Cross-correlation misfit measurement and adjoint source for full-waveform inversion

Ke Chen*

* *Energy Geosciences Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA.*

E-mail: kechen@lbl.gov

(March 26, 2019)

GEO-Example

Running head: ***C*ross-correlation misfit**

ABSTRACT

The waveform least-squares misfit function is often used in full-waveform inversion. However, it may lead to local minimum if the observed data and synthetic data are cycle-skipped. This document shows the cross-correlation travel time misfit function and the corresponding adjoint source for synthetic and real seismic data.

THEORY

Cross-correlation

The cross-correlation of observed and synthetic signal is defined as

$$
\begin{aligned}
C(x_r, \tau; x_s) &= \int_{t1}^{t2} d(x_r, t - \tau; x_s) d^{obs}(x_r, t; x_s) dt \\
&= \int_{t1}^{t2} d(x_r, t; x_s) d^{obs}(x_r, t + \tau; x_s) dt,
\end{aligned} \tag{1}
$$

where τ is the time shift, $d(x_r, t; x_s)$ and $d^{obs}(x_r, t; x_s)$ are the synthetic and observed seismic data at for source s and receiver r. The synthetic seismic data are dependent on the earth model parameter.

Travel time difference/peak time shift

The peak time shift is defined as the time shift that maximize the cross-correlation function

$$
\delta\tau(x_r, x_s) = \max_{\tau} C(x_r, \tau; x_s) = \max_{\tau} \int_{t1}^{t2} d(x_r, t; x_s) d^{obs}(x_r, t + \tau; x_s) dt. \tag{2}
$$

Implicit function between peak time shift and synthetic waveform data

To obtain the connection between the peak time shift $\delta\tau(x_r, x_s)$ and model parameter $m(x)$, one needs to first estimate the relation between the peak time shift $\delta\tau(x_r, x_s)$ and synthetic data $d(x_r, t; x_s)$. The derivative of $C(x_r, \tau; x_s)$ with respect to τ is given by

$$
\begin{aligned}
\frac{\partial C(x_r, \tau; x_s)}{\partial \tau} &= \int_{t1}^{t2} d(x_r, t; x_s) \frac{\partial d^{obs}(x_r, t + \tau; x_s)}{\partial \tau} dt \\
&= \int_{t1}^{t2} d(x_r, t; x_s) \frac{\partial d^{obs}(x_r, t + \tau; x_s)}{\partial (t + \tau)} \frac{\partial (t + \tau)}{\partial \tau} dt \\
&= \int_{t1}^{t2} d(x_r, t; x_s) \dot{d}^{obs}(x_r, t + \tau; x_s) dt,
\end{aligned} \tag{3}
$$

where over-dot means time derivative. It is obvious that the derivative $\partial C(x_r, \tau; x_s)/\partial\tau$ vanishes at peak time shift $\tau = \delta\tau$.

$$\frac{\partial C(x_r, \tau; x_s)}{\partial \tau} |_{\tau=\delta\tau} = \int_{t1}^{t2} d(x_r, t; x_s)\dot{d}^{obs}(x_r, t + \delta\tau; x_s)dt = 0 \tag{4}$$

Equation 4 indicates that the peak time shift $\delta\tau(x_r, x_s)$ is an implicit function of synthetic data $d(x_r, t; x_s)$. The connectivity can be denoted as

$$f[d(x_r, t; x_s), \delta\tau(x_r, x_s)] = \int_{t1}^{t2} d(x_r, t; x_s)\dot{d}^{obs}(x_r, t + \delta\tau; x_s)dt = 0. \tag{5}$$

Cross-correlation travel time misfit function for inversion

The cross-correlation travel time misfit function for full-waveform inversion is defined as the sum of the square of the peak time shift

$$J = \frac{1}{2}\sum_s\sum_r \delta\tau(x_r, x_s)^2, \tag{6}$$

where $\delta\tau(x_r, x_s)$ represents the peak time shift of observed and synthetic seismograms between source s and receiver r.

Gradient of misfit function

One of the key component of full-waveform inversion is the gradient of the misfit function

$$g(x) = \frac{\partial J}{\partial m(x)} = \sum_s\sum_r \frac{\partial\delta\tau(x_r, x_s)}{\partial m(x)}\delta\tau(x_r, x_s), \tag{7}$$

the gradient is obtained by applying the Frechet derivative on the peak time shift.

To obtain the Frechet derivative, we differentiate connectivity equation $f[d(x_r, t; x_s), \delta\tau(x_r, x_s)] = 0$ with respect to model parameter $m(x)$

$$\frac{\partial f[d(x_r, t; x_s), \delta\tau(x_r, x_s)]}{\partial d(x_r, t; x_s)}\frac{\partial d(x_r, t; x_s)}{\partial m(x)} + \frac{\partial f[d(x_r, t; x_s), \delta\tau(x_r, x_s)]}{\partial \delta\tau(x_r, x_s)}\frac{\partial\delta\tau(x_r, x_s)}{\partial m(x)} = 0 \tag{8}$$

The Frechet derivative is given by

$$\frac{\partial \delta\tau(x_r,x_s)}{\partial m(x)} = -\left[\frac{\partial f(d,\delta\tau)}{\partial \delta\tau(x_r,x_s)}\right]^{-1}\left[\frac{\partial f(d,\delta\tau)}{\partial d(x_r,t;x_s)}\frac{\partial d(x_r,t;x_s)}{\partial m(x)}\right] \tag{9}$$

Using equation 5, we have

$$\begin{aligned}\frac{\partial f(d,\delta\tau)}{\partial \delta\tau(x_r,x_s)} &= \int_{t1}^{t2} d(x_r,t;x_s)\frac{\partial \dot{d}^{obs}(x_r,t+\delta\tau;x_s)}{\partial \delta\tau(x_r,x_s)}dt \\ &= \int_{t1}^{t2} d(x_r,t;x_s)\frac{\partial \dot{d}^{obs}(x_r,t+\delta\tau;x_s)}{\partial (t+\delta\tau)}\frac{\partial (t+\delta\tau)}{\partial \delta\tau}dt \\ &= \int_{t1}^{t2} d(x_r,t;x_s)\ddot{d}^{obs}(x_r,t+\delta\tau;x_s)dt,\end{aligned} \tag{10}$$

and

$$\frac{\partial f(d,\delta\tau)}{\partial d(x_r,t;x_s)}\frac{\partial d(x_r,t;x_s)}{\partial m(x)} = \int_{t1}^{t2}\frac{\partial d(x_r,t;x_s)}{\partial m(x)}\dot{d}^{obs}(x_r,t+\delta\tau;x_s)dt. \tag{11}$$

With the above equations, the Frechet derivative can be expressed as (Luo and Schuster, 1991)

$$\frac{\partial \delta\tau(x_r,x_s)}{\partial m(x)} = -\frac{\int_{t1}^{t2}\frac{\partial d(x_r,t;x_s)}{\partial m(x)}\dot{d}^{obs}(x_r,t+\delta\tau;x_s)dt}{\int_{t1}^{t2} d(x_r,t;x_s)\ddot{d}^{obs}(x_r,t+\delta\tau;x_s)dt}. \tag{12}$$

If assuming that $d(x_r,t;x_s)^{obs}$ and $d(x_r,t;x_s)$ are purely time shifted ($d^{obs}(x_r,t{+}\delta\tau;x_s) \approx d(x_r,t;x_s)$) (Marquering et al., 1999; Dahlen et al., 2000; Tromp et al., 2005; Tape et al., 2007), the Frechet derivative can be rewritten as

$$\frac{\partial \delta\tau(x_r,x_s)}{\partial m(x)} \approx -\frac{\int_{t1}^{t2}\frac{\partial d(x_r;t;x_s)}{\partial m(x)}\dot{d}(x_r,t;x_s)dt}{\int_{t1}^{t2} d(x_r,t;x_s)\ddot{d}(x_r,t;x_s)dt}. \tag{13}$$

In this note, we follow the definition in equation 13. The gradient of misfit function (equation

7) changes to

$$
\begin{aligned}
g(x) &= \frac{\partial J}{\partial m(x)} \\
&= \sum_s \sum_r \frac{\partial \delta\tau(x_r, x_s)}{\partial m(x)} \delta\tau(x_r, x_s) \\
&\approx -\sum_s \sum_r \delta\tau(x_r, x_s) \frac{\int_{t1}^{t2} \frac{\partial d(x_r, t; x_s)}{\partial m(x)} \dot{d}(x_r, t; x_s) dt}{\int_{t1}^{t2} d(x_r, t; x_s) \ddot{d}(x_r, t; x_s) dt} \\
&= -\sum_s \sum_r \int_{t1}^{t2} \frac{\partial d(x_r, t; x_s)}{\partial m(x)} \left[\dot{d}(x_r, t; x_s) \frac{\delta\tau(x_r, x_s)}{N} \right] dt
\end{aligned} \tag{14}
$$

where $N = \int_{t1}^{t2} d(x_r, t; x_s)\ddot{d}(x_r, t; x_s)dt$ is normalizer and $\partial d(x_r, t; x_s)/\partial m(x)$ is the Frechet derivative of waveform data $d(x_r, t; x_s)$ with respect to $m(x)$. This formulation is very similar to the gradient of the ℓ_2 norm waveform misfit function

$$
g_{\ell_2}(x) = -\sum_s \sum_r \int_{t1}^{t2} \frac{\partial d(x_r, t; x_s)}{\partial m(x)} \left[\delta d(x_r, t; x_s)\right] dt \tag{15}
$$

where $\delta d(x_r, t; x_s) = d^{obs}(x_r, t; x_s) - d(x_r, t; x_s)$ is the waveform difference. The difference between equation 15 and equation 14 is the replacement of $[\delta d(x_r, t; x_s)]$ by $[\dot{d}(x_r, t; x_s)\frac{\delta\tau(x_r, x_s)}{N}]$.

Adjoint source

Instead of explicitly construct the Frechet derivative, the adjoint-state method is usually used to compute the action of Frechet derivative on data residuals. We first look at the Frechet derivative $\partial d(x_r, t; x_s)/\partial m(x)$.

The elastic wave equation can be abstractly written as

$$
\mathbf{S}\mathbf{u} = \mathbf{f}, \tag{16}
$$

where $\mathbf{S}$ denotes wave equation operator, $\mathbf{u}$ is wavefield vector and $\mathbf{f}$ is source vector. The solution of the wave equation is

$$
\mathbf{u} = \mathbf{S}^{-1}\mathbf{f}. \tag{17}
$$

The corresponding Green's function of wave equation 16 follows

$$S(x,t)G(x,t;x',t') = \delta(x-x')\delta(t-t') \tag{18}$$

The integration representation of the solution of the wave equation is

$$u(x,t) = \int_0^t \int_V G(x,t;x',t')f(x',t')d^3x'dt' \tag{19}$$

Assuming that the Green's function is time invariant

$$G(x,t;x',t') = G(x,t-t';x',0) = G(x,0;x',t'-t) \tag{20}$$

The wavefield can be written as

$$\begin{aligned} u(x,t) &= \int_0^t \int_V G(x,t-t';x',0)f(x',t')d^3x'dt' \\ &= \int_V G(x,t;x',0) * f(x',t)d^3x' \end{aligned} \tag{21}$$

The wavefield due to a point source located at $x' = x_s$ is

$$u(x,t;x_s) = G(x,t;x_s,0) * f(t;x_s), \tag{22}$$

where $f(x',t') = \delta(x'-x_s)f(t;x_s)$.

The data recorded at receivers are a sampled version of the wavefield

$$d(x_r,t;x_s) = u(x_r,t;x_s) = G(x_r,t;x_s,0) * f(t;x_s) \tag{23}$$

Differentiate the wave equation with respect to model vector

$$\frac{\partial \mathbf{S}}{\partial \mathbf{m}}\mathbf{u} + \mathbf{S}\frac{\partial \mathbf{u}}{\partial \mathbf{m}} = \mathbf{0}, \tag{24}$$

where the right side of the equation vanishes because the source term is not dependent on the model parameter. The Frechet derivative of $\mathbf{u}$ with respect to $\mathbf{m}$ can be expressed as

$$\frac{\partial \mathbf{u}}{\partial \mathbf{m}} = -\mathbf{S}^{-1}\left(\frac{\partial \mathbf{S}}{\partial \mathbf{m}}\mathbf{u}\right) \tag{25}$$

The Born scattered wavefield $\delta\mathbf{u}$ due to scatter $\delta\mathbf{m}$ can be written as

$$\delta\mathbf{u} = \frac{\partial\mathbf{u}}{\partial\mathbf{m}}\delta\mathbf{m} = -\mathbf{S}^{-1}\left(\frac{\partial\mathbf{S}}{\partial\mathbf{m}}\mathbf{u}\right)\delta\mathbf{m}, \tag{26}$$

where $\mathbf{u}$ is the incident wavefield. Using the Green's function, equation 26 changes to

$$\delta u(x,t) = -\int_V G(x,t;x',0) * \left[\frac{\partial S(x',t)}{\partial m(x')}u(x',t)\delta m(x')\right] d^3x' \tag{27}$$

The scattered data recorded at the location of receivers due to incident wavefield excited by source s are

$$\delta d(x_r,t;x_s) = \delta u(x_r,t;x_s) = -\int_V G(x_r,t;x,0) * \left[\frac{\partial S(x,t)}{\partial m(x)}u(x,t;x_s)\delta m(x)\right] d^3x \tag{28}$$

The integration kernel of Frechet derivative $\partial d/\partial m$ is

$$\frac{\partial d(x_r,t;x_s)}{\partial m(x)} = -G(x_r,t;x,0) * \left[\frac{\partial S(x,t)}{\partial m(x)}u(x,t;x_s)\right] \tag{29}$$

The gradient of cross-correlation travel time misfit function equation 14 can be written as

$$\begin{aligned} g(x) &= -\sum_s\sum_r \int_{t1}^{t2} \frac{\partial d(x_r,t;x_s)}{\partial m(x)}\left[\dot{d}(x_r,t;x_s)\frac{\delta\tau(x_r,x_s)}{N}\right] dt \\ &= \sum_s\sum_r \int_{t1}^{t2} \left\{G(x_r,t;x,0) * \left[\frac{\partial S(x,t)}{\partial m(x)}u(x,t;x_s)\right]\right\}\left[\dot{d}(x_r,t;x_s)\frac{\delta\tau(x_r,x_s)}{N}\right] dt. \end{aligned} \tag{30}$$

Using the property of convolution

$$\int [f(t) * g(t)]h(t)dt = \int g(t)[f(-t) * h(t)]dt, \tag{31}$$

equation 30 can be changed to

$$\begin{aligned} g(x) &= \sum_s\sum_r \int_{t1}^{t2} \left[\frac{\partial S(x,t)}{\partial m(x)}u(x,t;x_s)\right]\left\{G(x_r,-t;x,0) * \left[\dot{d}(x_r,t;x_s)\frac{\delta\tau(x_r,x_s)}{N}\right]\right\} dt \\ &= \sum_s\sum_r \int_{t1}^{t2} \left[\frac{\partial S(x,t)}{\partial m(x)}u(x,t;x_s)\right]\left\{G(x,0;x_r,t) * \left[\dot{d}(x_r,t;x_s)\frac{\delta\tau(x_r,x_s)}{N}\right]\right\} dt, \end{aligned} \tag{32}$$

where the time invariant property ($G(x_r,-t;x,0) = G(x_r,0;x,t)$) and reciprocity of Green's function is used ($G(x_r,0;x,t) = G(x,0;x_r,t)$). The physical meaning of the term in the curly bracket is taking the adjoint source, placing it in the receiver position x_r and propagating it in reversing time direction ($t \rightarrow 0$). The adjoint source is given by the formulation

$$s_{adj} = \dot{d}(x_r, t; x_s)\frac{\delta\tau(x_r, x_s)}{N}. \tag{33}$$

EXAMPLES

Synthetic example

In the synthetic data example, the observed data is a Ricker wavelet with peak frequency 1 Hz. The synthetic data is a shifted and amplitude reduced version of the observed data. The time shift measured by the cross-correlation is -0.2 s. It means that the observed data is 0.2 s advance than the synthetic data. The value of the cross-correlation travel time misfit function is 0.02.

Real data example

The observed seismogram is from earthquake in Berkeley January 4, 2018. The recording station is WENL - Wente Vineyards, Livermore, CA. The observed data is low-pass filtered with corner frequency 0.02 Hz. The synthetic data is simulated using $SW4$. The time shift measure by cross-correlation is 1.025 s. The value of the misfit function is 0.525

REFERENCES

Dahlen, F. A., S.-H. Hung, and G. Nolet, 2000, Fréchet kernels for finite-frequency traveltimes—i. theory: Geophysical Journal International, **141**, 157–174.

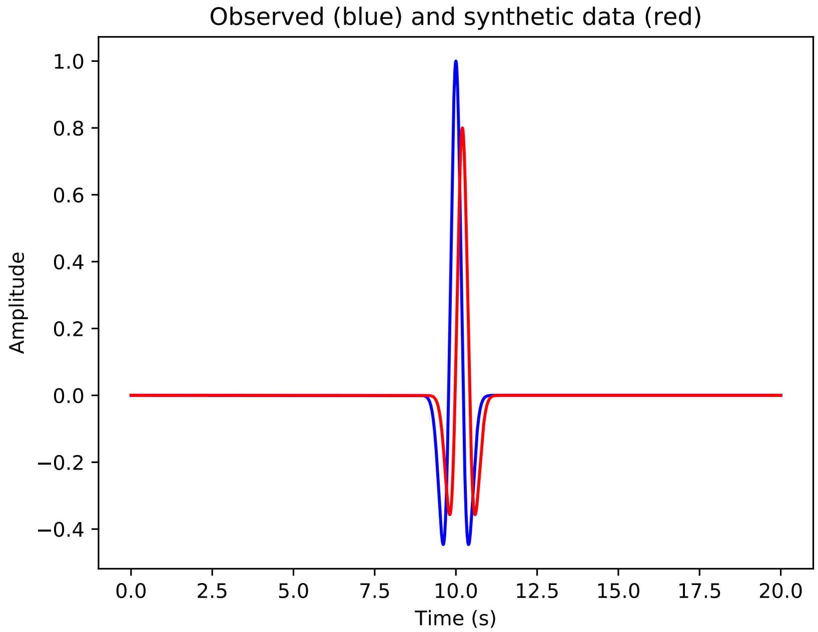
Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Marquering, H., F. Dahlen, and G. Nolet, 1999, Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana–doughnut paradox: Geophysical Journal International, **137**, 805–815.

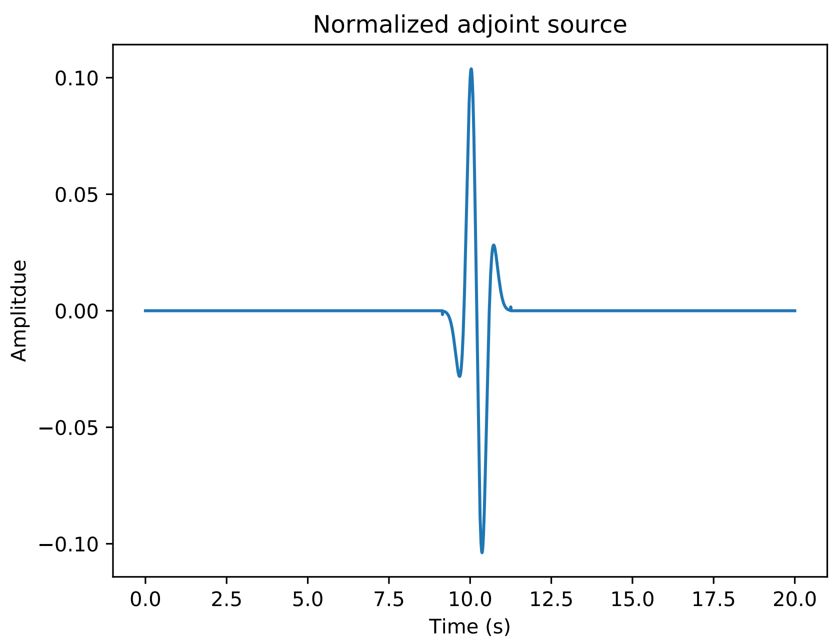
Tape, C., Q. Liu, and J. Tromp, 2007, Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves: Geophysical Journal International, **168**, 1105–1129.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, 195–216.

LIST OF FIGURES



(a)



(b)

Figure 1: (a) The observed and synthetic data. The observed data is 0.1 sec in advance of the synthetic data. (b) The normalized adjoint source for cross-correlation travel time misfit.

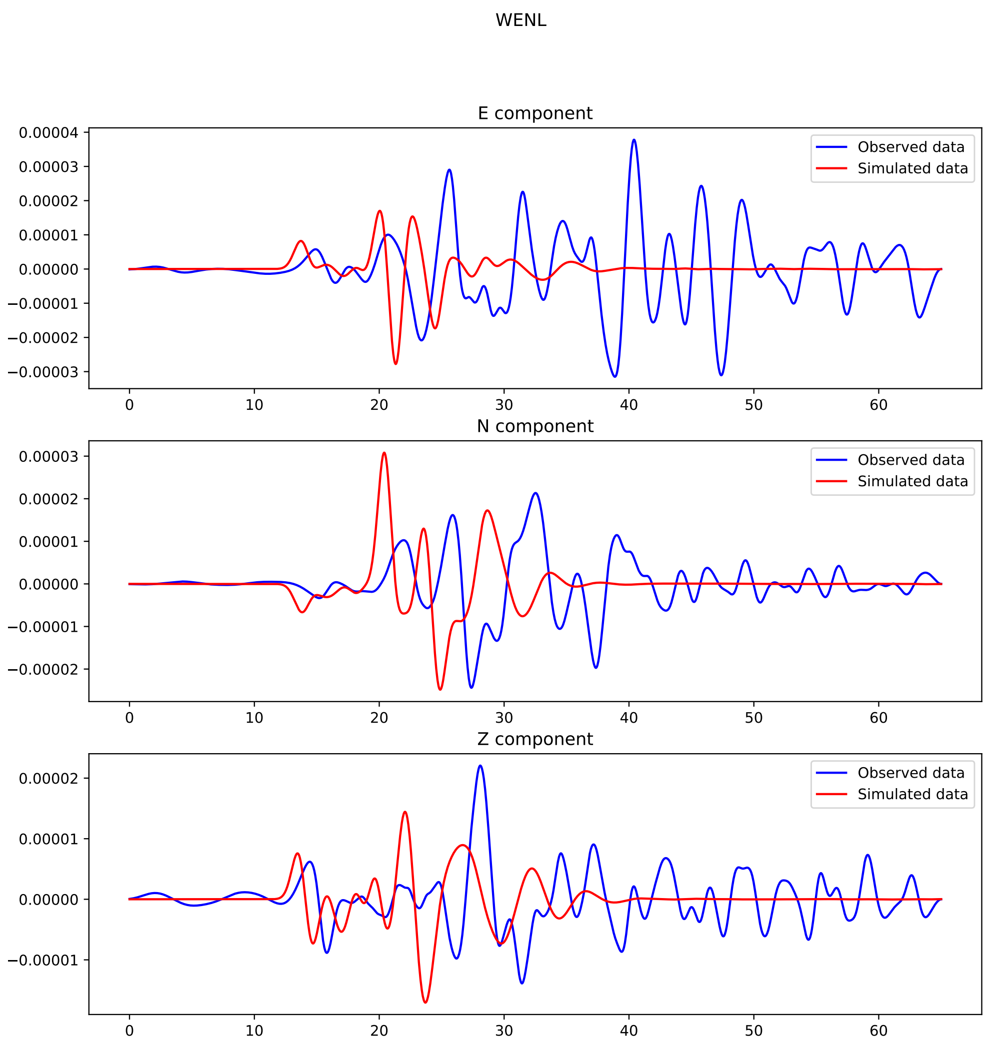
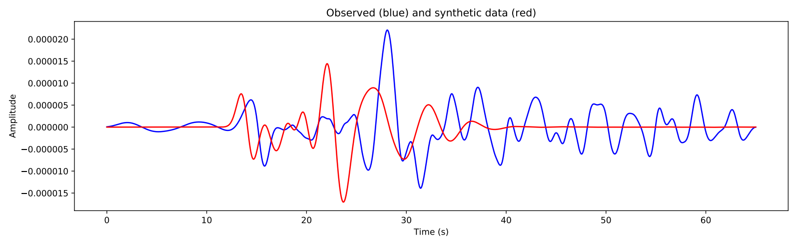
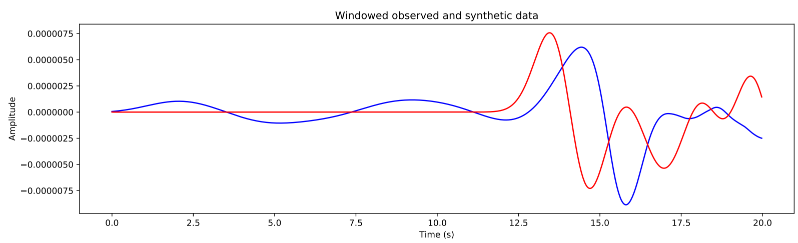


Figure 2: The Jan. 4, 2018 Berkeley earthquake data recorded at station WENL.The observed (blue) and synthetic (red) data. The data are low-pass filtered with corner frequency 0.02 Hz.

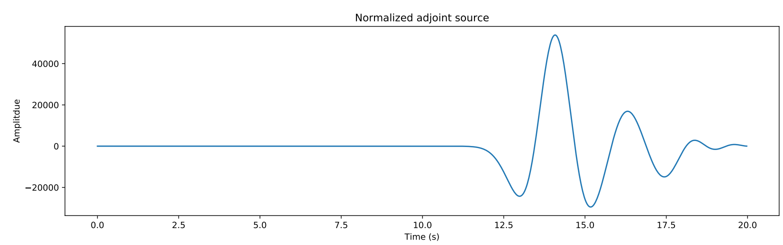
– **GEO-Example**



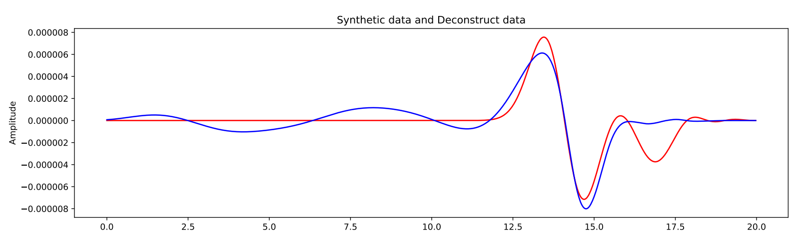
(a)



(b)



(c)



(d)

Figure 3: (a) Horizontal component of the synthetic and real earthquake data (Jan 4, 2018 Berkeley event). (b) Windowed version of the data. (c) The normalized adjoint source for cross-correlation travel time misfit. (d) The observed data is shifted by the time delay estimated from cross-correlation for comparison.

– **GEO-Example**