^*(t)\, e^{i\omega t}\, \mathrm{d}t.
\end{aligned}$$

Consider the following derivation:

$$\begin{aligned}
\int_{-\infty}^{\infty} f(t)\, g^*(t)\, \mathrm{d}t &= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \widetilde{f}(\omega)\, e^{i\omega t}\, \mathrm{d}\omega\right] \left[\int_{-\infty}^{\infty} \widetilde{g}^*(\omega')\, e^{-i\omega' t}\, \mathrm{d}\omega'\right] \mathrm{d}t \\
&= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \widetilde{f}(\omega)\, \widetilde{g}^*(\omega')\, e^{it(\omega-\omega')} \mathrm{d}\omega\, \mathrm{d}\omega'\, \mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \widetilde{f}(\omega)\, \widetilde{g}^*(\omega') \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it(\omega-\omega')} \mathrm{d}t\right] \mathrm{d}\omega\, \mathrm{d}\omega' \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(\omega-\omega')\, \widetilde{f}(\omega)\, \widetilde{g}^*(\omega')\, \mathrm{d}\omega\, \mathrm{d}\omega' \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(\omega)\, \widetilde{g}^*(\omega)\, \mathrm{d}\omega.
\end{aligned}$$

If $g(t) = f(t)$ (and thus $g^*(t) = f^*(t)$), then we obtain Plancherel's Theorem,

$$\int_{-\infty}^{\infty} |f(t)|^2 \, \mathrm{d}t = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \widetilde{f}(\omega) \right|^2 \, \mathrm{d}\omega, \tag{105}$$

which states that the integral of the squared modulus of a function is equal to the integral of the squared modulus of its spectrum.

3.6 Units on adjoint quantities and kernels

Table 2 summarizes the units for quantities in this paper or in *Tromp et al.* (2005). Note that in the case of traveltime and amplitude tomography using cross-correlation (xcorr) measurements, the Fréchet derivative of the misfit function is obtained by integrating the sampling kernels over the volume and weighting them by the measurement, i.e., ΔT for traveltimes or ΔA for amplitudes. Thus for traveltimes (xcorr), we have: 3D sampling kernel (m^{-3} s) $\times$ volume (m^3) $\times$ measurement (s) = Fréchet (s^2). For amplitudes (xcorr), we have: 3D sampling kernel (m^{-3}) $\times$ volume (m^3) $\times$ measurement (none) = Fréchet (none). The same is true for the 2D expressions.

A Miscellaneous formulas

Consider two complex numbers,

$$\begin{aligned} z &= a + bi \\ w &= c + di \end{aligned}$$

$$\begin{aligned}
\mathrm{Re}\,[iz] &= \mathrm{Re}\,[i(a+bi)] = \mathrm{Re}\,[ai-b] = -b = \mathrm{Im}[-ai-b] = \mathrm{Im}[-z]. && (106)\\
\mathrm{Re}\left[i^{-1}z\right] &= \mathrm{Re}\,[-iz] = \mathrm{Im}[z]. && (107)\\
-iz^* &= -i(a-bi) = -b+ai = (-b+ai)^* = [i(a+bi)]^* = (iz)^*. && (108)\\
\mathrm{Re}\,[zw^*] &= \mathrm{Re}\,[(a+bi)(c-di)] \\
&= \mathrm{Re}\,[ac-adi+bci+bd] \\
&= ac+bd \\
&= \mathrm{Re}\,[ac+adi-bci+bd] \\
&= \mathrm{Re}\,[(a-bi)(c+di)] \\
&= \mathrm{Re}\,[z^*w] && (109)\\
\mathrm{Re}\,[izw^*] &= \mathrm{Re}\,[i(a+bi)(c-di)] \\
&= \mathrm{Re}\,[i\,(ac-adi+bci+bd)] \\
&= \mathrm{Re}\,[aci+ad-bc+bdi] \\
&= ad-bc \\
&= \mathrm{Re}\,[-aci+ad-bc-bdi] \\
&= \mathrm{Re}\,[-i(ac+adi-bci+bd)] \\
&= \mathrm{Re}\,[-i(a-bi)(c+di)] \\
&= \mathrm{Re}\,[-iz^*w] && (110)
\end{aligned}$$

B Expanding the transfer function using basis functions (OBSOLETE)

We now expand (50)–(51) in terms of the test functions $B_\alpha(\omega)$:

$$\delta\tau(\omega) \;=\; \sum_\alpha B_\alpha(\omega)\,\delta\tau_\alpha \;=\; \mathrm{Re}\left(\frac{1}{i\omega}\,\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \tag{111}$$

$$\delta \ln A(\omega) \;=\; \sum_\alpha B_\alpha(\omega)\,\delta \ln A_\alpha \;=\; \mathrm{Re}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right)\;. \tag{112}$$

Multiplying both sides by $B_\beta(\omega)$ and integrating, we get

$$\sum_\alpha H_{\beta\alpha}\;\delta\tau_\alpha \;=\; \int B_\beta(\omega)\;\mathrm{Re}\left(\frac{1}{i\omega}\,\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right)\mathrm{d}\omega \tag{113}$$

$$\sum_\alpha H_{\beta\alpha}\;\delta \ln A_\alpha \;=\; \int B_\beta(\omega)\;\mathrm{Re}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right)\mathrm{d}\omega\;, \tag{114}$$

where

$$H_{\beta\alpha} \;=\; \int B_\beta(\omega) B_\alpha(\omega)\,\mathrm{d}\omega\;. \tag{115}$$

Note that $H_{\beta\alpha}$ is the identity matrix, if $B_\alpha(\omega)$ are orthogonal. We then determine the expressions for the discrete measurements, similar to the steps used in obtaining (52)–(53):

$$\begin{aligned}
\delta\tau_\alpha &= \sum_\beta H^{-1}_{\beta\alpha} \int B_\beta(\omega)\, \mathrm{Re}\left(\frac{1}{i\omega}\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \mathrm{d}\omega \\
&= \int \sum_\beta H^{-1}_{\beta\alpha} B_\beta(\omega)\, \mathrm{Re}\left(\frac{-1}{i\omega}\frac{\sum_j s_j\, \delta s_j^*}{\sum_j s_j\, s_j^*}\right) \mathrm{d}\omega \\
&= \mathrm{Re}\left[\int \sum_j \sum_\beta H^{-1}_{\beta\alpha} B_\beta(\omega) \left(\frac{-1}{i\omega}\frac{s_j}{\sum_k s_k\, s_k^*}\right) \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \\
&= \mathrm{Re}\left[\int \sum_j p_j^\alpha(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \qquad (116) \\
\delta \ln A_\alpha &= \sum_\beta H^{-1}_{\beta\alpha} \int B_\beta(\omega)\, \mathrm{Re}\left(\frac{\sum_j \delta s_j\, s_j^*}{\sum_j s_j\, s_j^*}\right) \mathrm{d}\omega \\
&= \int \sum_\beta H^{-1}_{\beta\alpha} B_\beta(\omega)\, \mathrm{Re}\left(\frac{\sum_j s_j\, \delta s_j^*}{\sum_j s_j\, s_j^*}\right) \mathrm{d}\omega \\
&= \mathrm{Re}\left[\int \sum_j \sum_\beta H^{-1}_{\beta\alpha} B_\beta(\omega) \left(\frac{s_j}{\sum_k s_k\, s_k^*}\right) \delta s_j^*(\omega)\, \mathrm{d}\omega\right] \\
&= \mathrm{Re}\left[\int \sum_j q_j^\alpha(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega\right] . \qquad (117)
\end{aligned}$$

We now define

$$\begin{aligned}
p_j^\alpha(\omega) &= \sum_\beta H^{-1}_{\beta\alpha}\, B_\beta(\omega)\, p_j(\omega) \qquad (118) \\
q_j^\alpha(\omega) &= \sum_\beta H^{-1}_{\beta\alpha}\, B_\beta(\omega)\, q_j(\omega) \;=\; -i\omega\, p_j^\alpha(\omega)\,, \qquad (119)
\end{aligned}$$

where $p_j(\omega)$ and $q_j(\omega)$ are defined in (54)–(55).

Using Parseval's theorem with (116)–(119) and following similar steps to (58)–(59), we have

$$\delta\tau_\alpha = \mathrm{Re}\left[\sum_j \int p_j^\alpha(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega\right] = \sum_j \int p_j^\alpha(t)\, \delta s_j(t)\, \mathrm{d}t = \int \left[\sum_j p_j^\alpha(t) h_j(t)\right] \delta s(t)\, \mathrm{d}t \qquad (120)$$

$$\delta \ln A_\alpha = \mathrm{Re}\left[\sum_j \int q_j^\alpha(\omega)\, \delta s_j^*(\omega)\, \mathrm{d}\omega\right] = \sum_j \int q_j^\alpha(t)\, \delta s_j(t)\, \mathrm{d}t = \int \left[\sum_j q_j^\alpha(t) h_j(t)\right] \delta s(t)\, \mathrm{d}t \qquad (121)$$

where $\delta s_j(t) = h_j(t)\, \delta s(t)$ is the tapered, perturbed time series. From (120) and (121), we can

express the adjoint sources for traveltime (P) and amplitude (Q) as

$$f^\dagger_{\rm P}(t) \;=\; \sum_\alpha \delta\tau_\alpha \; \bar{f}^\dagger_{{\rm P}\alpha}(t) \tag{122}$$

$$f^\dagger_{\rm Q}(t) \;=\; \sum_\alpha \delta \ln A_\alpha \; \bar{f}^\dagger_{Q\alpha}(t) \;, \tag{123}$$

where

$$\bar{f}^\dagger_{{\rm P}\alpha}(t) \;=\; \sum_j p^\alpha_j(T-t)\; h_j(T-t) \tag{124}$$

$$\bar{f}^\dagger_{Q\alpha}(t) \;=\; \sum_j q^\alpha_j(T-t)\; h_j(T-t) \;=\; \sum_j \dot{p}^\alpha_j(T-t)\; h_j(T-t) \tag{125}$$

represent the unweighted adjoint sources for the test function $B_\alpha(\omega)$. (The bar-notation, as used in *Tromp et al.* (2005), indicates a quantity derived from the synthetics alone.) Note that the adjoint sources in (122)–(123) consist of sums over the the tapers (j) and sums over the test functions (α).

We express the multitaper traveltime adjoint source (122) in the frequency domain:

$$\begin{aligned}
f^\dagger_{\rm P}(\omega) \;&=\; \sum_\alpha \delta\tau_\alpha \; \bar{f}^\dagger_{{\rm P}\alpha}(\omega) \\
&=\; \sum_\alpha \delta\tau_\alpha \; \sum_j p^\alpha_j(\omega) * h_j(\omega) \\
&=\; \sum_\alpha \delta\tau_\alpha \; \sum_j \left[\sum_\beta H^{-1}_{\beta\alpha} \; B_\beta(\omega)\; p_j(\omega) \right] * h_j(\omega) \\
&=\; \left[\sum_\alpha \delta\tau_\alpha \; \sum_\beta H^{-1}_{\beta\alpha} \; B_\beta(\omega) \right] \sum_j p_j(\omega) * h_j(\omega) \\
&=\; \delta\tau(\omega) \sum_j p_j(\omega) * h_j(\omega) \\
&=\; \delta\tau(\omega) \; \bar{f}^\dagger_{{\rm P}\alpha}(\omega) \;,
\end{aligned} \tag{126}$$

which is the frequency domain version of (60). Similarly, for the multitaper amplitude adjoint source, we also derive (61). Note that there is no longer any dependence on α in these expressions, and that we have assumed that the $B_\alpha(\omega)$ are orthogonal.

B.1 Selecting the $B_\alpha(\omega)$

As shown in (113)–(114), we use the basis functions $B_\alpha(\omega)$ to represent the continuous (real) functions that comprise the transfer function $T(\omega)$. If the $B_\alpha(\omega)$ are orthogonal, then $H_{\alpha\beta} = \delta_{\alpha\beta}$ (115), and (118)–(119) reduce to

$$p_j^\alpha(\omega) \;=\; \sum_\beta H^{-1}_{\beta\alpha}\, B_\beta(\omega)\, p_j(\omega) \;=\; B_\alpha(\omega)\, p_j(\omega) \tag{127}$$

$$q_j^\alpha(\omega) \;=\; \sum_\beta H^{-1}_{\beta\alpha}\, B_\beta(\omega)\, q_j(\omega) \;=\; B_\alpha(\omega)\, q_j(\omega)\;. \tag{128}$$

Here we note some possible choices of $B_\alpha(\omega)$:

1. Just "picking off" discrete points on $\tau(\omega)$ and $\delta \ln A(\omega)$ is equivalent to using the function

$$B_\alpha^\Pi(\omega) \;=\; \frac{1}{\Omega}\,\Pi\left(\frac{\omega-\omega_\alpha}{\Omega}\right), \tag{129}$$

   with $\Omega = \Delta\omega$, where $\Pi(x)$ is the rectangle function, defined by

$$\Pi(x) \;=\; \begin{cases} 1 & |x| \le \frac{1}{2} \\ 0 & |x| > \frac{1}{2} \end{cases}\;. \tag{130}$$

   Note that $\lim_{\Omega\to 0} B_\alpha^\Pi(\omega) = \delta(\omega-\omega_\alpha)$.

2. By selecting the appropriate ω_α and $\Omega \ge \Delta\omega$, we can cover the entire frequency range with the orthogonal $B_\alpha^\Pi(\omega)$.

3. The Welch window function is defined as

$$B_\alpha^W(\omega) \;=\; 1 - 4\,W\left(\frac{\omega-\omega_\alpha}{\Omega}\right), \tag{131}$$

   where $W(x)$ is a truncated parabola, defined by

$$W(x) \;=\; \begin{cases} x^2 & |x| \le \frac{1}{2} \\ 0 & |x| > \frac{1}{2} \end{cases}\;. \tag{132}$$

   Note that overlapping Welch windows are not orthogonal.

4. B-splines. Note that overlapping B-splines are not orthogonal.

References

Dahlen, F. A., and A. M. Baig (2002), Fréchet kernels for body-wave amplitudes, *Geophys. J. Int.*, *150*, 440–446.

Dahlen, F. A., and J. Tromp (1998), *Theoretical Global Seismology*, Princeton U. Press, Princeton, New Jersey, USA.

Dahlen, F. A., S.-H. Hung, and G. Nolet (2000), Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.*, *141*, 157–174.

Ekström, G., J. Tromp, and E. W. F. Larson (1997), Measurements and global models of surface wave propagation, *J. Geophys. Res.*, *102*(B4), 8137–8157.

Laske, G., and G. Masters (1996), Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, *101*(B7), 16,059–16,075.

Maggi, A., C. Tape, M. Chen, D. Chao, and J. Tromp (2009), An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.*, *178*, 257–281.

Marquering, H., F. A. Dahlen, and G. Nolet (1999), Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox, *Geophys. J. Int.*, *137*, 805–815.

Percival, D., and A. Walden (1993), *Spectral Analysis for Physical Applications*, Cambridge U. Press, Cambridge, UK.

Slepian, D. (1978), Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V: The discrete case, *Bell Syst. Tech.*, *57*, 1371–1430.

Tape, C., Q. Liu, and J. Tromp (2007), Finite-frequency tomography using adjoint methods—Methodology and examples using membrane surface waves, *Geophys. J. Int.*, *168*, 1105–1129.

Thomson, D. J. (1982), Spectrum estimation and harmonic analysis, *IEEE Proc.*, *70*, 1055–1096.

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

Zhou, Y., F. A. Dahlen, and G. Nolet (2004), Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, *158*, 142–168.

Zhou, Y., F. A. Dahlen, G. Nolet, and G. Laske (2005), Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.*, *163*, 1087–1111.

Table 2: Units associated with adjoint quantities and sensitivity kernels. The top half corresponds to quantities used in generating 3D kernels; the bottom half is for 2D kernels. "Misfit" quantities (e.g., $\mathbf{f}^\dagger$) depend on the data; "banana-doughnut" (BD) quantities adopt the bar-notation (e.g., $\bar{\mathbf{f}}^\dagger$) and do not depend on the data (i.e., the measurement). The labels "xcorr" and "mt" refer to a cross-correlation measurement and a multitaper measurement, respectively. δS is the variation of a particular objective function with respect to a variation of model parameters. Most of the expressions for these quantities are presented in *Tromp et al.* (2005). The rows "*Tape et al.* (2007)" indicate the misfit function example used in that study, i.e., a traveltime cross-correlation misfit function that does *not* have the weighting according to the error in the measurement, σ.

Name	Misfit $[S]$, $[\delta S]$	Adjoint source $[\bar{\mathbf{f}}^\dagger]$, $[\mathbf{f}^\dagger]$	Adjoint wavefield $[\bar{\mathbf{s}}^\dagger]$, $[\mathbf{s}^\dagger]$	Kernels $[\bar{K}_p]$, $[K_p]$	Measurement
3D		$\mathrm{m}^{-4}\,\mathrm{s}^{-1}$ [S]	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}$ [S]	m^{-3} [S]	
waveform (misfit)	$\mathrm{m}^2\,\mathrm{s}$	m^{-2}	$\mathrm{kg}^{-1}\,\mathrm{m}\,\mathrm{s}^2$	$\mathrm{m}^{-1}\,\mathrm{s}$	—
traveltime (xcorr), misfit	none	$\mathrm{m}^{-4}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}$	m^{-3}	s
traveltime (mt), misfit	none	$\mathrm{m}^{-4}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}$	m^{-3}	s
traveltime (xcorr), BD	s	m^{-4}	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}^2$	$\mathrm{m}^{-3}\,\mathrm{s}$	NA
amplitude (xcorr), misfit	none	$\mathrm{m}^{-4}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}$	m^{-3}	none
amplitude (mt), misfit	none	$\mathrm{m}^{-4}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}$	m^{-3}	none
amplitude (xcorr), BD	none	$\mathrm{m}^{-4}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}$	m^{-3}	NA
Tape et al. (2007)	s^2	$\mathrm{m}^{-4}\,\mathrm{s}$	$\mathrm{kg}^{-1}\,\mathrm{m}^{-1}\,\mathrm{s}^3$	$\mathrm{m}^{-3}\,\mathrm{s}^2$	s
2D		$\mathrm{m}^{-3}\,\mathrm{s}^{-1}$ [S]	$\mathrm{kg}^{-1}\,\mathrm{s}$ [S]	m^{-2} [S]	
waveform (misfit)	$\mathrm{m}^2\,\mathrm{s}$	m^{-1}	$\mathrm{kg}^{-1}\,\mathrm{m}^2\,\mathrm{s}^2$	s	—
traveltime (xcorr), misfit	none	$\mathrm{m}^{-3}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{s}$	m^{-2}	s
traveltime (mt), misfit	none	$\mathrm{m}^{-3}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{s}$	m^{-2}	s
traveltime (xcorr), BD	s	m^{-3}	$\mathrm{kg}^{-1}\,\mathrm{s}^2$	$\mathrm{m}^{-2}\,\mathrm{s}$	NA
amplitude (xcorr), misfit	none	$\mathrm{m}^{-3}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{s}$	m^{-2}	none
amplitude (mt), misfit	none	$\mathrm{m}^{-3}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{s}$	m^{-2}	none
amplitude (xcorr), BD	none	$\mathrm{m}^{-3}\,\mathrm{s}^{-1}$	$\mathrm{kg}^{-1}\,\mathrm{s}$	m^{-2}	NA
Tape et al. (2007)	s^2	$\mathrm{m}^{-3}\,\mathrm{s}$	$\mathrm{kg}^{-1}\,\mathrm{s}^3$	$\mathrm{m}^{-2}\,\mathrm{s}^2$	s

Table 3: Examples of misfit functions, $S(\mathbf{m})$, and their variations, δS, for different tomographic approaches (*Tromp et al.*, 2005). The model vector is given by $\mathbf{m}$; the corresponding data quantities are denoted by the obs superscript. Each misfit function expression is for a single event, single receiver, single component, and single phase; the actual misfit function is a summation over events, receivers, components, and phases.

Tomographic Approach	Misfit function $S(\mathbf{m})$	Variation of misfit function δS	Measurement ΔM
Waveform	$\frac{1}{2}\int W(t)\ \left\|\mathbf{s}(\mathbf{x},t,\mathbf{m})-\mathbf{d}(\mathbf{x},t)\right\|^2\ dt$	$\int W(t)\,\mathbf{\Delta M}\cdot\delta\mathbf{s}(\mathbf{x},t,\mathbf{m})dt$	$\mathbf{s}(\mathbf{x},t,\mathbf{m})-\mathbf{d}(\mathbf{x},t)$
Traveltime (xcorr)	$\frac{1}{2}\left[\dfrac{T^{\rm obs}-T(\mathbf{m})}{\sigma_T}\right]^2$	$-\dfrac{\Delta M}{\sigma_T^2}\,\delta T$	$T^{\rm obs}-T(\mathbf{m})$
Amplitude (xcorr)	$\frac{1}{2}\left[\dfrac{\ln A^{\rm obs}-\ln A(\mathbf{m})}{\sigma_A}\right]^2$	$-\dfrac{\Delta M}{\sigma_A^2}\,\delta\ln A$	$\ln A^{\rm obs}-\ln A(\mathbf{m})$
Traveltime (multitaper)	$\frac{1}{2}\int W_P(\omega)\left[\tau^{\rm obs}(\omega)-\tau(\omega,\mathbf{m})\right]^2 d\omega$	$-\int W_P(\omega)\,\Delta M\,\delta\tau(\omega,\mathbf{m})\ d\omega$	$\tau^{\rm obs}(\omega)-\tau(\omega,\mathbf{m})$
Amplitude (multitaper)	$\frac{1}{2}\int W_Q(\omega)\left[\ln A^{\rm obs}(\omega)-\ln A(\omega,\mathbf{m})\right]^2 d\omega$	$-\int W_Q(\omega)\,\Delta M\,\delta\ln A(\omega,\mathbf{m})\ d\omega$	$\ln A^{\rm obs}(\omega)-\ln A(\omega,\mathbf{m})$
Differential TT (xcorr)	$\frac{1}{2}\left[\Delta T^{\rm obs}-\Delta T(\mathbf{m})\right]^2$	$-\Delta M\,\delta\Delta T$	$\Delta T^{\rm obs}-\Delta T(\mathbf{m})$

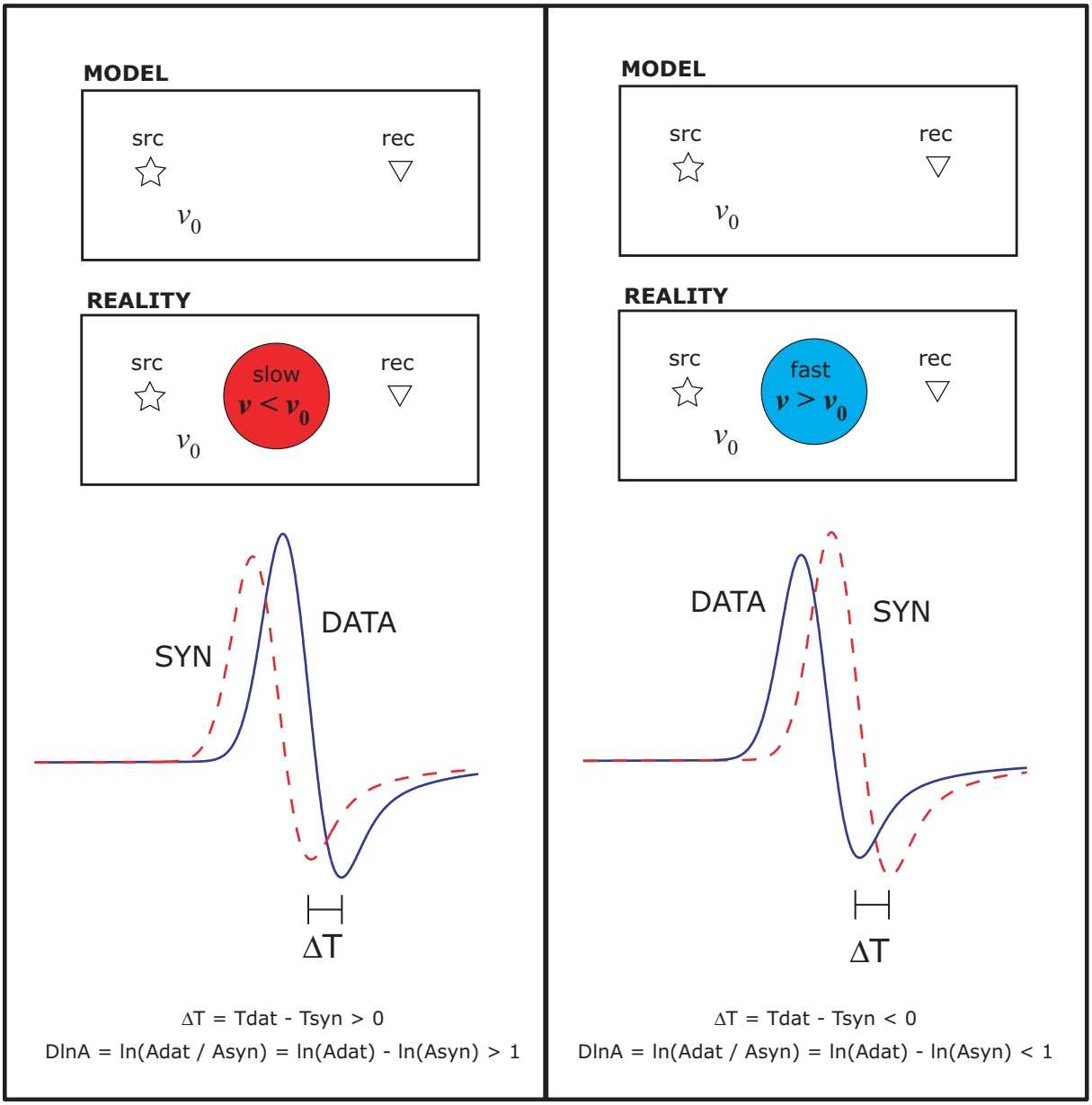

Figure 1: The measurement convention. See Section 2.1.

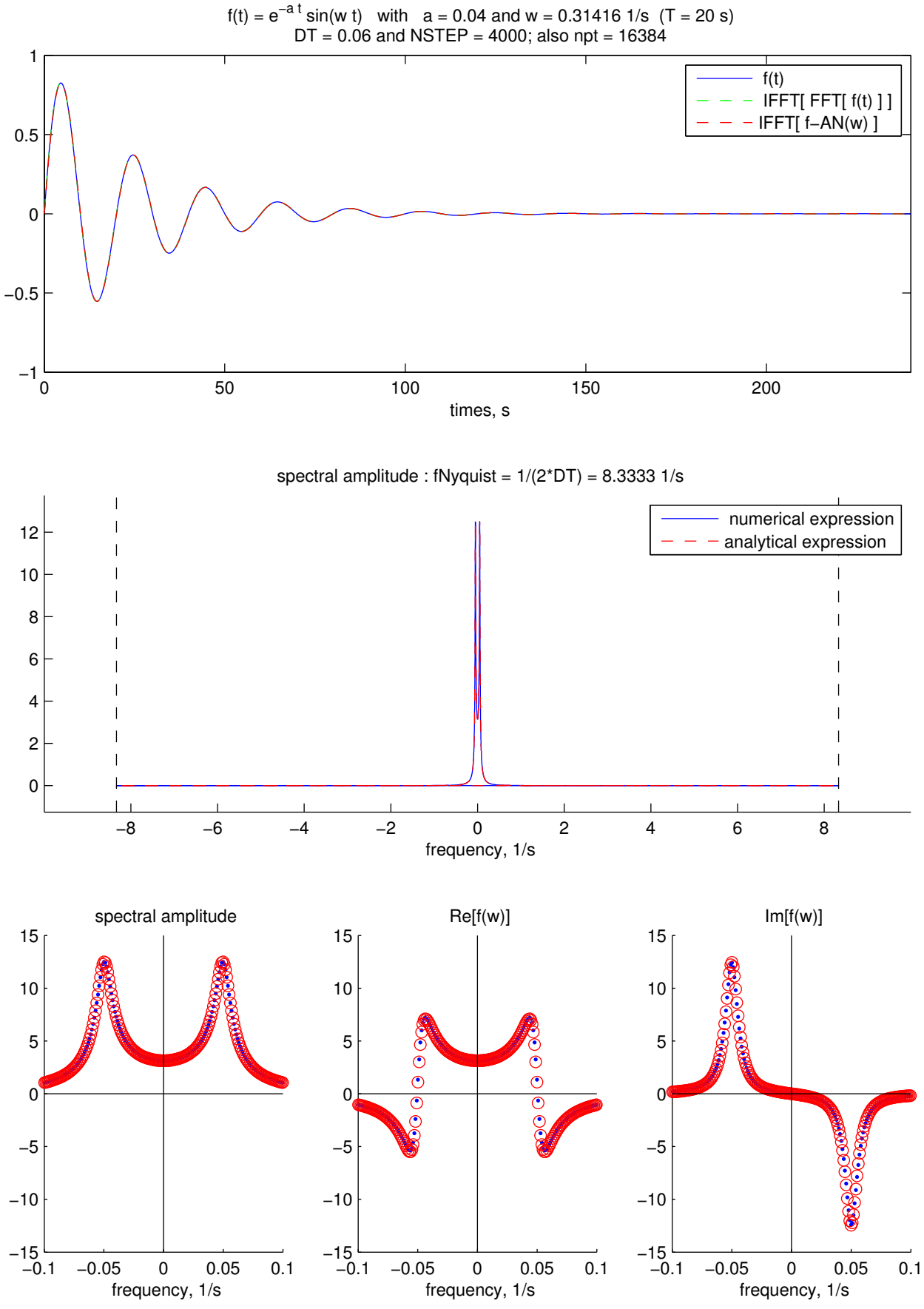

Figure 2: Example figure showing the results of the forward and inverse Fourier transform, using the subroutines in the multitaper code. See Section 3.2.

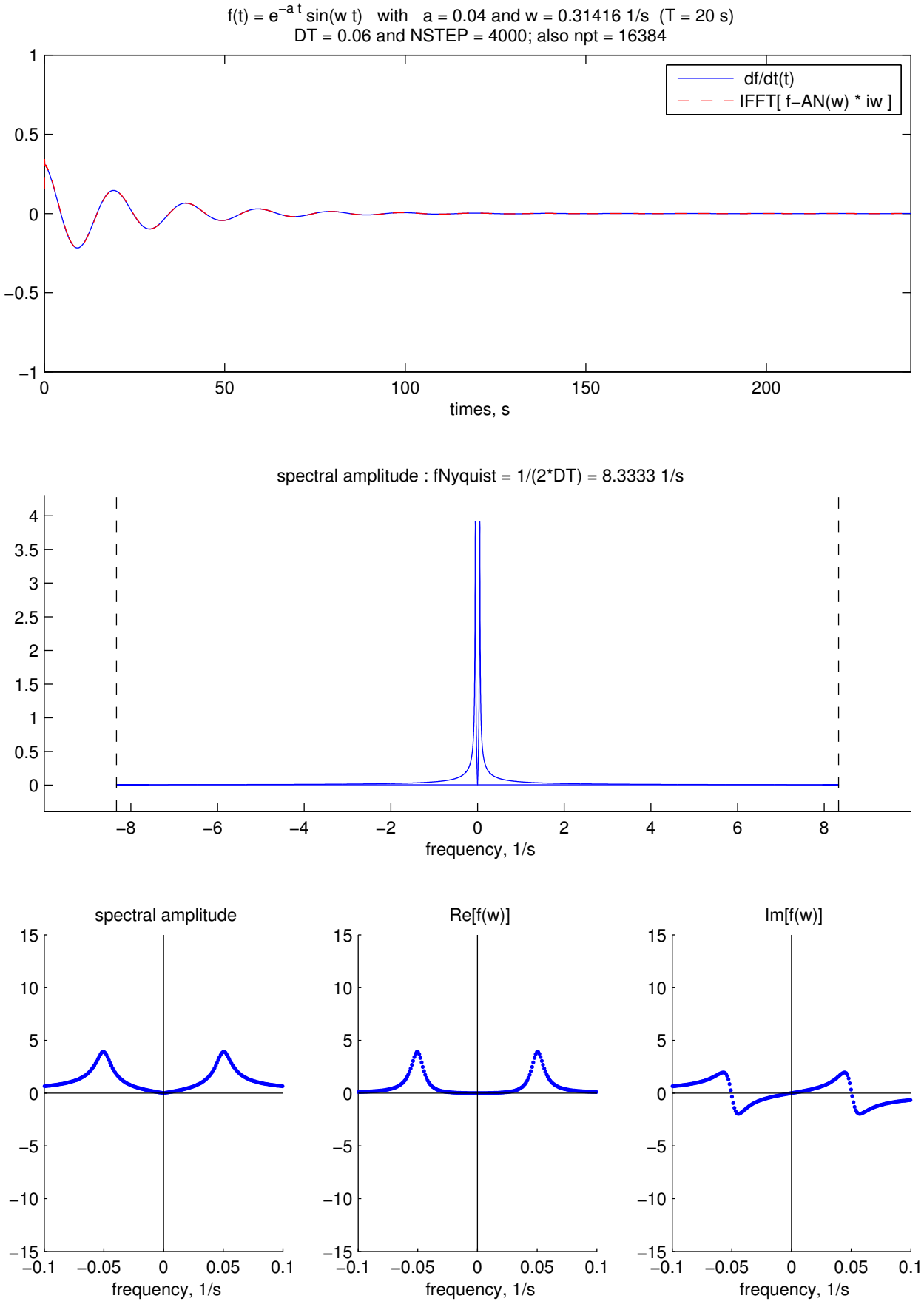

Figure 3: Example figure showing the effect of multiplying by $i\omega$ in the frequency domain. Using the Fourier convention in Section 3.2, this corresponds to differentiation in the time domain. See Section 3.2.

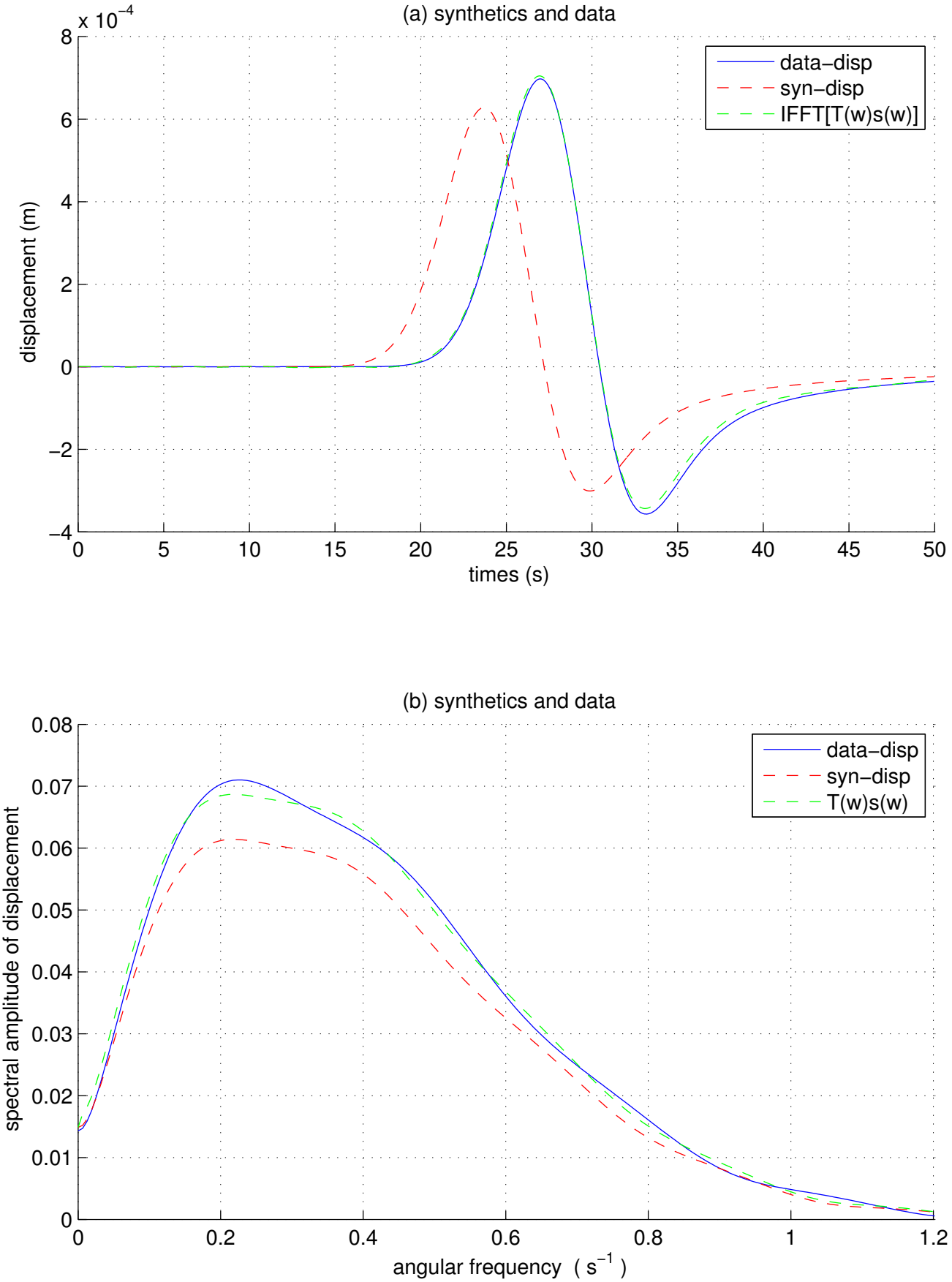

Figure 4: Pre-processed data and synthetics arrival, both in the time domain (top) and the frequency domain (bottom).

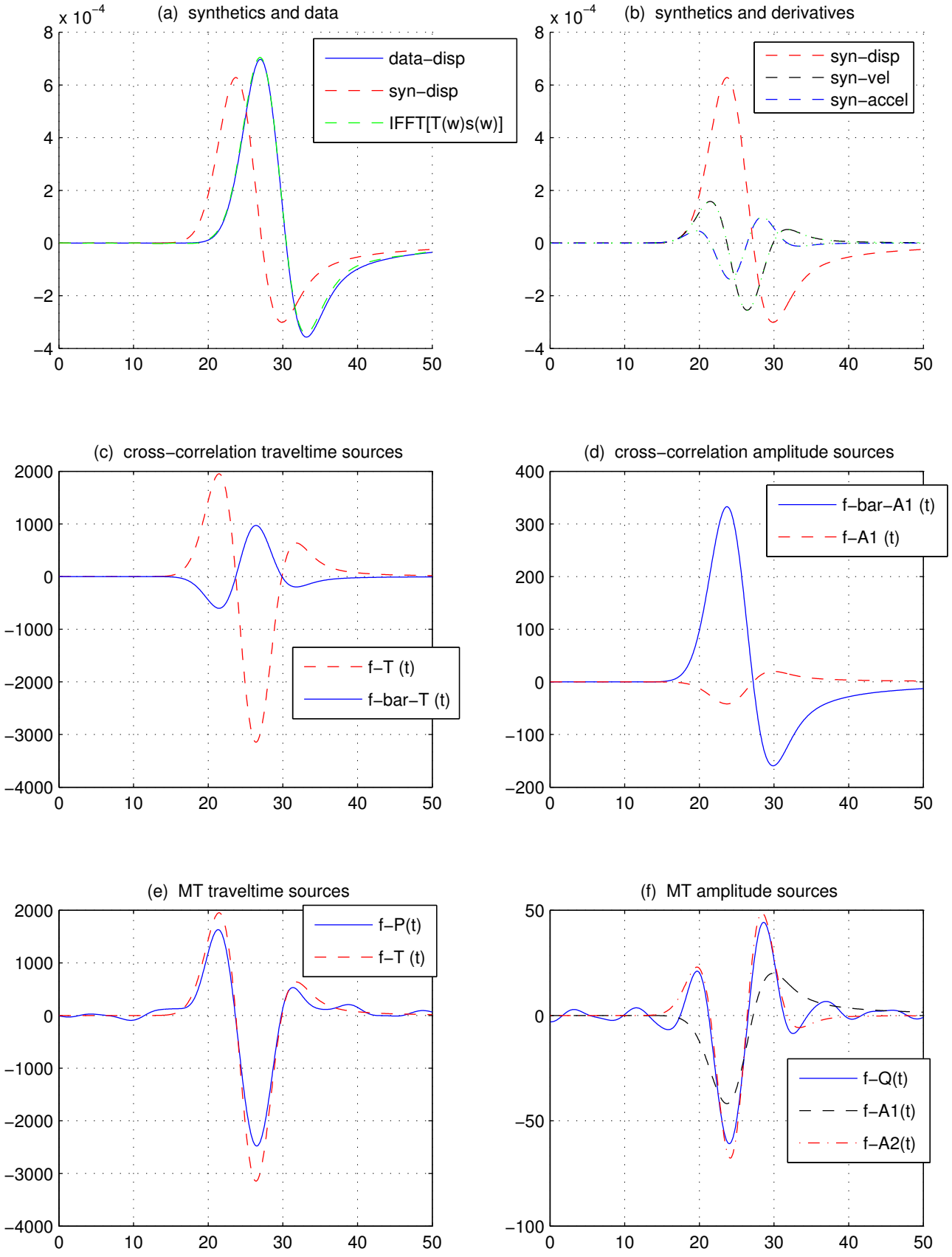

Figure 5: Pre-processed data and synthetics arrival, the construction of cross-correlation traveltime and amplitude adjoint sources (*Tromp et al.*, 2005), and the transfer function between data and synthetics. See Section 2.6.

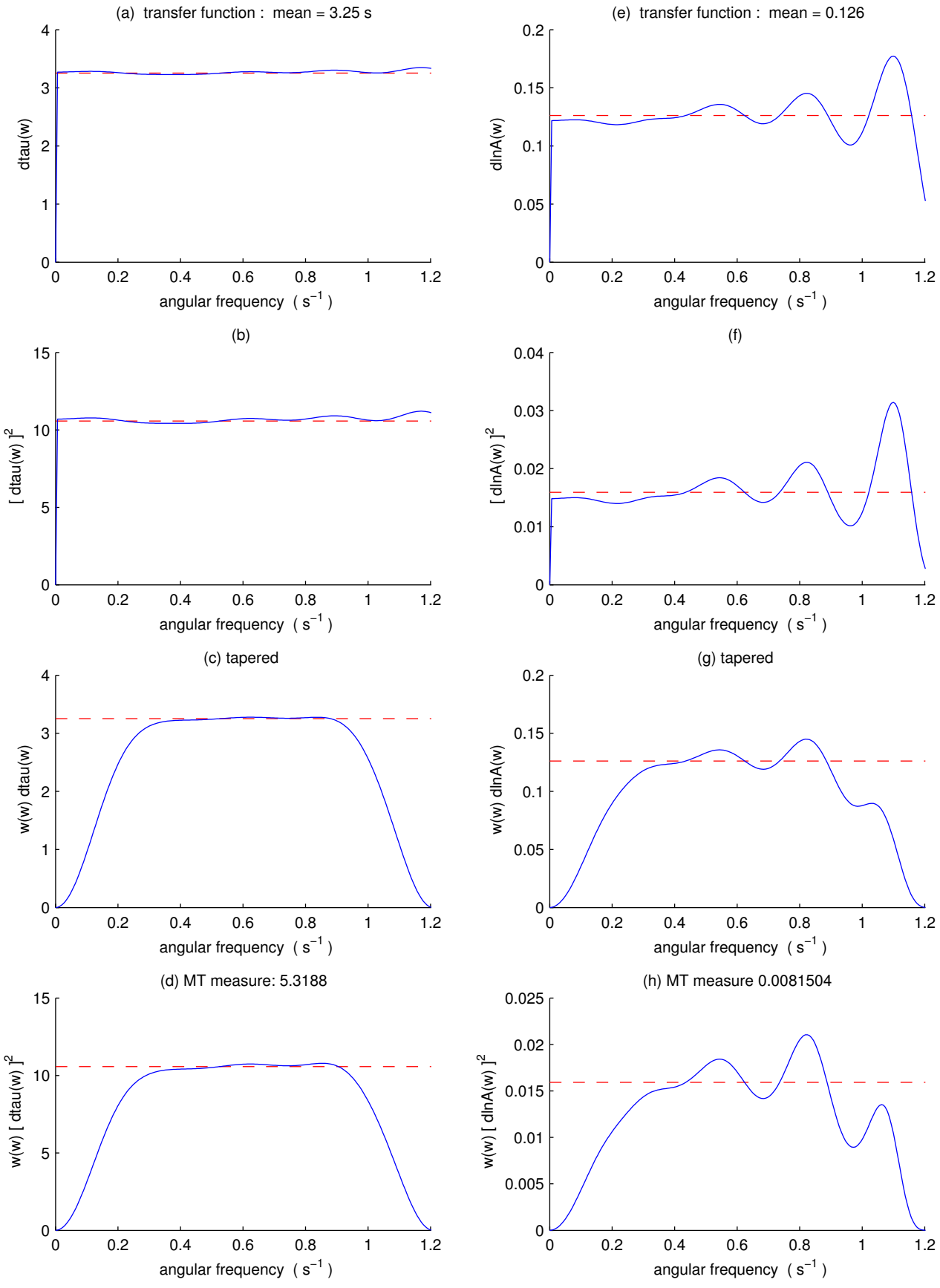

Figure 6: Transfer function and measurements.

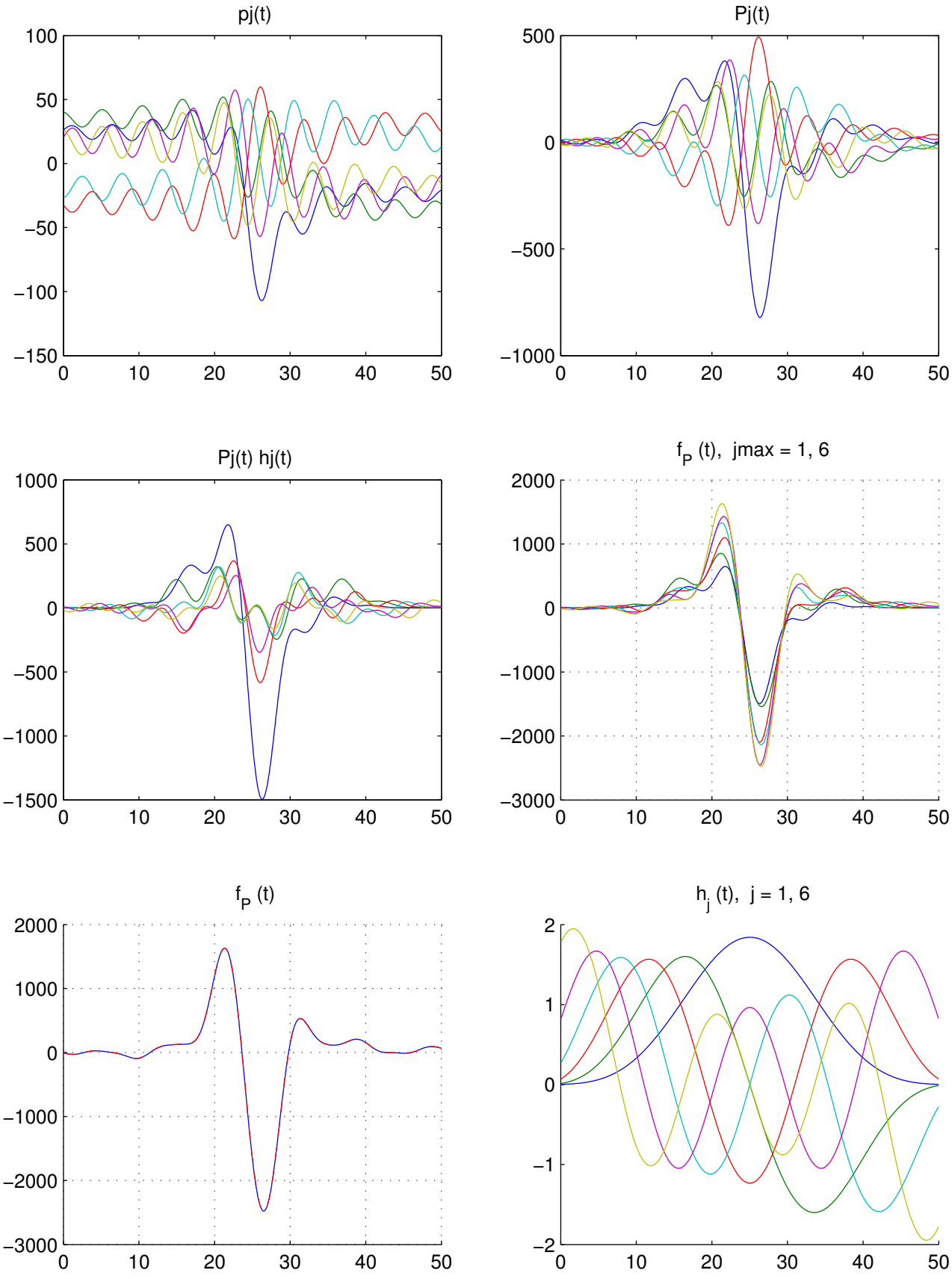

Figure 7: Time series used in constructing the multitaper traveltime adjoint source. Lower right shows the first five eigentapers, $h_j(t)$. See Section 2.6.

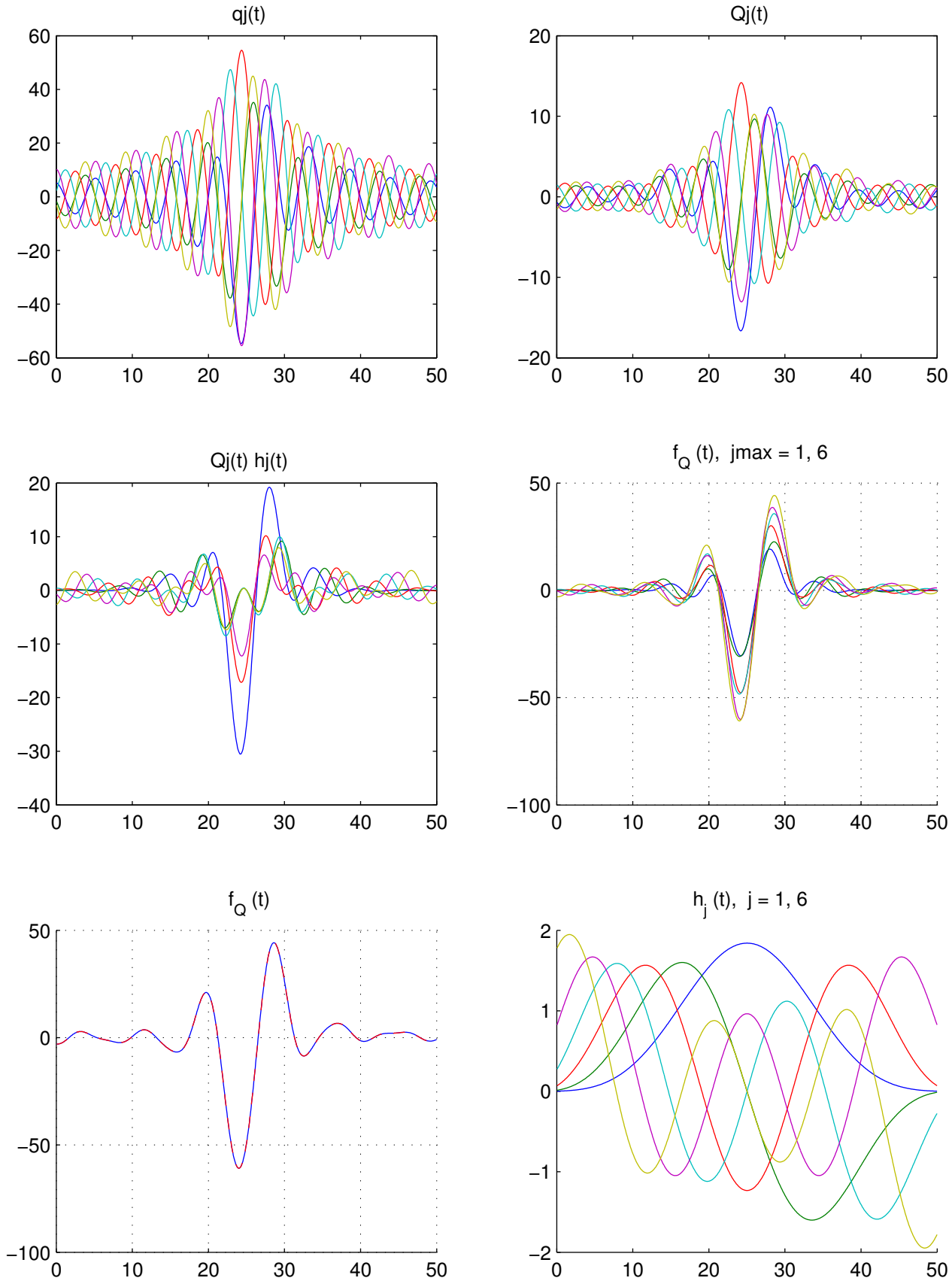

Figure 8: Time series used in constructing the multitaper amplitude adjoint source. Lower right shows the first five eigentapers, $h_j(t)$. See Section 2.6.

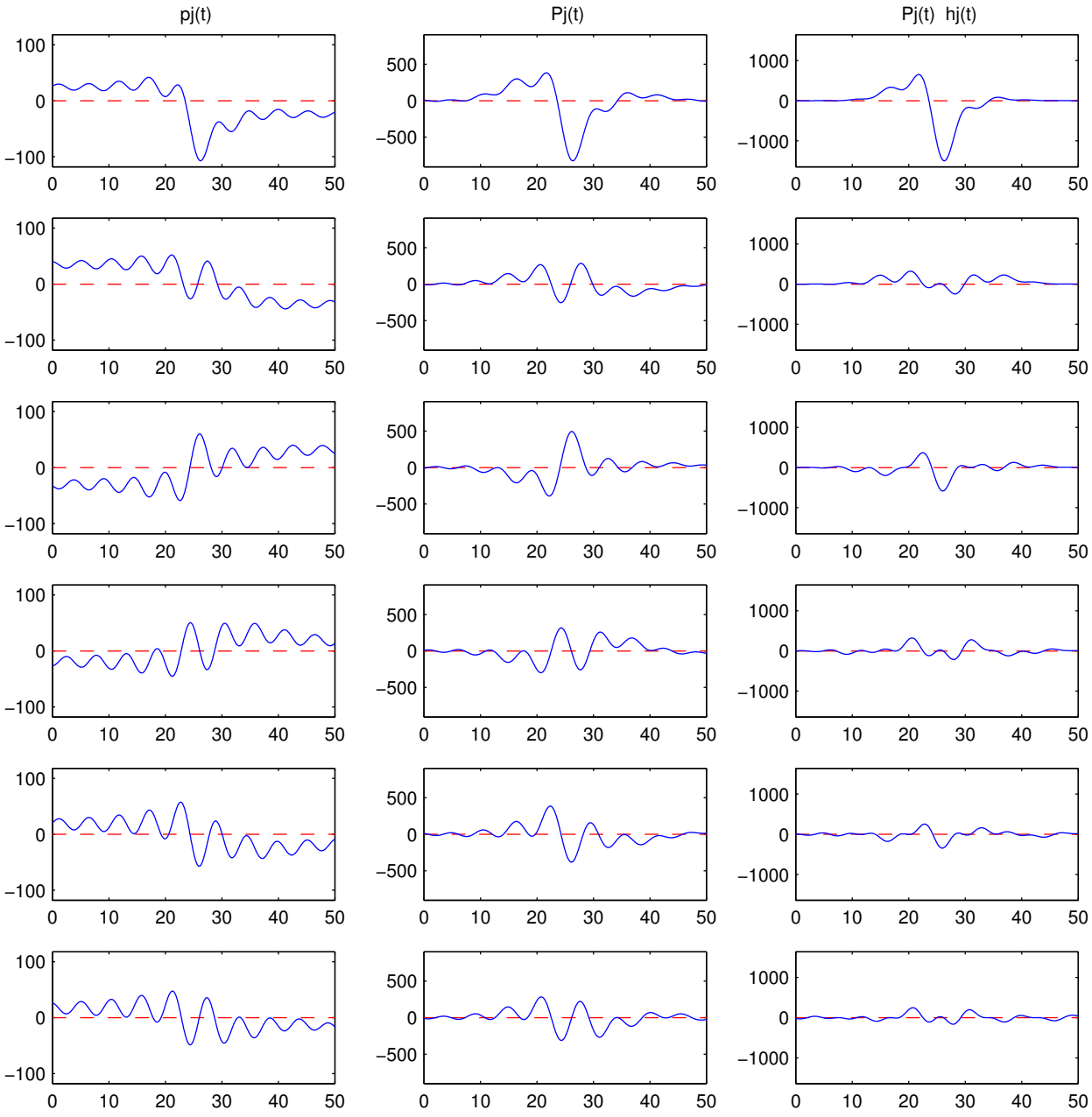

Figure 9: Time series used in constructing the multitaper traveltime adjoint source. See Section 2.6.

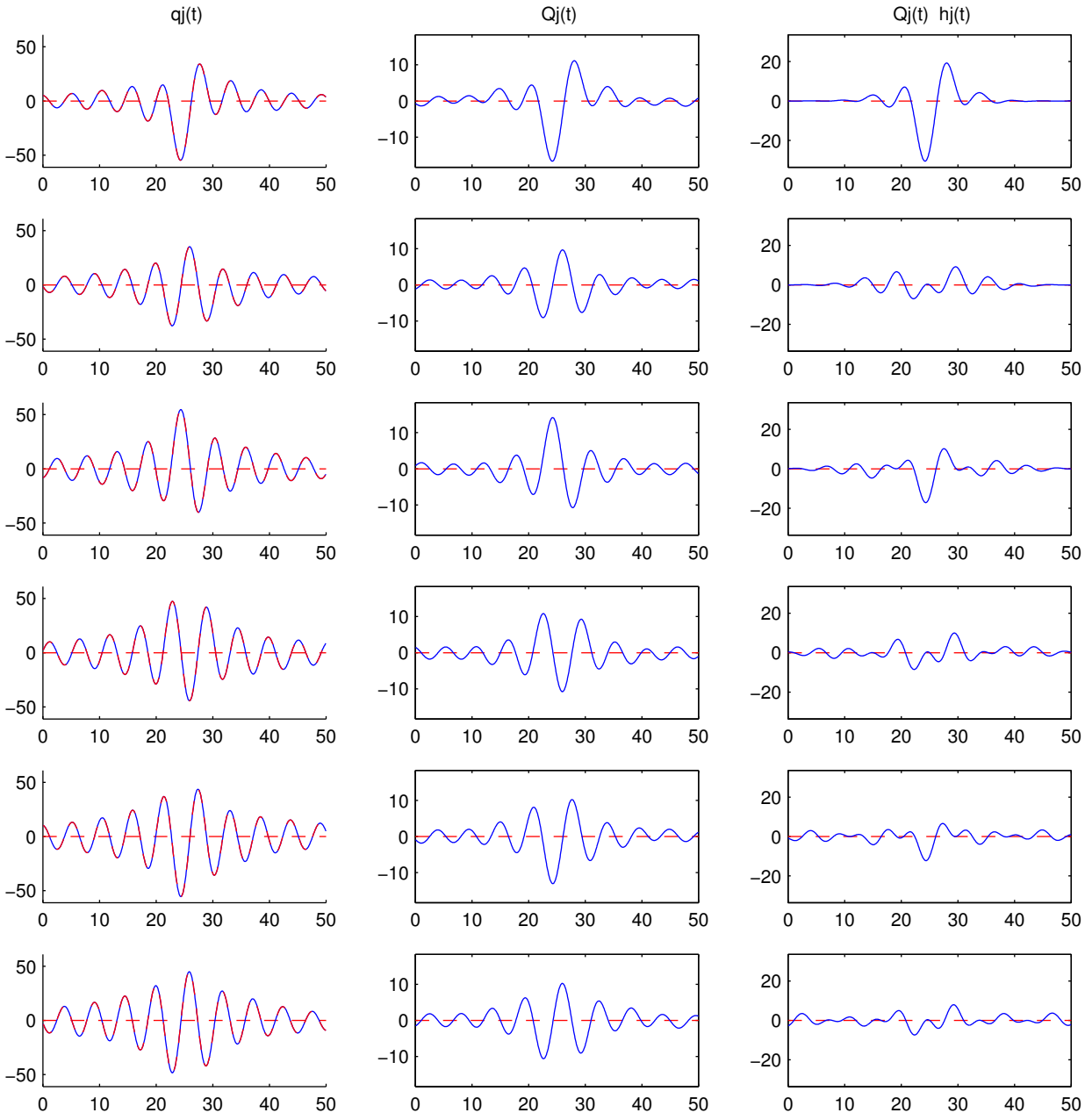

Figure 10: Time series used in constructing the multitaper amplitude adjoint source. See Section 2.6.

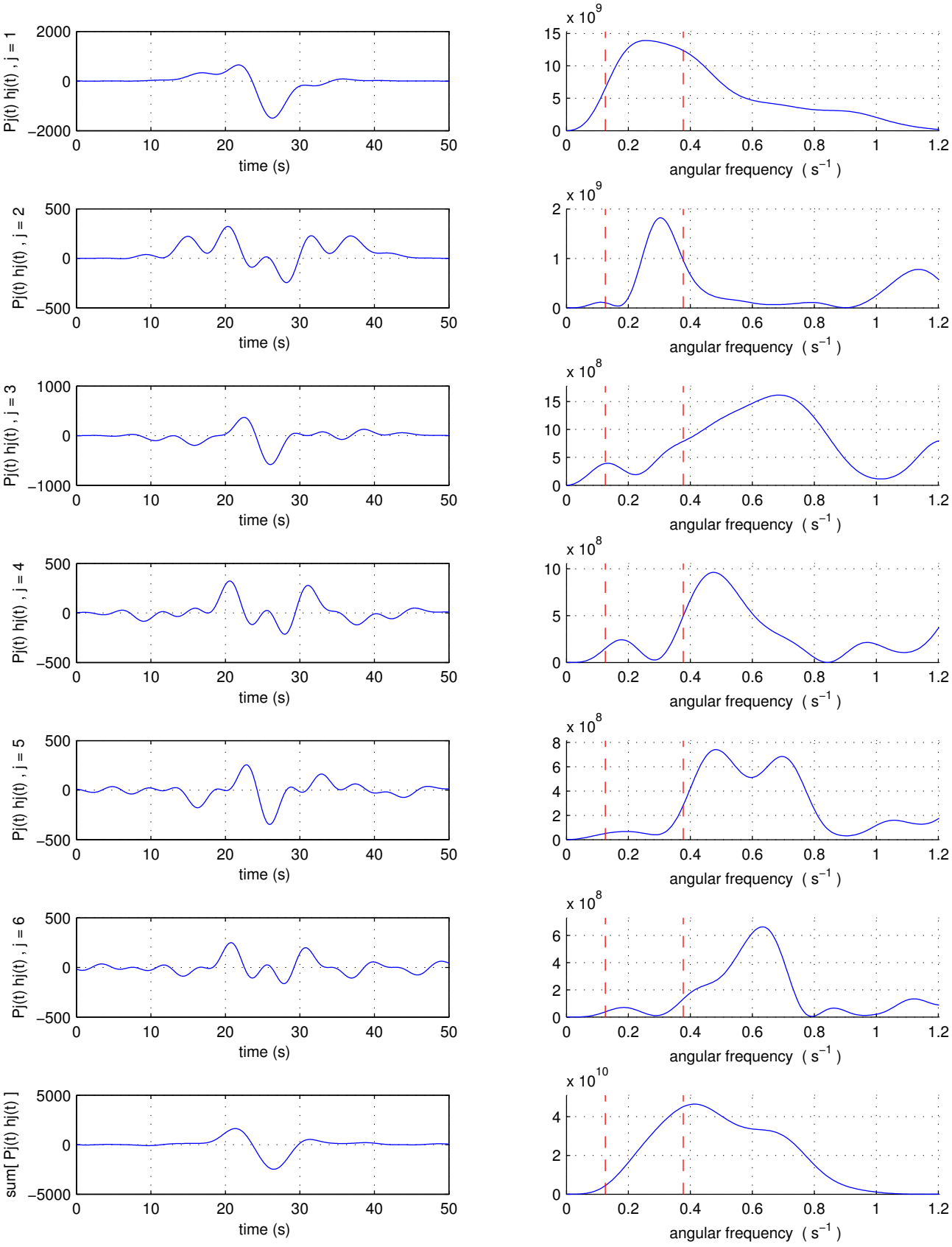

Figure 11: Construction of the multitaper traveltime adjoint source. Left column is time domain, right column is frequency domain. Bottom left shows the composite signal used for the adjoint source. See Section 2.6.

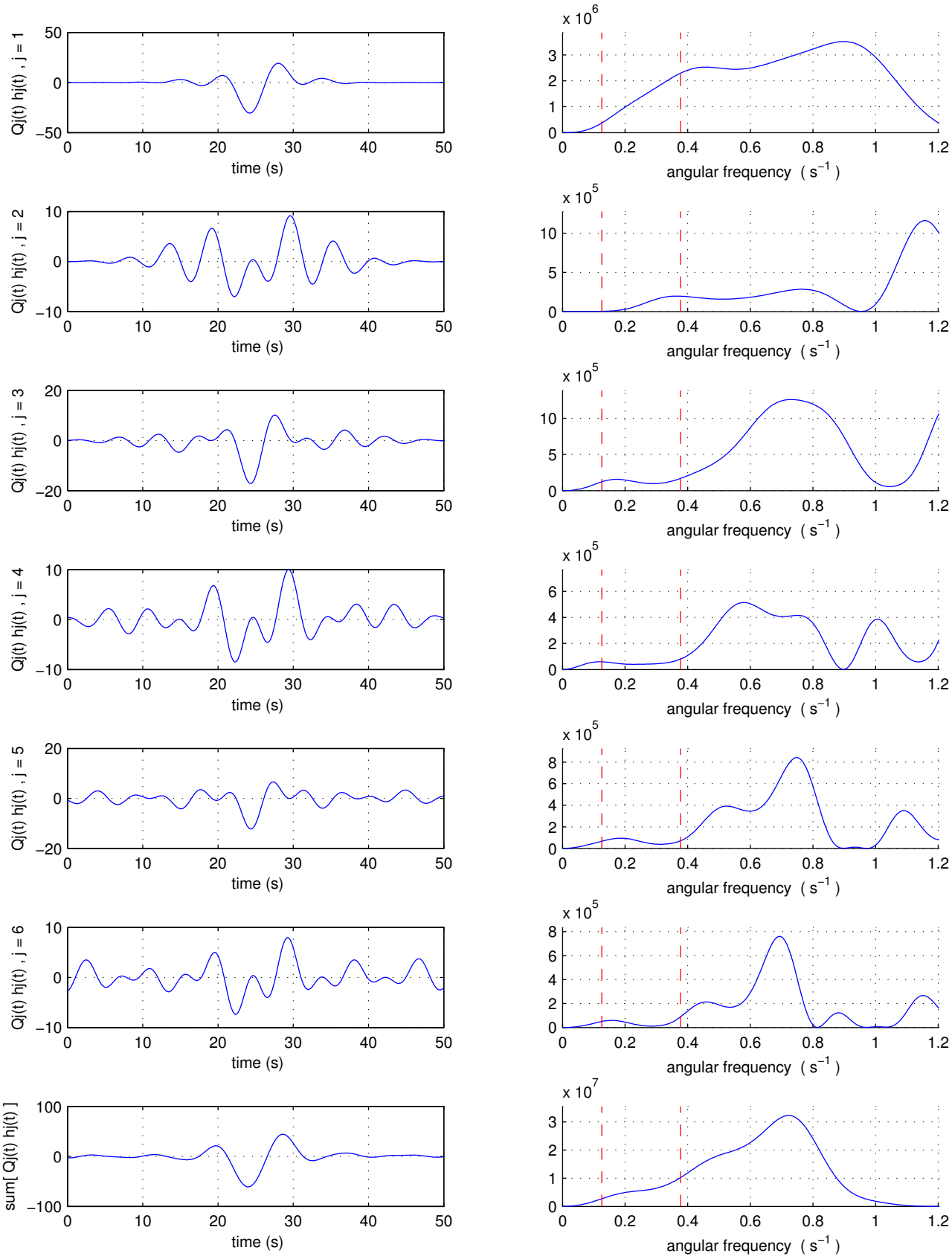

Figure 12: Construction of the multitaper amplitude adjoint source. Left column is time domain, right column is frequency domain. Bottom left shows the composite signal used for the adjoint source. See Section 2.6.
rm mt