Geophysical Journal International

Geophys. J. Int. (2010) **183,** 791–819 doi: 10.1111/j.1365-246X.2010.04721.x

Noise cross-correlation sensitivity kernels

Jeroen Tromp,[1] Yang Luo,[2] Shravan Hanasoge[2,3] and Daniel Peter[2]

[1]*Department of Geosciences and Program in Applied & Computational Mathematics, Princeton University, Princeton, NJ, USA. E-mail: jtromp@princeton.edu*

[2] *Department of Geosciences, Princeton University, Princeton, NJ, USA*

[3]*Max-Planck-Institut für Sonnensystemforschung, Katlenburg-Lindau, Germany*

Accepted 2010 June 29. Received 2010 June 29; in original form 2010 May 20

SUMMARY

We determine finite-frequency sensitivity kernels for seismic interferometry based upon noise cross-correlation measurements. Under the assumptions that noise is spatially uncorrelated but non-uniform, we determine ensemble-averaged cross correlations between synthetic seismograms at two geographically distinct locations. By minimizing a measure of the difference between observed and simulated ensemble cross correlations—subject to the constraint that the simulated wavefield satisfies the seismic wave equation—we obtain ensemble sensitivity kernels. These ensemble kernels reflect the sensitivity of ensemble cross-correlation measurements to variations in model parameters, for example, mass density, shear and compressional wave speeds and the spatial distribution of noise. Ensemble kernels are calculated based upon the interaction between two wavefields: an ensemble forward wavefield and an ensemble adjoint wavefield. To obtain the ensemble forward wavefield, one first calculates a generating wavefield obtained by inserting a signal determined by the characteristics of the noise at the location of the first receiver, saving the results of this calculation at locations where noise is generated, that is, typically on (a portion of) the Earth's surface. Next, one uses this generating wavefield as the source of the ensemble forward wavefield associated with the first receiver. The ensemble adjoint wavefield is obtained by using measurements between simulated and observed ensemble cross correlations as a seismic source located at the second receiver. The interaction between ensemble forward and adjoint wavefields 'paints' ensemble sensitivity kernels. We illustrate the construction of ensemble kernels and their nature in two and three dimensions using a spectral-element method. In addition to a 'banana-doughnut' feature connecting the two receivers, as in traditional finite-frequency earthquake tomography, some noise cross-correlation sensitivity kernels exhibit hyperbolic 'jets' protruding from each receiver in a direction away from the other receiver. Ensemble sensitivity kernels for long-period ($T > \sim$50 s) non-uniform noise in global models exhibit sensitivity along the minor and major arcs. These kernels reflect the fact that measurements typically involve long time-series that include multi-orbit surface waves. Like free oscillations, such measurements are sensitive to structure along the great circle through the two receivers. From the perspective of noise cross-correlation tomography, we discuss strategies for inversions in terrestrial and helioseismology.

Key words: Interferometry; Seismic tomography; Computational seismology; Theoretical seismology.

1 INTRODUCTION

Diffuse seismic wavefields created by stochastic wave excitation at the Earth's surface are known to contain significant information about the interior. Seismograms of these wavefields look much like band-limited noise, reflecting the stochastic nature of wave excitation. However, cross correlations between seismograms at different seismographic stations show statistically significant signals to be present. A large body of work has attempted to model such seismic interferometry measurements and perform inversions for the Earth's structure based on various interpretations of these signals (e.g. Campillo & Paul 2003; Shapiro *et al.* 2005; Yang *et al.* 2007; Bensen *et al.* 2009). Other applications include inverting for the structure of the lunar subsurface (e.g. Larose *et al.* 2005) and monitoring dynamic changes underneath volcanoes (e.g. Brenguier *et al.* 2007).

A common interpretation of noise cross correlations is to relate them to a form of the Green's function between two receivers (e.g. Wapenaar *et al.* 2006; Fan & Snieder 2009). Depending on the manner in which the problem is posed, various idealizations, such as spatial uniformity of the noise distribution, zero attenuation, etc., have to be met for this interpretation to be valid (see e.g. Tanimoto 2008; Kimman & Trampert 2010). However, it is well known that noise sources, such as tropical storms and ocean-wave breaking, are spatially non-uniform. Furthermore, attenuation, albeit small, is certainly present in the Earth. How then are these cross-correlation measurements to be interpreted?

Helioseismology, which involves the study of wavefields at the surface (photosphere) of the Sun to infer its interior properties, also relies on the precise interpretation of correlations of a diffuse wavefield (e.g. Christensen-Dalsgaard 2002; Gizon & Birch 2005; Gizon *et al.* 2010). Waves are stochastically excited by the action of strongly turbulent convective cells that inhabit the near-surface layers of the Sun. Time-distance helioseismology, introduced by Duvall *et al.* (1993), is a common measurement technique applied to helioseismic data that requires computing cross correlations of observed surface velocities at disparate spatial points. To theoretically interpret these measurements, Woodard (1997) and Gizon & Birch (2002) suggest treating excitation as a random process in Fourier space, whereby a given realization of sources is produced by filling its Fourier transform with a set of random coefficients and multiplying by the average power spectral distribution of the sources (an assumption consistent with observations). This leads to the idea of directly modeling ensemble-averaged cross correlations, where the average is taken over all possible realizations of the source. In other words, we seek the central limit of a sufficiently large number of samples of the cross correlation, where the term 'sufficient' implies that the cross correlations must be computed over a substantial multiple of source-correlation times (assuming this is greater than the wave crossing time). Using this idea, Gizon & Birch (2002) computed ensemble-averaged sensitivity kernels for a variety of anomalies in an underlying laterally homogeneous medium, such as wave speed and attenuation perturbations and non-uniformities in the source distribution.

In this paper we determine ensemble-averaged sensitivity kernels based upon a Lagrange multiplier method. The approach is valid for general 3-D earth models, accommodating lateral variations in model parameters and a non-uniform distribution of noise. Wavefields are generated based upon a spectral-element method (e.g. Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999; Komatitsch *et al.* 2002), and ensemble finite-frequency sensitivity kernels are calculated based upon an adjoint method (Tromp *et al.* 2005; Liu & Tromp 2006, 2008).

2 GOVERNING EQUATIONS

Let us consider an earth model with volume Ω and outer free surface $\partial\Omega$. The synthetic wavefield $\mathbf{s}(\mathbf{x}, t)$ is governed by the seismic wave equation

$$\rho\partial_t^2\mathbf{s} - \nabla\cdot\mathbf{T} = \mathbf{f}, \tag{1}$$

where ρ denotes the mass density. In an elastic medium, the stress $\mathbf{T}$ is related to the displacement gradient through Hooke's law:

$$\mathbf{T} = \mathbf{c}:\nabla\mathbf{s}, \tag{2}$$

where $\mathbf{c}$ denotes the fourth-order elastic tensor. On the earth's free surface $\partial\Omega$ the traction must vanish:

$$\hat{\mathbf{n}}\cdot\mathbf{T} = \mathbf{0} \quad \text{on } \partial\Omega, \tag{3}$$

where $\hat{\mathbf{n}}$ denotes the unit outward normal on the surface. The force $\mathbf{f}$ in (1) represents the source of the wavefield. In this paper, we will be interested in uniform and non-uniform 'noise' sources.

3 GREEN'S TENSOR

The Green's tensor $\mathbf{G}(\mathbf{x}, \mathbf{x}'; t - t')$ is the response to the source $\mathbf{f}(\mathbf{x}, t) = \mathbf{I}\,\delta(\mathbf{x} - \mathbf{x}')\,\delta(t - t')$, where $\mathbf{I}$ denotes the identity tensor. In terms of the Green's tensor, the solution to the boundary value problem introduced in the previous section may be expressed as

$$\mathbf{s}(\mathbf{x}, t) = \int_{-\infty}^{t}\int_{\Omega}\mathbf{G}(\mathbf{x}, \mathbf{x}'; t - t')\cdot\mathbf{f}(\mathbf{x}', t')\,\mathrm{d}^3\mathbf{x}'\,\mathrm{d}t'. \tag{4}$$

The Green's tensor satisfies the source–receiver reciprocity relationship (e.g. Aki & Richards 1980; Dahlen & Tromp 1998)

$$\mathbf{G}(\mathbf{x}, \mathbf{x}'; t - t') = \mathbf{G}^T(\mathbf{x}', \mathbf{x}; t - t'), \tag{5}$$

where a superscript T denotes the transpose.

We will be switching between frequency- and time-domain representations of wavefields. Let us use the Fourier transform convention

$$\mathbf{s}(\mathbf{x}, \omega) = \int \mathbf{s}(\mathbf{x}, t)\,\exp(-\mathrm{i}\omega t)\,\mathrm{d}t, \qquad \mathbf{s}(\mathbf{x}, t) = \frac{1}{2\pi}\int \mathbf{s}(\mathbf{x}, \omega)\,\exp(\mathrm{i}\omega t)\,\mathrm{d}\omega. \tag{6}$$

The frequency-domain response may now be expressed as

$$\mathbf{s}(\mathbf{x}, \omega) = \int_{\Omega}\mathbf{G}(\mathbf{x}, \mathbf{x}'; \omega)\cdot\mathbf{f}(\mathbf{x}', \omega)\,\mathrm{d}^3\mathbf{x}'. \tag{7}$$

4 NOISE CROSS CORRELATION

We are interested in observed and simulated cross correlations between time-series recorded at two distinct locations. In practice, one uses an 'ensemble average' of many cross correlations, which we will refer to as the 'ensemble cross correlation'. From an observational perspective, obtaining an ensemble average involves stacking hours (Sun) or months (Earth) worth of noise cross-correlation data. In principle, one can obtain simulated ensemble cross correlations by stacking numerical simulations involving random noise (Cupillard & Capdeville 2010), but for routine seismic interferometry the computational cost is prohibitive. Therefore, our first goal is to devise a practical approach to calculating synthetic ensemble cross correlations.

Let us consider the $\hat{\boldsymbol{\nu}}^\alpha$ component of the displacement at location $\mathbf{x}^\alpha$, and the $\hat{\boldsymbol{\nu}}^\beta$ component of the displacement at location $\mathbf{x}^\beta$:

$$s^\alpha(t) \equiv \hat{\boldsymbol{\nu}}^\alpha \cdot \mathbf{s}(\mathbf{x}^\alpha, t), \qquad s^\beta(t) \equiv \hat{\boldsymbol{\nu}}^\beta \cdot \mathbf{s}(\mathbf{x}^\beta, t). \tag{8}$$

The cross correlation $C^{\alpha\beta}$ between the two time-series s^α and s^β is given by

$$C^{\alpha\beta}(t) = \int s^\alpha(t+\tau)\, s^\beta(\tau)\, \mathrm{d}\tau. \tag{9}$$

Note that, based upon this definition, we always have the symmetry

$$C^{\alpha\beta}(t) = C^{\beta\alpha}(-t). \tag{10}$$

We seek to determine the ensemble average of the cross correlation (9).

Using the Fourier convention (6), we have

$$C^{\alpha\beta}(t) = \frac{1}{2\pi} \int s^\alpha(\omega)\, s^{\beta *}(\omega)\, \exp(\mathrm{i}\omega t)\, \mathrm{d}\omega, \tag{11}$$

where the asterisk denotes complex conjugation. In terms of the Green's tensor, we have

$$s^\alpha(\omega) = \hat{\boldsymbol{\nu}}^\alpha \cdot \mathbf{s}(\mathbf{x}^\alpha, \omega) = \hat{\nu}_i^\alpha \int G_{ij}(\mathbf{x}^\alpha, \mathbf{x}'; \omega)\, f_j(\mathbf{x}', \omega)\, \mathrm{d}^3\mathbf{x}'. \tag{12}$$

Now the cross correlation (9) may be rewritten as

$$C^{\alpha\beta}(t) = \frac{1}{2\pi}\, \hat{\nu}_i^\alpha\, \hat{\nu}_\ell^\beta \iiint G_{ij}(\mathbf{x}^\alpha, \mathbf{x}'; \omega)\, f_j(\mathbf{x}', \omega)\, G^*_{\ell m}(\mathbf{x}^\beta, \mathbf{x}''; \omega)\, f^*_m(\mathbf{x}'', \omega)\, \exp(\mathrm{i}\omega t)\, \mathrm{d}^3\mathbf{x}'\, \mathrm{d}^3\mathbf{x}''\, \mathrm{d}\omega. \tag{13}$$

Throughout this paper, we will assume that noise is 'spatially uncorrelated'. In terms of the source $\mathbf{f}$ in the wave eq. (1) this implies that

$$\langle f_j(\mathbf{x}', \omega)\, f^*_m(\mathbf{x}'', \omega)\rangle = S_{jm}(\mathbf{x}', \omega)\, \delta(\mathbf{x}' - \mathbf{x}''), \tag{14}$$

where $\langle\cdot\rangle$ denotes an ensemble average (Woodard 1997; Gizon & Birch 2002). Note that we have the symmetry $S_{jm} = S_{mj}$. In the time domain, the assumption is that

$$\langle f_j(\mathbf{x}', t')\, f_m(\mathbf{x}'', t'')\rangle = S_{jm}(\mathbf{x}', t' - t'')\, \delta(\mathbf{x}' - \mathbf{x}''), \tag{15}$$

where the time of interaction is short, that is, $S_{jm}(\mathbf{x}', t' - t'')$ is narrowly concentrated around $t' - t'' = 0$. In both terrestrial and helioseismology, S_{jm} is generally non-zero only on the surface, that is, where the noise is generated. The forcing is frequently in the vertical direction, in which case S_{jm} is of the form $S_{jm}(\mathbf{x}, \omega) = S(\mathbf{x}, \omega)\, \hat{n}_j(\mathbf{x})\, \hat{n}_m(\mathbf{x})$, where $\hat{\mathbf{n}}$ is the unit outward normal to the surface. The assumed spatial behaviour of the ensemble-averaged noise, (14), is similar to the delta-function behaviour of the sum of the products of spherical harmonics $Y_{\ell m}$ of degree ℓ and order m at two distinct locations $\hat{\mathbf{r}}$ and $\hat{\mathbf{r}}'$ on the unit sphere: $\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\hat{\mathbf{r}})\, Y^*_{\ell m}(\hat{\mathbf{r}}') = \delta(\hat{\mathbf{r}}, \hat{\mathbf{r}}')$.

Upon taking the ensemble average of (13) and using reciprocity (5), we find the 'ensemble cross correlation'

$$\langle C^{\alpha\beta}\rangle(t) \;=\; \frac{1}{2\pi}\, \hat{\nu}_i^\alpha\, \hat{\nu}_\ell^\beta \iint S_{jm}(\mathbf{x}, \omega)\, G_{ji}(\mathbf{x}, \mathbf{x}^\alpha; \omega)\, G^*_{m\ell}(\mathbf{x}, \mathbf{x}^\beta; \omega)\, \exp(\mathrm{i}\omega t)\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}\omega. \tag{16}$$

The spatial integration is limited to regions where noise is generated, that is, the Earth's surface. Note that, like the cross correlation $C^{\alpha\beta}$, the ensemble cross correlation has the symmetry

$$\langle C^{\alpha\beta}\rangle(t) = \langle C^{\beta\alpha}\rangle(-t). \tag{17}$$

We emphasize that the symmetry (17) is a direct consequence of the cross-correlation symmetry (10), valid for any earth model and any noise distribution (even if the distribution is spatially correlated). Consequently, measuring $\langle C^{\alpha\beta}\rangle(t)$ is identical to measuring $\langle C^{\beta\alpha}\rangle(-t)$. One usually takes the option of measuring the positive branches ($t > 0$) of $\langle C^{\alpha\beta}\rangle(t)$ and $\langle C^{\beta\alpha}\rangle(t)$, since the negative branches ($t < 0$) contain no further information given the symmetry (17). In Section 9, we demonstrate how to take advantage of this symmetry to facilitate noise cross-correlation tomography.

5 OPTIMIZATION

Suppose we have observed ensemble cross correlation $\langle C^{\alpha\beta}\rangle_{\mathrm{obs}}$ and simulated ensemble cross correlation $\langle C^{\alpha\beta}\rangle_{\mathrm{sim}}$, calculated based upon (16). Our objective is to minimize their differences,

$$\Delta\langle C^{\alpha\beta}\rangle = \langle C^{\alpha\beta}\rangle_{\mathrm{sim}} - \langle C^{\alpha\beta}\rangle_{\mathrm{obs}}, \tag{18}$$

based upon the misfit function

$$\chi = \frac{1}{2} \int \left[\langle C^{\alpha\beta} \rangle_{\rm sim} - \langle C^{\alpha\beta} \rangle_{\rm obs} \right]^2 \, {\rm d}t - \left\langle \iint_\Omega \boldsymbol{\lambda} \cdot \left(\rho \partial_t^2 \mathbf{s} - \boldsymbol{\nabla} \cdot \mathbf{T} - \mathbf{f} \right) \, {\rm d}^3\mathbf{x} \, {\rm d}t \right\rangle. \tag{19}$$

The second term in (19) reflects the fact that we wish to minimize the difference $\Delta\langle C^{\alpha\beta}\rangle$, subject to the ensemble-averaged constraint that the displacement field $\mathbf{s}$ satisfies the boundary value problem discussed in Section 2, that is, the wave equation.

The time integration in (19) is over all times, but in practice we may wish to window the ensemble cross-correlation difference $\Delta\langle C^{\alpha\beta}\rangle$, for example, by selecting only the branch for positive times. Rather than building such a windowing operation explicitly into the definition of the misfit function, for brevity and to avoid clutter, we work under the assumption that the ensemble cross correlation may have been windowed with a function $w(t)$, that is, that $\Delta\langle C^{\alpha\beta}\rangle(t)$ implicitly means $w(t) \times \Delta\langle C^{\alpha\beta}\rangle(t)$. The symmetry of the ensemble cross correlation (17) remains valid. In other words, measuring the positive branch of $\langle C^{\alpha\beta}\rangle$ leads to the same result as measuring the negative branch of $\langle C^{\beta\alpha}\rangle$. Filtering operations may be treated in the same fashion, thereby enabling multifrequency measurements.

We demonstrate in Appendices A & B that the change in the misfit function (19) may be written in the form

$$\delta\chi = \left\langle \int_\Omega (\delta \ln \rho \, K_\rho + \delta\mathbf{c} :: \mathbf{K}_\mathbf{c} + \delta\Sigma) \, {\rm d}^3\mathbf{x} \right\rangle, \tag{20}$$

where $\delta \ln \rho = \delta\rho/\rho$ denotes relative perturbations in density, and where we have introduced the kernels (Liu & Tromp 2006; Liu & Tromp 2008)

$$K_\rho = -\int \rho \, \mathbf{s}^\dagger(-t) \cdot \partial_t^2 \mathbf{s}(t) \, {\rm d}t, \tag{21}$$

$$\mathbf{K}_\mathbf{c} = -\int \boldsymbol{\nabla}\mathbf{s}^\dagger(-t) \, \boldsymbol{\nabla}\mathbf{s}(t) \, {\rm d}t. \tag{22}$$

and where

$$\delta\Sigma = \int \mathbf{s}^\dagger \cdot \delta\mathbf{f} \, {\rm d}t. \tag{23}$$

Realizing that $\delta\mathbf{c}$ and $\mathbf{K}_\mathbf{c}$ are both fourth-order tensors, we use the notation $\delta\mathbf{c} :: \mathbf{K}_\mathbf{c} = \delta c_{ijkl} K_{c_{ijkl}}$ in (20). The symmetries of the elastic tensor $\mathbf{c}$ imply the same symmetries in the fourth-order kernel $\mathbf{K}_\mathbf{c}$. The calculation (20) requires access to the forward wavefield $\mathbf{s}(\mathbf{x}, t)$ generated by the noise source $\mathbf{f}$, and an 'adjoint wavefield' $\mathbf{s}^\dagger(\mathbf{x}, -t)$ generated by the 'adjoint source'

$$\mathbf{f}^\dagger(\mathbf{x}, t) = \hat{\mathbf{v}}^\beta \int \Delta\langle C^{\alpha\beta}\rangle(\tau) \, s^\alpha(-t+\tau) \, {\rm d}\tau \, \delta(\mathbf{x} - \mathbf{x}^\beta) + \hat{\mathbf{v}}^\alpha \int \Delta\langle C^{\beta\alpha}\rangle(\tau) \, s^\beta(-t+\tau) {\rm d}\tau \, \delta(\mathbf{x} - \mathbf{x}^\alpha), \tag{24}$$

which is symmetric in α and β. The adjoint wavefield $\mathbf{s}^\dagger$ is related to the Lagrange multiplier $\boldsymbol{\lambda}$ by

$$\mathbf{s}^\dagger(\mathbf{x}, t) = \boldsymbol{\lambda}(\mathbf{x}, -t). \tag{25}$$

As it stands, eq. (20) is not useable because it involves an ensemble average $\langle\cdot\rangle$. We demonstrate in Appendix C that, after taking this average, the variation in the misfit function (20) may be written in the practical form

$$\delta\chi = \int_\Omega (\delta \ln \rho \, \langle K_\rho \rangle + \delta\mathbf{c} :: \langle \mathbf{K}_\mathbf{c} \rangle + \langle \delta\Sigma \rangle) \, {\rm d}^3\mathbf{x}, \tag{26}$$

where we have introduced the 'ensemble sensitivity kernels'

$$\langle K_\rho \rangle = -\int \rho \, \left[\boldsymbol{\Phi}^{\dagger\,\alpha\beta}(-t) \cdot \partial_t^2 \boldsymbol{\Phi}^\alpha(t) + \boldsymbol{\Phi}^{\dagger\,\beta\alpha}(-t) \cdot \partial_t^2 \boldsymbol{\Phi}^\beta(t) \right] {\rm d}t, \tag{27}$$

$$\langle \mathbf{K}_\mathbf{c} \rangle = -\int \left[\boldsymbol{\nabla}\boldsymbol{\Phi}^{\dagger\,\alpha\beta}(-t) \, \boldsymbol{\nabla}\boldsymbol{\Phi}^\alpha(t) + \boldsymbol{\nabla}\boldsymbol{\Phi}^{\dagger\,\beta\alpha}(-t) \, \boldsymbol{\nabla}\boldsymbol{\Phi}^\beta(t) \right] {\rm d}t, \tag{28}$$

and where

$$\langle \delta\Sigma \rangle = \int \left[\boldsymbol{\Phi}^{\dagger\,\alpha\beta}(-t) \cdot \delta\mathbf{F}^\alpha(t) + \boldsymbol{\Phi}^{\dagger\,\beta\alpha}(-t) \cdot \delta\mathbf{F}^\beta(t) \right] {\rm d}t. \tag{29}$$

The fourth-order kernel $\langle \mathbf{K}_\mathbf{c} \rangle$ has the same symmetries as the elastic tensor $\mathbf{c}$. One may use (27) and (28) to determine structural model parameters. The characteristics of the noise may be determined based upon (29), analogous to kinematic source inversions in the earthquake case based upon eq. (98) of Tromp *et al.* (2005).

The calculation of (27)–(29) requires access to two types of wavefields: $\boldsymbol{\Phi}^\alpha$ and $\boldsymbol{\Phi}^\beta$ in regular time, t, and $\boldsymbol{\Phi}^{\dagger\,\alpha\beta}$ and $\boldsymbol{\Phi}^{\dagger\,\beta\alpha}$ in reverse time, $-t$. The 'ensemble forward wavefield' $\boldsymbol{\Phi}^\alpha$ is generated based upon the source

$$F_i^\alpha(\mathbf{x}, \omega) \equiv G_{jk}^*(\mathbf{x}, \mathbf{x}^\alpha; \omega) \, \nu_k^\alpha \, S_{ij}(\mathbf{x}, \omega). \tag{30}$$

We may define the 'generating wavefield' as follows:

$$\eta_i^\alpha(\mathbf{x}, \omega) \equiv F_i^{\alpha *}(\mathbf{x}, \omega) = G_{jk}(\mathbf{x}, \mathbf{x}^\alpha; \omega) \, \nu_k^\alpha S_{ij}^*(\mathbf{x}, \omega). \tag{31}$$

Thus the source of the ensemble forward wavefield is simply the time-reversed generating wavefield

$$\mathbf{F}^\alpha(\mathbf{x}, t) = \boldsymbol{\eta}^\alpha(\mathbf{x}, -t). \tag{32}$$

The 'ensemble adjoint wavefield' $\mathbf{\Phi}^{\dagger\,\alpha\beta}$ is generated based upon the 'ensemble adjoint source'

$$\mathbf{F}^{\dagger\,\alpha\beta}(\mathbf{x},t) = \hat{\boldsymbol{\nu}}^{\beta}\,\Delta\langle C^{\alpha\beta}\rangle(t)\,\delta(\mathbf{x}-\mathbf{x}^{\beta}). \tag{33}$$

Finally, the source perturbation (29) is determined by

$$\delta F_i^{\alpha}(\mathbf{x},\omega) \equiv G^{*}_{jk}(\mathbf{x},\mathbf{x}^{\alpha};\omega)\,\nu_k^{\alpha}\,\delta S_{ij}(\mathbf{x},\omega). \tag{34}$$

We conclude that, after taking the ensemble average, the ensemble forward source $\mathbf{F}^{\alpha}$ plays the role of the original force $\mathbf{f}$, and $\mathbf{\Phi}^{\alpha}$ plays the role of the regular wavefield $\mathbf{s}$. Similarly, the ensemble adjoint source $\mathbf{F}^{\dagger\,\alpha\beta}$ plays the role of the original adjoint source $\mathbf{f}^{\dagger}$, and $\mathbf{\Phi}^{\dagger\,\alpha\beta}$ plays the role of the adjoint wavefield $\mathbf{s}^{\dagger}$.

Note that the ensemble kernels (27) and (28) each consist of two parts. Take the ensemble density kernel (27) as an example: the first contribution is from interactions between $\partial_t^2\mathbf{\Phi}^{\alpha}(t)$ and $\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)$, whereas the second contribution involves $\partial_t^2\mathbf{\Phi}^{\beta}(t)$ and $\mathbf{\Phi}^{\dagger\,\beta\alpha}(-t)$. In what follows (especially in the figures), we will use superscripts $\alpha\beta$ and $\beta\alpha$ to denote the first and second contributions, respectively, for example,

$$\left\langle K_{\rho}^{\alpha\beta}\right\rangle = -\int \rho\,\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)\cdot\partial_t^2\mathbf{\Phi}^{\alpha}(t)\,\mathrm{d}t, \tag{35}$$

$$\left\langle K_{\rho}^{\beta\alpha}\right\rangle = -\int \rho\,\mathbf{\Phi}^{\dagger\,\beta\alpha}(-t)\cdot\partial_t^2\mathbf{\Phi}^{\beta}(t)\,\mathrm{d}t. \tag{36}$$

Note that $\langle K_{\rho}^{\alpha\beta}\rangle$ and $\langle K_{\rho}^{\beta\alpha}\rangle$ each involve 'the same measurement', for example, the positive branch of the ensemble cross correlation. The final ensemble kernel $\langle K_{\rho}\rangle$ is just the sum of $\langle K_{\rho}^{\alpha\beta}\rangle$ and $\langle K_{\rho}^{\beta\alpha}\rangle$

$$\langle K_{\rho}\rangle = \left\langle K_{\rho}^{\alpha\beta}\right\rangle + \left\langle K_{\rho}^{\beta\alpha}\right\rangle. \tag{37}$$

It is shown in Appendix C that the $\hat{\boldsymbol{\nu}}^{\alpha}$ component of the ensemble forward wavefield $\mathbf{\Phi}^{\beta}$ evaluated at location $\mathbf{x}^{\alpha}$ is equal to the ensemble cross correlation $\langle C^{\alpha\beta}\rangle$:

$$\langle C^{\alpha\beta}\rangle(t) = \hat{\boldsymbol{\nu}}^{\alpha}\cdot\mathbf{\Phi}^{\beta}(\mathbf{x}^{\alpha},t). \tag{38}$$

The symmetry (17) and the relationship (38) imply that

$$\hat{\boldsymbol{\nu}}^{\alpha}\cdot\mathbf{\Phi}^{\beta}(\mathbf{x}^{\alpha},t) = \hat{\boldsymbol{\nu}}^{\beta}\cdot\mathbf{\Phi}^{\alpha}(\mathbf{x}^{\beta},-t), \tag{39}$$

illustrating how either $\mathbf{\Phi}^{\alpha}$ or $\mathbf{\Phi}^{\beta}$ may be used to calculate the ensemble cross correlation $\langle C^{\alpha\beta}\rangle$.

In an isotropic earth model we have $c_{jklm} = (\kappa - {}^{2}/_{3}\,\mu)\,\delta_{jk}\delta_{lm} + \mu\,(\delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl})$, where μ and κ denote the shear and bulk moduli, respectively. Thus we may write

$$\delta\mathbf{c}::\langle\mathbf{K}_{\mathbf{c}}\rangle = \delta\ln\mu\,\langle K_{\mu}\rangle + \delta\ln\kappa\,\langle K_{\kappa}\rangle, \tag{40}$$

where the ensemble isotropic kernels $\langle K_{\mu}\rangle$ and $\langle K_{\kappa}\rangle$ represent Fréchet derivatives with respect to relative shear and bulk moduli perturbations $\delta\ln\mu = \delta\mu/\mu$ and $\delta\ln\kappa = \delta\kappa/\kappa$, respectively. These isotropic kernels are given by

$$\langle K_{\mu}\rangle = -\int 2\mu\,\left[\mathbf{D}^{\dagger\,\alpha\beta}(-t):\mathbf{D}^{\alpha}(t) + \mathbf{D}^{\dagger\,\beta\alpha}(-t):\mathbf{D}^{\beta}(t)\right]\,\mathrm{d}t, \tag{41}$$

$$\langle K_{\kappa}\rangle = -\int \kappa\,\left[\boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)\,\boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\alpha}(t) + \boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\dagger\,\beta\alpha}(-t)\,\boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\beta}(t)\right]\,\mathrm{d}t, \tag{42}$$

where

$$\mathbf{D}^{\alpha} = \frac{1}{2}[\boldsymbol{\nabla}\mathbf{\Phi}^{\alpha} + (\boldsymbol{\nabla}\mathbf{\Phi}^{\alpha})^{T}] - \frac{1}{3}(\boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\alpha})\,\mathbf{I}, \qquad \mathbf{D}^{\dagger\,\alpha\beta} = \frac{1}{2}[\boldsymbol{\nabla}\mathbf{\Phi}^{\dagger\,\alpha\beta} + (\boldsymbol{\nabla}\mathbf{\Phi}^{\dagger\,\alpha\beta})^{T}] - \frac{1}{3}(\boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\dagger\,\alpha\beta})\,\mathbf{I}, \tag{43}$$

denote the traceless ensemble strain deviator and corresponding adjoint.

We may express the Fréchet derivatives in an isotropic earth model in terms of relative variations in density $\delta\ln\rho$, shear wave speed $\delta\ln\beta$ and compressional wave speed $\delta\ln\alpha$ based upon the relationship (Tromp *et al.* 2005).

$$\delta\ln\rho\,\langle K_{\rho}\rangle + \delta\ln\mu\,\langle K_{\mu}\rangle + \delta\ln\kappa\,\langle K_{\kappa}\rangle = \delta\ln\rho\,\langle K'_{\rho}\rangle + \delta\ln\beta\,\langle K_{\beta}\rangle + \delta\ln\alpha\,\langle K_{\alpha}\rangle, \tag{44}$$

where

$$\langle K'_{\rho}\rangle = \langle K_{\rho}\rangle + \langle K_{\kappa}\rangle + \langle K_{\mu}\rangle, \tag{45}$$

$$\langle K_{\beta}\rangle = 2\left[\langle K_{\mu}\rangle - \frac{4}{3}\,\frac{\mu}{\kappa}\,\langle K_{\kappa}\rangle\right], \tag{46}$$

$$\langle K_{\alpha}\rangle = 2\left(\frac{\kappa + \frac{4}{3}\,\mu}{\kappa}\right)\langle K_{\kappa}\rangle. \tag{47}$$

In Section 9 we rewrite these gradients and related ensemble kernels in a more compact form, suitable from a noise cross-correlation tomography perspective.

6 ALTERNATIVE MEASUREMENTS

As discussed by Tromp *et al.* (2005) and Fichtner & Igel (2009), there are numerous other measurements one may choose to use. In each case the ensemble adjoint source needs to be modified, and the resulting finite-frequency sensitivity kernels will reflect the new measurement. For example, rather than using differences between observed and simulated ensemble cross correlations, one may opt for the non-dimensional misfit

$$\chi = \frac{1}{N^{\alpha\beta}} \int \left[\langle C^{\alpha\beta} \rangle_{\rm sim} - \langle C^{\alpha\beta} \rangle_{\rm obs} \right]^2 \, {\rm d}t, \tag{48}$$

where

$$N^{\alpha\beta} = \int \langle C^{\alpha\beta} \rangle_{\rm obs}^2 \, {\rm d}t. \tag{49}$$

This would result in the ensemble adjoint source

$$\mathbf{F}^{\dagger\,\alpha\beta}(\mathbf{x}, t) = \hat{\mathbf{v}}^\beta \, \Delta \langle C^{\alpha\beta} \rangle(t) \, \delta(\mathbf{x} - \mathbf{x}^\beta) / N^{\alpha\beta}. \tag{50}$$

We note again that in practice the ensemble cross-correlation differences may have been windowed with a function $w(t)$, for example, by selecting the main arrival for positive times.

Perhaps a better choice involves measuring ensemble cross-correlation delay times between observed and simulated ensemble cross correlations, just like one determines cross-correlation traveltime anomalies for earthquake-generated seismic arrivals, such as *P* and *S* waves. The traveltime misfit function is of the form

$$\chi = \frac{1}{2} \, (\Delta T^{\alpha\beta})^2, \tag{51}$$

which would result in the ensemble adjoint source

$$\mathbf{F}^{\dagger\,\alpha\beta}(\mathbf{x}, t) = -\hat{\mathbf{v}}^\beta \, \Delta T^{\alpha\beta} \, \langle \dot{C}^{\alpha\beta} \rangle_{\rm sim}(t) \, \delta(\mathbf{x} - \mathbf{x}^\beta) / N'^{\alpha\beta}, \tag{52}$$

where a dot denotes differentiation with respect to time and $\Delta T^{\alpha\beta}$ the cross-correlation traveltime anomaly $T_{\rm sim}^{\alpha\beta} - T_{\rm obs}^{\alpha\beta}$, and where the normalization constant is given by

$$N'^{\alpha\beta} = \int \left[\langle \dot{C}^{\alpha\beta} \rangle_{\rm sim} \right]^2 {\rm d}t. \tag{53}$$

In this case $\langle \dot{C}^{\alpha\beta} \rangle_{\rm sim}(t)$ is implicitly windowed by a function $w(t)$ defining the interval over which the cross-correlation delay time is measured. It is important to note that $\Delta T^{\alpha\beta} \, \langle \dot{C}^{\alpha\beta} \rangle_{\rm sim}(t) = \Delta T^{\beta\alpha} \, \langle \dot{C}^{\beta\alpha} \rangle_{\rm sim}(-t)$, and that both $\Delta T^{\alpha\beta}$ and $\langle \dot{C}^{\alpha\beta} \rangle_{\rm sim}(t)$ change sign when exchanging α and β.

$$\Delta T^{\alpha\beta} = -\Delta T^{\beta\alpha}, \qquad \langle \dot{C}^{\alpha\beta} \rangle_{\rm sim}(t) = -\langle \dot{C}^{\beta\alpha} \rangle_{\rm sim}(-t). \tag{54}$$

Finally, recognizing the fact that observed and simulated ensemble cross correlations correspond pre-dominantly to dispersive surface waves, one may choose to measure frequency-dependent phase and amplitude anomalies. This may be accomplished based upon generalized data functionals (Gee & Jordan 1992), phase-matched filtering (Ekström *et al.* 1997), multitaper dispersion measurements (Laske & Masters 1996; Tape *et al.* 2009, 2010) or time-frequency misfits (Fichtner *et al.* 2009).

7 2-D EXAMPLES

To develop intuition for the construction and nature of ensemble sensitivity kernels, we start with some simple 2-D examples.

Consider the 2-D wave equation for 'membrane' surface waves, as in Tape *et al.* (2007). Wave propagation is governed by the 2-D wave equation

$$\rho \partial_t^2 s = \boldsymbol{\nabla} \cdot (\mu \boldsymbol{\nabla} s) + f, \tag{55}$$

where s denotes the vertical component of waves in a horizontal membrane, and waves travel with shear wave speed $\beta = \sqrt{\mu/\rho}$.

The frequency-domain Green's function for this problem may be expressed as a sum over the normal modes of the system, $s_k(\mathbf{x})$, with associated eigenfrequencies, ω_k.

$$G(\mathbf{x}, \mathbf{x}'; \omega) = \sum_k \left(\omega_k^2 - \omega^2 \right)^{-1} \, s_k(\mathbf{x}) \, s_k(\mathbf{x}'), \tag{56}$$

or in the time domain

$$G(\mathbf{x}, \mathbf{x}'; t) = \sum_k s_k(\mathbf{x}) \, s_k(\mathbf{x}') \, \omega_k^{-1} \, \sin(\omega_k t) \, H(t), \tag{57}$$

where $H(t)$ denotes the Heaviside function. The modes are orthonormal in the sense

$$\int \rho(\mathbf{x}) \, s_k(\mathbf{x}) \, s_{k'}(\mathbf{x}) \, {\rm d}^2\mathbf{x} = \delta_{kk'}. \tag{58}$$

For simplicity, we will assume that the ensemble-averaged noise (14) is uniform: $S(\mathbf{x}, \omega) = S(\omega)$. In the frequency domain, the ensemble cross correlation (16) becomes

$$\langle C^{\alpha\beta}\rangle(\omega) = \int G(\mathbf{x}, \mathbf{x}^{\alpha}; \omega)\, G^{*}(\mathbf{x}, \mathbf{x}^{\beta}; \omega)\, S(\omega)\, \mathrm{d}^2\mathbf{x}$$
$$= \int \sum_{k} \left(\omega_k^2 - \omega^2\right)^{-1} s_k(\mathbf{x})\, s_k(\mathbf{x}^{\alpha}) \sum_{k'} \left(\omega_{k'}^2 - \omega^2\right)^{-1} s_{k'}(\mathbf{x})\, s_{k'}(\mathbf{x}^{\beta})\, S(\omega)\, \mathrm{d}^2\mathbf{x}. \tag{59}$$

Now we assume that the density ρ is constant, and use mode orthonormality (58):

$$\langle C^{\alpha\beta}\rangle(\omega) = \rho^{-1} \sum_{k} \left(\omega_k^2 - \omega^2\right)^{-2} s_k(\mathbf{x}^{\alpha})\, s_k(\mathbf{x}^{\beta})\, S(\omega). \tag{60}$$

Upon differentiating (56) with respect to ω we find that

$$\tilde{G}(\mathbf{x}, \mathbf{x}'; \omega) \equiv (2\,\omega)^{-1} \partial_{\omega} G(\mathbf{x}, \mathbf{x}'; \omega) = \sum_{k} \left(\omega_k^2 - \omega^2\right)^{-2} s_k(\mathbf{x})\, s_k(\mathbf{x}'), \tag{61}$$

which has the time domain expression

$$\tilde{G}(\mathbf{x}, \mathbf{x}'; t) = \frac{1}{2} \sum_{k} s_k(\mathbf{x})\, s_k(\mathbf{x}')\, \omega_k^{-3} \sin(\omega_k t)\, H(t). \tag{62}$$

Thus we determine that the ensemble cross correlation is given by

$$\langle C^{\alpha\beta}\rangle(\omega) = \rho^{-1}\, (2\,\omega)^{-1} \partial_{\omega} G(\mathbf{x}^{\alpha}, \mathbf{x}^{\beta}; \omega)\, S(\omega). \tag{63}$$

We make the important observation that in this case the ensemble cross correlation is not equal to the Green's function, but rather a derivative thereof. Distinctions between the ensemble cross correlation and the Green's function are also discussed in Tanimoto (2008) and Kimman & Trampert (2010).

The ensemble forward wavefield Φ^{α} (C8) is given by

$$\Phi^{\alpha}(\mathbf{x}, \omega) = \int G(\mathbf{x}, \mathbf{x}'; \omega)\, G^{*}(\mathbf{x}', \mathbf{x}^{\alpha}; \omega)\, S(\omega)\, \mathrm{d}^2\mathbf{x}' = \rho^{-1}\, (2\,\omega)^{-1} \partial_{\omega} G(\mathbf{x}, \mathbf{x}^{\alpha}; \omega)\, S(\omega), \tag{64}$$

and the ensemble adjoint wavefield (C11) is

$$\Phi^{\dagger\,\alpha\beta}(\mathbf{x}, \omega) = G(\mathbf{x}, \mathbf{x}^{\beta}; \omega)\, \Delta\langle C^{\alpha\beta}\rangle(\omega). \tag{65}$$

Now we have all the ingredients we need to calculate 2-D finite-frequency sensitivity kernels for noise cross-correlation measurements.

7.1 Rectangular drum

Let us consider a rectangular drum with dimensions $0 \leq x \leq L$ and $0 \leq y \leq W$. The eigenfrequencies are

$$\omega_{nm} = \pi\beta\, (n^2/L^2 + m^2/W^2)^{1/2}, \qquad n, m = 1, 2, 3, \ldots, \tag{66}$$

and the orthonormal eigenfunctions are

$$\mathbf{s}_{nm}(\mathbf{x}) = \frac{2}{\sqrt{\rho L W}} \sin(n\pi x/L) \sin(m\pi y/W), \tag{67}$$

if the displacement on the boundaries is kept zero, or

$$\mathbf{s}_{nm}(\mathbf{x}) = \frac{2}{\sqrt{\rho L W}} \cos(n\pi x/L) \cos(m\pi y/W), \tag{68}$$

if the stress is zero. Now we have an analytical expression for the Green's function and thus the ensemble cross correlation.

7.2 Numerical results

In this section we simulate membrane surface wave propagation, that is, a 2-D *SH* problem, using a 2-D spectral-element method, and we illustrate the construction of ensemble sensitivity kernels. For illustrative and instructional purposes, we only consider cross-correlation measurements between two distinct locations ('one-to-one'), similar to classic 'banana-doughnut' kernels where one computes one kernel for a given pick or arrival. As demonstrated by Tape *et al.* (2007), the beauty of the adjoint method is that it facilitates the calculation of the gradient of the misfit function based upon one adjoint simulation per earthquake. Basically, this gradient consists of a sum of weighted banana-doughnut kernels, with weights determined by the traveltime anomaly for a given pick. Thus, the calculation of the gradient of the misfit function scales linearly with the number of earthquakes, but is independent of the number of stations, components or picks. We shall see in Section 9 that a similar feature applies to ensemble noise cross-correlation kernels. In other words, we will discuss how to accommodate 'one-to-many' cross-correlation measurements later, but focus on 'one-to-one' measurements in the current section.

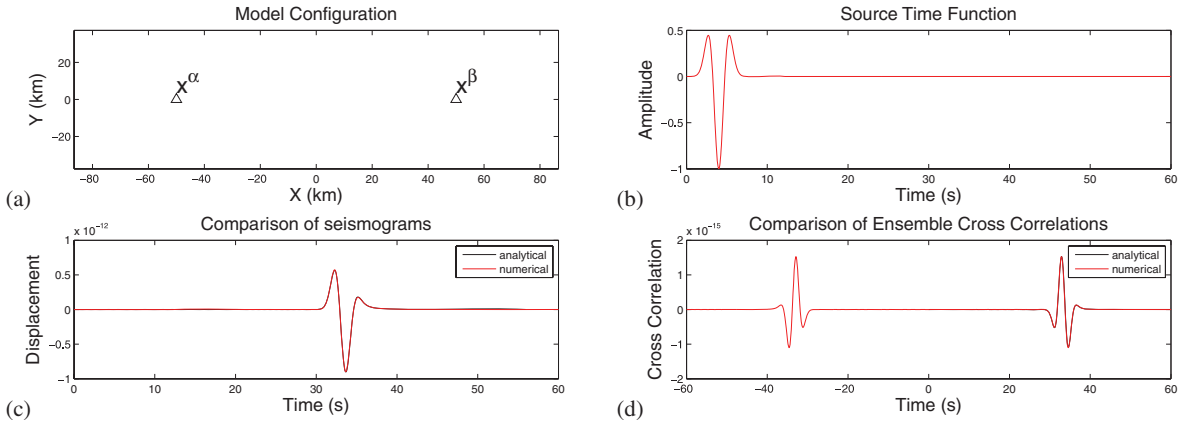

Figure 1. (a) Model set-up for the 2-D SH example. Two receivers, $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$, are denoted by triangles, and the distance between them is 100 km. The medium is homogeneous, with a mass density ρ of 3000 kg m^{-3} and a shear wave speed β of 3 km s^{-1}. All four boundaries are absorbing boundaries. (b) The source time function used in the benchmark is a Ricker wavelet with a half duration of 3 s. (c) Comparison of simulated and analytical seismograms observed at $\mathbf{x}^\beta$ due to a point source located at $\mathbf{x}^\alpha$. The black line is the analytical result obtained by summing the normal modes of a rectangular drum, whereas the red line is the numerical result calculated using a 2-D spectral-element code. (d) Comparison between analytical and simulated ensemble cross correlations from uniformly distributed noise. The analytical result is based upon normal-mode summation using (63), while (38) is used to obtain the numerical ensemble cross correlation.

7.2.1 Model set-up and ensemble cross correlation calculation

As shown in Fig. 1(a), the region of interest extends approximately 200 km by 80 km. Two observation points, $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$, denoted by triangles, are located 100 km away from each other. The model is homogeneous with density $\rho = 3000\,\mathrm{kg\,m^{-3}}$ and shear wave speed $\beta = 3\,\mathrm{km\,s^{-1}}$. All four boundaries are absorbing boundaries. The computational domain contains 4692 elements, resolving a 3 s Ricker wavelet. The time step is 0.04 s and the simulation length is ∼160 s. All 2-D simulations are carried out on a single processor.

Before calculating ensemble sensitivity kernels, we need to validate our ensemble cross-correlation calculation. For this benchmark, we assume a uniform noise distribution, as in the analytical expression (63). Figs 1(c) and (d) illustrate that analytical and simulated displacement seismograms and ensemble cross correlations fit perfectly.

7.2.2 Ensemble sensitivity kernel construction

In our first experiment, we consider uniformly distributed noise, within the approximately 200 km by 80 km model domain. The experimental set up is shown in Fig. 1. Fig. 2 depicts the ingredients and procedures for constructing $\langle K_\rho^{\alpha\beta}\rangle$, the first contribution to the ensemble density kernel $\langle K_\rho\rangle$, as defined by (35). Given the spectrum of the ensemble-averaged noise, we may obtain the source time function for the generating wavefield $\eta^\alpha(\mathbf{x},t)$, which is shown in Fig. 2(a). As expected, this source is narrowly concentrated around zero time.

First, we simulate the generating wavefield $\eta^\alpha(\mathbf{x},t)$ based upon (31), injecting a source with the spectrum of the ensemble-averaged noise at location $\mathbf{x}^\alpha$. This simulation is exactly the same as an ordinary earthquake simulation, but uses a point force instead of a moment tensor.

After obtaining the generating wavefield, we time-reverse it, based upon (32), to obtain the ensemble forward source $\mathbf{F}^\alpha(\mathbf{x},t)$ — the driving force of the ensemble forward wavefield $\mathbf{\Phi}^\alpha$. The driving force is shown in the first column of Fig. 2(e), where time increases from -40 s in the top row to 40 s in the bottom row. We see that the driving force collapses onto the point $\mathbf{x}^\alpha$ and vanishes around $t = 0$ s, which is a consequence of time-reversing the generating wavefield $\eta^\alpha(\mathbf{x},t)$, which starts at $t = 0$. The ensemble forward source consists of isotropic concentric circles centred on $\mathbf{x}^\alpha$.

Now that we have the ensemble forward source $\mathbf{F}^\alpha(\mathbf{x},t)$, we may generate the ensemble forward wavefield $\mathbf{\Phi}^\alpha$. This simulation differs from typical forward simulations in that the source is located where the noise is generated, in this 2-D case everywhere. From a practical perspective, one saves a 'movie' of the generating wavefield $\eta^\alpha(\mathbf{x},t)$ within the noise region at each time step. The simulated ensemble forward wavefield is shown in the second column of Fig. 2(e). Unlike the driving force shown in the first column, which vanishes for positive times, the ensemble forward field first converges on $\mathbf{x}^\alpha$, and then passes through the point and spreads out. If the model domain were the entire plane, and the noise uniform across this plane, then the ensemble forward wavefield $\mathbf{\Phi}^\alpha$ would consist of isotropic concentric circles centred on $\mathbf{x}^\alpha$, just like the ensemble forward source $\mathbf{F}^\alpha$. However, because we are restricting the simulation domain, the ensemble forward wavefield is not isotropic, reflecting the fact that, numerically, the noise distribution is not uniform—the noise exists only within our computational domain. In the 3-D examples we consider uniformly distributed noise at the scale of the globe; the 2-D examples in this section are mainly for illustrative and educational purposes.

Next, we need to calculate the ensemble adjoint wavefield $\mathbf{\Phi}^{\dagger\,\alpha\beta}$. To accomplish this, we need the ensemble adjoint source $\mathbf{F}^{\dagger\,\alpha\beta}$, which is related to measurements between observed and simulated ensemble cross correlations. The simulated ensemble cross correlation may be obtained based upon (38). Note that $\langle C^{\alpha\beta}\rangle$ involves $\mathbf{\Phi}^\beta$, not $\mathbf{\Phi}^\alpha$, but due to the symmetry (39) the calculation of $\mathbf{\Phi}^\beta$ is unnecessary. The

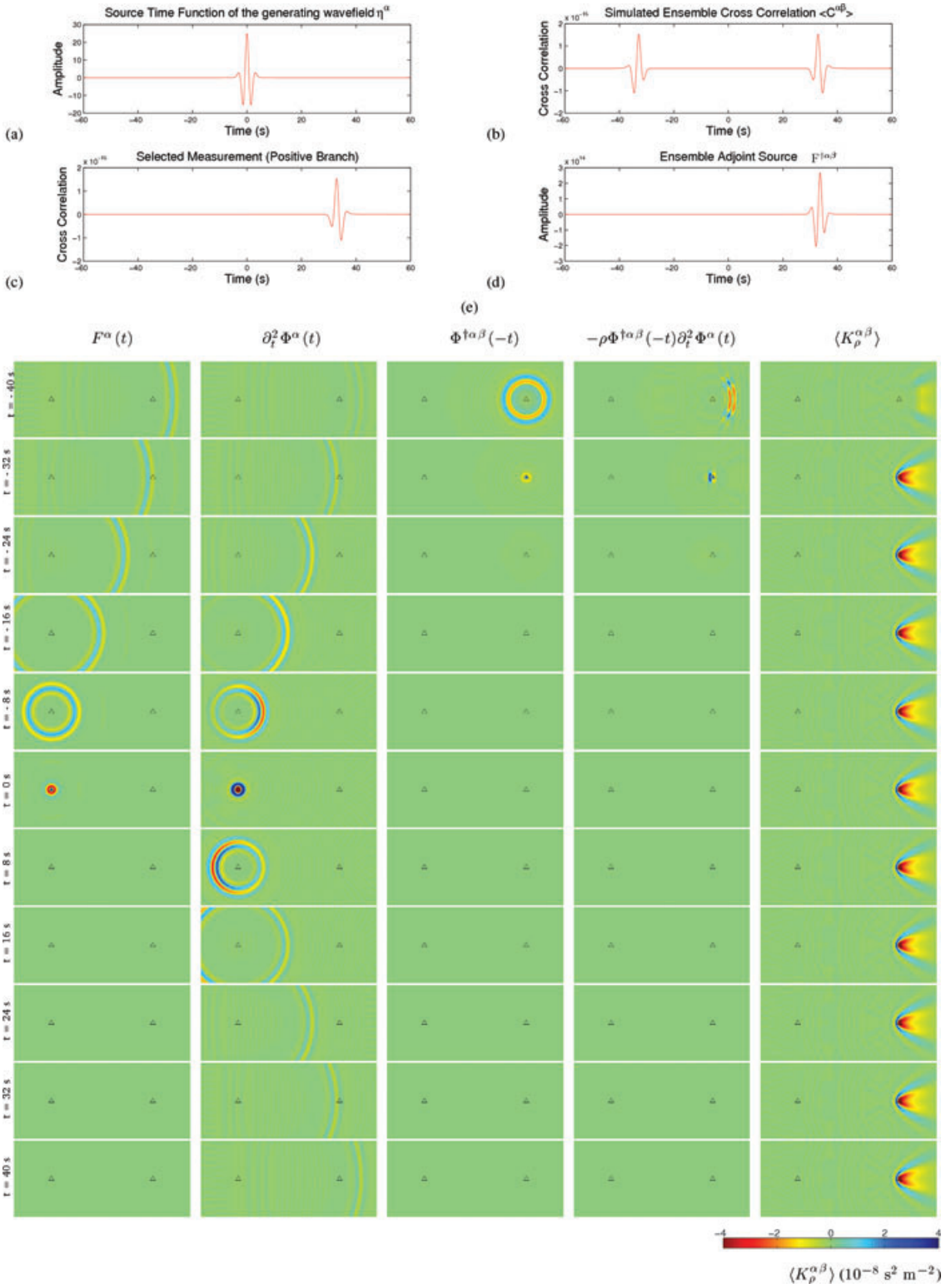

Figure 2. Ingredients and procedures for constructing the ensemble density sensitivity kernel $\langle K_\rho^{\alpha\beta}\rangle$ given by (35). The construction of the second contribution, $\langle K_\rho^{\beta\alpha}\rangle$, is shown in Fig. 3. (a) The source time function, with spectrum $S(\omega)$, used for the generating wavefield. (b) Synthetic ensemble cross correlation obtained from (38). (c) Time window of interests, in this case the positive branch of the ensemble cross correlation. Here we try to minimize the traveltime anomaly between observed and simulated ensemble cross correlations based upon (51). (d) The ensemble adjoint source $\mathbf{F}^{\dagger\alpha\beta}$ given by (52). (e) Sequence of interactions between the ensemble forward wavefield $\partial_t^2\Phi^\alpha$ and the ensemble adjoint field $\Phi^{\dagger\alpha\beta}$ during the construction of the ensemble density sensitivity kernel $\langle K_\rho^{\alpha\beta}\rangle$. First column: driving force F^α given by (32). Second column: ensemble forward wavefield $\partial_t^2\Phi^\alpha$ generated by the ensemble forward source F^α. Third column: ensemble adjoint wavefield $\Phi^{\dagger\alpha\beta}$ generated by the ensemble adjoint source (52). Fourth column: instantaneous interaction between the ensemble forward wavefield $\partial_t^2\Phi^\alpha$ and the ensemble adjoint wavefield $\Phi^{\dagger\alpha\beta}$. Fifth column: integration of the instantaneous interactions. Time increases from -40 (top row) to 40 s (bottom row) for all columns except the third one, since the ensemble adjoint wavefield needs to be time reversed.

resulting ensemble cross correlation is illustrated in Fig. 2(b). Due to the uniform noise distribution and constant model parameters, the ensemble cross correlation has an additional symmetry, namely

$$\langle C^{\alpha\beta}\rangle(t) = \langle C^{\alpha\beta}\rangle(-t). \tag{69}$$

Effectively, this additional symmetry requires translational invariance of the Green's function and a symmetric noise distribution.

We construct the ensemble adjoint source $\mathbf{F}^{\dagger\,\alpha\beta}$ by measuring some kind of difference between simulated and observed ensemble cross correlations. In global seismology, traveltime delay is a better measure of misfit between simulations and data than waveform difference. Therefore, we measure the traveltime delay between the observed and simulated ensemble cross correlations in the positive branch ($t > 0$), as shown in Fig. 2(c). This leads to the traveltime ensemble adjoint source (52). To obtain a banana-doughnut kernel, one sets $\Delta T^{\alpha\beta} = 1$. The traveltime ensemble adjoint source involves the first-order temporal derivative of the simulated ensemble cross correlation $\langle \dot{C}^{\alpha\beta}\rangle$, similar to the ordinary traveltime adjoint source, which depends on the synthetic velocity field $\dot{s}$ (Luo & Schuster 1990; Tromp *et al.* 2005). The corresponding ensemble adjoint source is shown in Fig. 2(d). Note that the amplitude of the ensemble adjoint source is many orders larger than the cross correlation, because of the normalization factor (53). It is this normalization factor that makes the ensemble kernels independent of the magnitude of the earthquake, or in this case the amplitude of the noise.

Given the ensemble adjoint source $\mathbf{F}^{\dagger\,\alpha\beta}$, calculation of the ensemble adjoint wavefield $\mathbf{\Phi}^{\dagger\,\alpha\beta}$ is straightforward. It is generated by injecting the ensemble adjoint source at some location—one of the receivers $\mathbf{x}^\beta$. The corresponding (time-reversed) ensemble adjoint wavefield is shown in the third column of Fig. 2(e). In other words, instead of $\mathbf{\Phi}^{\dagger\,\alpha\beta}(t)$, $\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)$ is plotted, because kernel construction (35) involves interactions between $\partial_t^2\mathbf{\Phi}^\alpha(t)$ and $\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)$. The ensemble adjoint wavefield is zero at the beginning from 40 to -24 s (note the reversed time), since the ensemble adjoint source kicks in very late.

The fourth column in Fig. 2(e) shows sequential interactions between the ensemble forward wavefield and the ensemble adjoint wavefield, that is, $-\rho\,\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)\cdot\partial_t^2\mathbf{\Phi}^\alpha(t)$. The fifth column in Fig. 2(e) is the integration of the interaction based upon (35). We see a hyperbolic fan tail starting at $\mathbf{x}^\beta$ and expanding away from $\mathbf{x}^\alpha$. In this and all following kernel figures, red denotes negative values, whereas blue represents positive values. All kernels are plotted on the same scale, except for the source strength kernels.

Fig. 3 depicts the ingredients and procedures for constructing $\langle K_\rho^{\beta\alpha}\rangle$, the second contribution to the ensemble density kernel $\langle K_\rho\rangle$, as defined by (36). Since Fig. 3 is in the same format as Fig. 2, we only highlight their differences. The first difference lies in the adjoint source, that is, Fig. 3(d). As implied by (17) & (33), the new adjoint source is reversed in time compared to Fig. 2(d). In Fig. 2, the ensemble forward source is located at $\mathbf{x}^\alpha$ and the ensemble adjoint source is injected at $\mathbf{x}^\beta$. For the second contribution, the locations of the two ensemble sources are interchanged, as shown in the first three columns of Fig. 3(e). The pattern of the ensemble forward wavefield remains basically the same, but the ensemble adjoint wavefield lasts much longer, because the adjoint source initiates earlier, that is, in the branch of negative times instead of positive times. As a consequence, the ensemble adjoint wavefield interacts with the ensemble forward wavefield more effectively, painting not only the 'jet' we see in Fig. 2(e), but also the classic 2-D banana-doughnut shape.

The final ensemble density sensitivity kernel associated with a traveltime measurement in the positive ensemble cross-correlation branch is obtained by summing the two contributions (35) and (36). In Fig. 4 we illustrate this summation process. Figs 4(a) and (b) collect the two contributions we obtained in Figs 2 and 3. Fig. 4(c) is the summation of (a) & (b). Interestingly, although still visible, the 'jets' in the two contributions almost cancel, leaving only the classic banana-doughnut kernel.

Other kernels may be constructed in a similar fashion.

7.2.3 Validation of the ensemble sensitivity kernel

Intuitively, the final ensemble density kernel shown in Fig. 4(c) makes sense: for uniformly distributed noise, it should be largely controlled by structure between the two receivers, as generally assumed in ambient noise tomography. In this section we quantitatively assess the validity of this result.

The traveltime, T, may be related to the wave speed structure along the ray path, v, as follows:

$$T = \int_{\text{ray}} v^{-1}\,\mathrm{d}l. \tag{70}$$

Fermat's principle tells us that the traveltime anomaly, δT, may be related to wave speed perturbations, δv, along the original ray by

$$\delta T = -\int_{\text{ray}} v^{-2}\,\delta v\,\mathrm{d}l. \tag{71}$$

Given a homogeneous material and a uniform perturbation, that is, $v = $ const. and $\delta v = $ const., we have:

$$\delta T = -L v^{-2}\,\delta v, \tag{72}$$

where L is the distance between the two points. In our *SH* problem, $v = \sqrt{\mu/\rho}$. Suppose we keep μ fixed; then the perturbation in v is only dependent on $\delta\rho$.

$$\delta v = -\frac{1}{2}\,\frac{v}{\rho}\,\delta\rho. \tag{73}$$

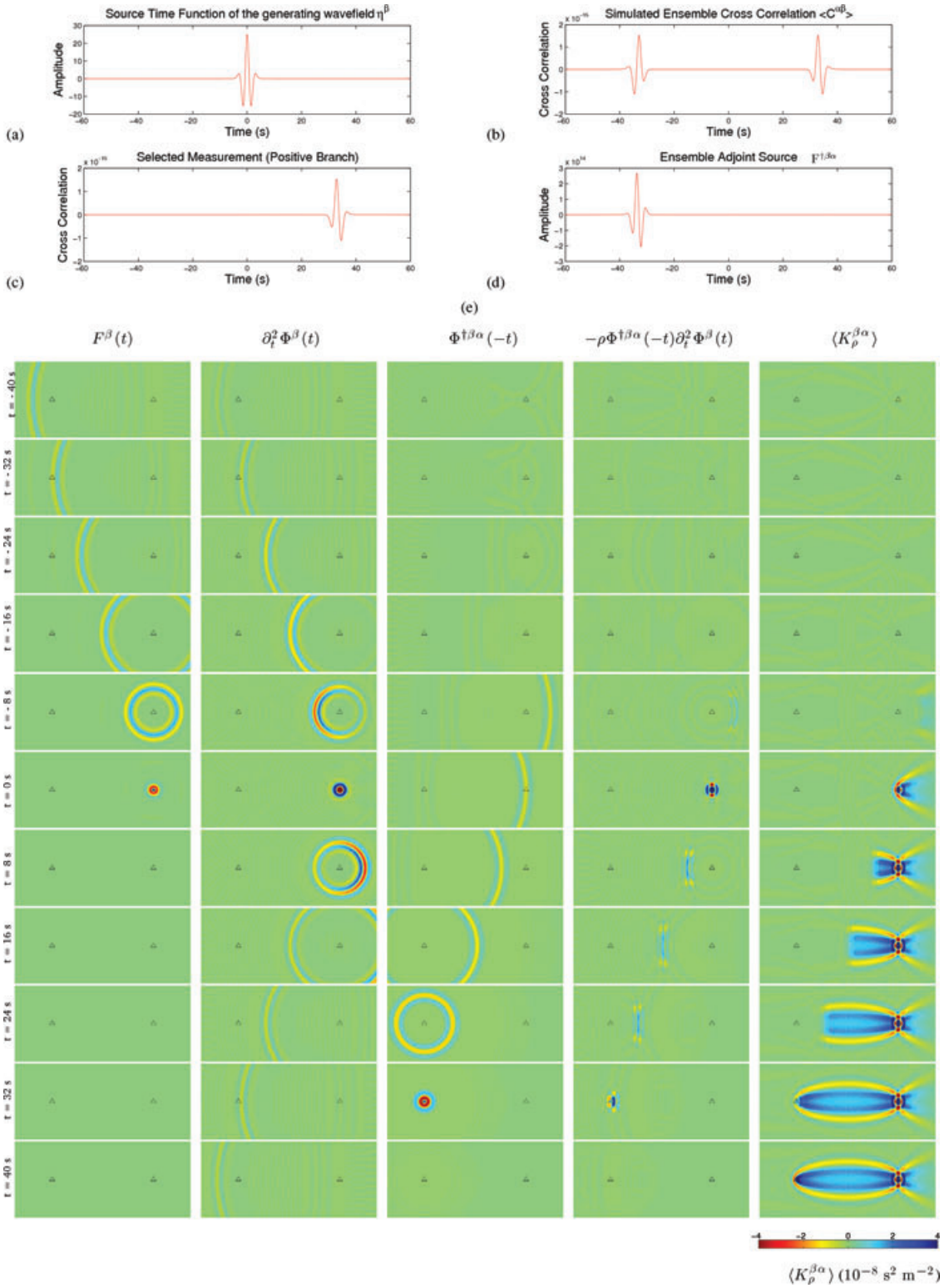

Figure 3. Ingredients and procedures for constructing the ensemble density kernel $\langle K_\rho^{\beta\alpha}\rangle$ defined in (36), that is, the complement of Fig. 2. (a)–(c) The same as in Fig. 2. Note that we are still measuring the traveltime delay in the positive branch, as illustrated in (c). (d) The ensemble adjoint source $\mathbf{F}^{\dagger\,\beta\alpha}$ takes the form (52). Upon comparing Figs 2(d) and 3(d), we see that the temporal dependence of the adjoint source $\mathbf{F}^{\dagger\,\beta\alpha}$ is reversed relative to $\mathbf{F}^{\dagger\,\alpha\beta}$. Also note that the location of the adjoint source is changed from $\mathbf{x}^\beta$ to $\mathbf{x}^\alpha$, as indicated by exchanging α and β in (52). (e) Sequence of interactions between the ensemble forward field $\partial_t^2\Phi^\beta$ and the ensemble adjoint field $\Phi^{\dagger\,\beta\alpha}$ during the construction of the ensemble sensitivity kernel $\langle K_\rho^{\beta\alpha}\rangle$. First column: driving force F^β given by (32). Second column: ensemble forward wavefield $\partial_t^2\Phi^\beta$ generated by the ensemble forward source F^β. Third column: ensemble adjoint wavefield $\Phi^{\dagger\,\beta\alpha}$ generated by the ensemble adjoint source (52). Fourth column: instantaneous interaction between the ensemble forward wavefield $\partial_t^2\Phi^\beta$ and the ensemble adjoint wavefield $\Phi^{\dagger\,\beta\alpha}$. Fifth column: integration of the instantaneous interactions. Time increases from -40 s (top row) to 40 s (bottom row) for all columns except the third one, since the ensemble adjoint wavefield needs to be time reversed.

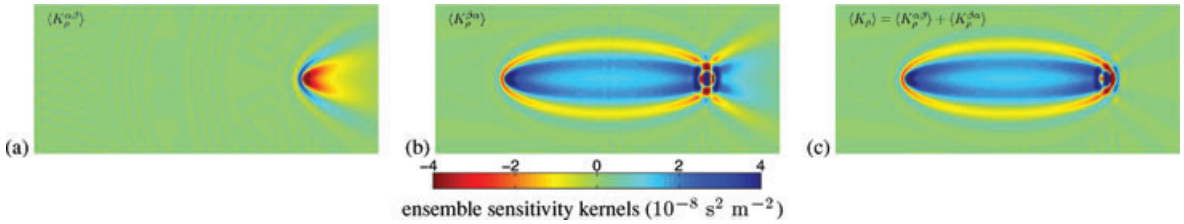

Figure 4. The final ensemble density kernel $\langle K_\rho \rangle$ shown in (c) is obtained by summing the two contributions $\langle K_\rho^{\alpha\beta} \rangle$ shown in (a) & $\langle K_\rho^{\beta\alpha} \rangle$ shown in (b). Except for the weak tails, the ensemble sensitivity kernel $\langle K_\rho \rangle$ looks similar to a classic banana-doughnut kernel in this simple 2-D example.

Thus, when only the density is uniformly perturbed, traveltime and density perturbations are related via

$$\delta T = \frac{1}{2} \frac{L}{v} \delta \ln \rho. \tag{74}$$

We also have the finite-frequency relationship between traveltime and density perturbations:

$$\delta T = \int \langle K_\rho \rangle \, \delta \ln \rho \, \mathrm{d}^2\mathbf{x}, \tag{75}$$

where the shear modulus, μ, is kept fixed. Under the assumption that the density perturbation is uniform, this expression reduces to

$$\delta T = \delta \ln \rho \int \langle K_\rho \rangle \, \mathrm{d}^2\mathbf{x}. \tag{76}$$

Therefore, upon comparing (74) and (76), we expect that

$$\int \langle K_\rho \rangle \, \mathrm{d}^2\mathbf{x} = \frac{1}{2} \frac{L}{v}. \tag{77}$$

Given the distance L and the wave speed v, we may evaluate how accurately the numerical ensemble density kernel $\langle K_\rho \rangle$ satisfies (77). It turns out that the left-hand side of (77) equals 16.7001 s, while the right-hand side has an analytical value of 16.6667 s. The discrepancy is merely 0.2 per cent, thereby validating our ensemble kernel calculation, in particular the combination of (35) and (36), as illustrated in Fig. 4; the wrong sign for (35) would enhance rather than diminish the signature of the jet, leading to the wrong value on the left-hand side of (77).

7.2.4 Ensemble sensitivity kernel for the negative cross-correlation branch

The ensemble density sensitivity kernels discussed in the preceding sections are associated with traveltime measurements in the positive ensemble cross-correlation branch; from hereon, we add a superscript '+' to denote such ensemble kernels, for example, $\langle K_\rho \rangle^+$. Of course, we may measure traveltime delays in the negative branch as well, resulting in an ensemble density kernel denoted by $\langle K_\rho \rangle^-$. Generally, one only makes measurements in the positive branch, since the symmetry (17) guarantees that negative branch measurements contain no additional information, provided the positive branches of $\langle C^{\alpha\beta} \rangle$ and $\langle C^{\beta\alpha} \rangle$ are both measured.

Figs 5(a) and (c) reiterate what we have already illustrated, that is, the ensemble density kernel $\langle K_\rho \rangle^+$ associated with measurements in the positive branch of the ensemble cross correlation $\langle C^{\alpha\beta} \rangle$. Figs 5(b) and (d) show the ensemble density kernel $\langle K_\rho \rangle^-$ associated with traveltime delay measurements in the negative branch of $\langle C^{\alpha\beta} \rangle$. We identify two main differences. First, the weak 'jets' change positions. This is natural, since measurements in the positive and negative branches correspond to waves travelling from $\mathbf{x}^\beta$ to $\mathbf{x}^\alpha$ and from $\mathbf{x}^\alpha$ to $\mathbf{x}^\beta$, respectively. Second, the signs of the two ensemble sensitivity kernels differ. This is because of our definition of the traveltime delay:

$$\Delta T^{\alpha\beta} = T^{\alpha\beta}_{\mathrm{sim}} - T^{\alpha\beta}_{\mathrm{obs}}, \tag{78}$$

regardless of which branch we are measuring. This definition, which is the same as in earthquake tomography, is convenient in the sense that we are able to treat ensemble cross correlations exactly the same as seismograms without worrying about in which branch we make our measurement. Therefore, a positive traveltime delay in the positive branch becomes a negative delay in the negative branch, provided the model is translationally invariant. Fig. 5(e) illustrates such an example: a positive traveltime delay in the positive branch means the wave speed model is too slow, whereas a positive traveltime delay in the negative branch indicates the model is too fast—exactly the same as implied by the signs of the ensemble kernels shown in Figs 5(c) and (d).

7.2.5 Other types of ensemble sensitivity kernels

Thus far we have only shown ensemble density kernels $\langle K_\rho \rangle$. The ensemble moduli and wave speeds kernels are similar to density kernels, and they will be illustrated in a 3-D example later, to avoid redundancy. In this section we show two additional ensemble sensitivity kernels.

For illustrative purposes, we assume that the frequency dependence of $S(\mathbf{x}, \omega)$ defined by (15) is fixed, but that its spatial distribution may vary, $\delta S(\mathbf{x}, \omega) = S(\mathbf{x}, \omega) \, \delta\sigma(\mathbf{x})$, where $\delta\sigma$ represents a perturbation in the spatial distribution of the noise. Then we have from (29)

$$\langle \delta\Sigma \rangle = \delta\sigma \langle K_\sigma \rangle, \tag{79}$$

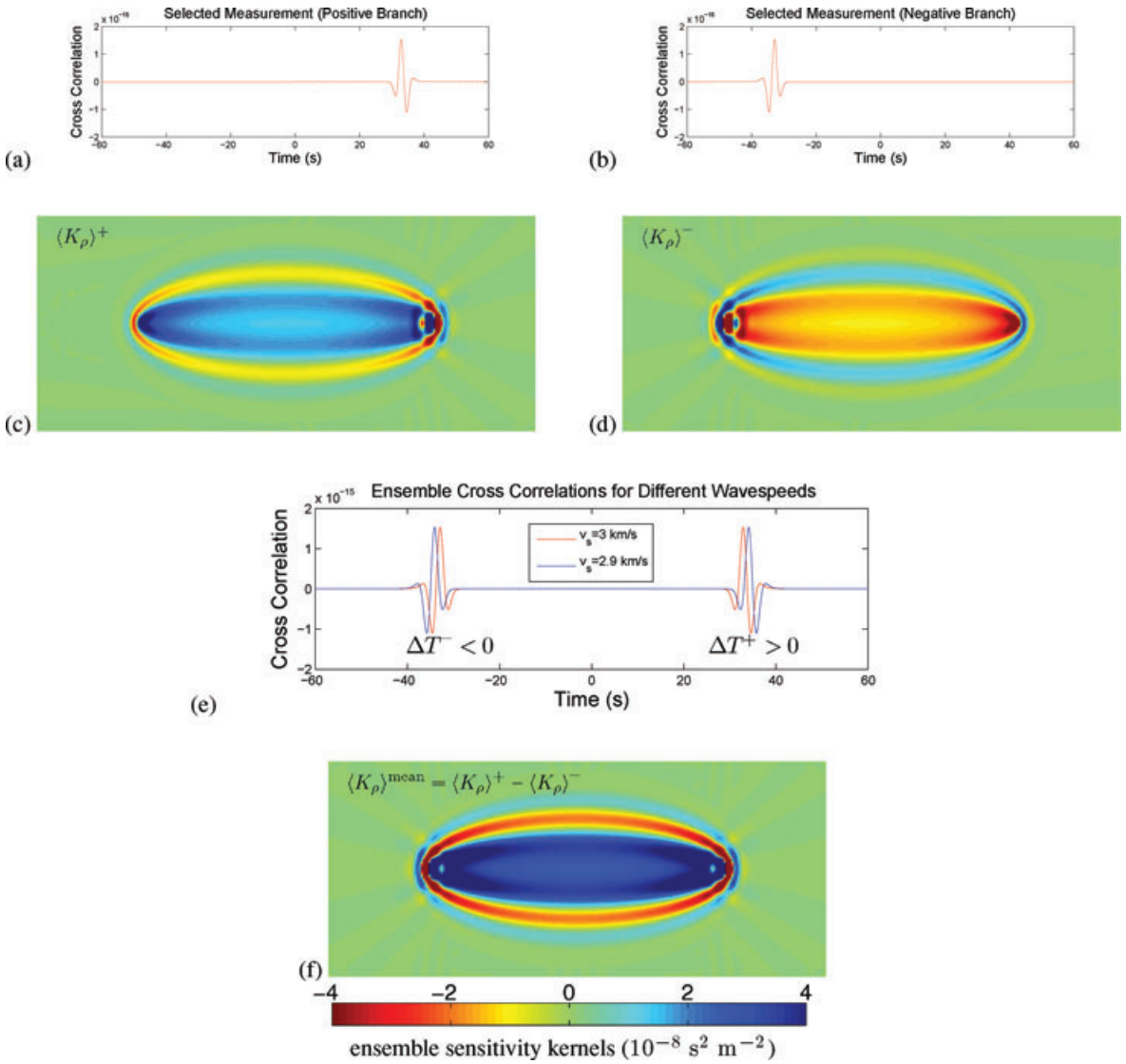

Figure 5. Ensemble density kernels associated with positive and negative ensemble cross-correlation branch measurements, $\langle K_\rho \rangle^+$ and $\langle K_\rho \rangle^-$, respectively: (a) & (c) repeat what we have already seen, that is, traveltime measurements made in the positive branch and the corresponding ensemble density sensitivity kernel $\langle K_\rho \rangle^+$; (b) & (d) involve negative branch measurements and the resulting ensemble density kernel $\langle K_\rho \rangle^-$. The signs of (c) & (d) are different, due to the definition of the traveltime delay. As shown in (e), when $\Delta T^+ > 0$ for the positive branch, the negative branch has $\Delta T^- = -\Delta T^+ < 0$, provided the Green's tensor is translationally invariant and the noise distribution is symmetric. If these strong constraints are met, as in our simple example, the ensemble kernels associated with negative and positive branches may be combined, forming a completely symmetric kernel, as shown in (f). When $\Delta T^- \neq -\Delta T^+$, the ensemble kernel $\langle K_\rho \rangle^{\text{mean}}$ represents the sensitivity with respect to the mean traveltime anomaly associated with the positive and negative branches, that is, $\Delta T^{\text{mean}} = (\Delta T^+ - \Delta T^-)/2$.

where $\langle K_\sigma \rangle$ denotes the 'ensemble source strength' kernel:

$$\langle K_\sigma \rangle = \int \left[\mathbf{\Phi}^{\dagger\, \alpha\beta}(-t) \cdot \mathbf{F}^\alpha(t) + \mathbf{\Phi}^{\dagger\, \beta\alpha}(-t) \cdot \mathbf{F}^\beta(t) \right] \mathrm{d}t. \tag{80}$$

We consider the source strength kernel (80) measured in both the positive and negative ensemble cross-correlation branches, $\langle K_\sigma \rangle^+$ and $\langle K_\sigma \rangle^-$, respectively.

Next, we consider the effects of adding a small damping term to the left-hand side of the membrane wave equation:

$$\rho \partial_t^2 s + \Gamma\, \partial_t s = \boldsymbol{\nabla} \cdot (\mu \boldsymbol{\nabla} s) + f. \tag{81}$$

The resulting ensemble sensitivity kernel with respect to the damping is given by

$$\langle K_\Gamma \rangle = -\int \left[\mathbf{\Phi}^{\dagger\, \alpha\beta}(-t) \cdot \partial_t \mathbf{\Phi}^\alpha(t) + \mathbf{\Phi}^{\dagger\, \beta\alpha}(-t) \cdot \partial_t \mathbf{\Phi}^\beta(t) \right] \mathrm{d}t. \tag{82}$$

In helioseismology, Gizon & Birch (2002) used a simple damping mechanism such as this to capture dissipation in the Sun.

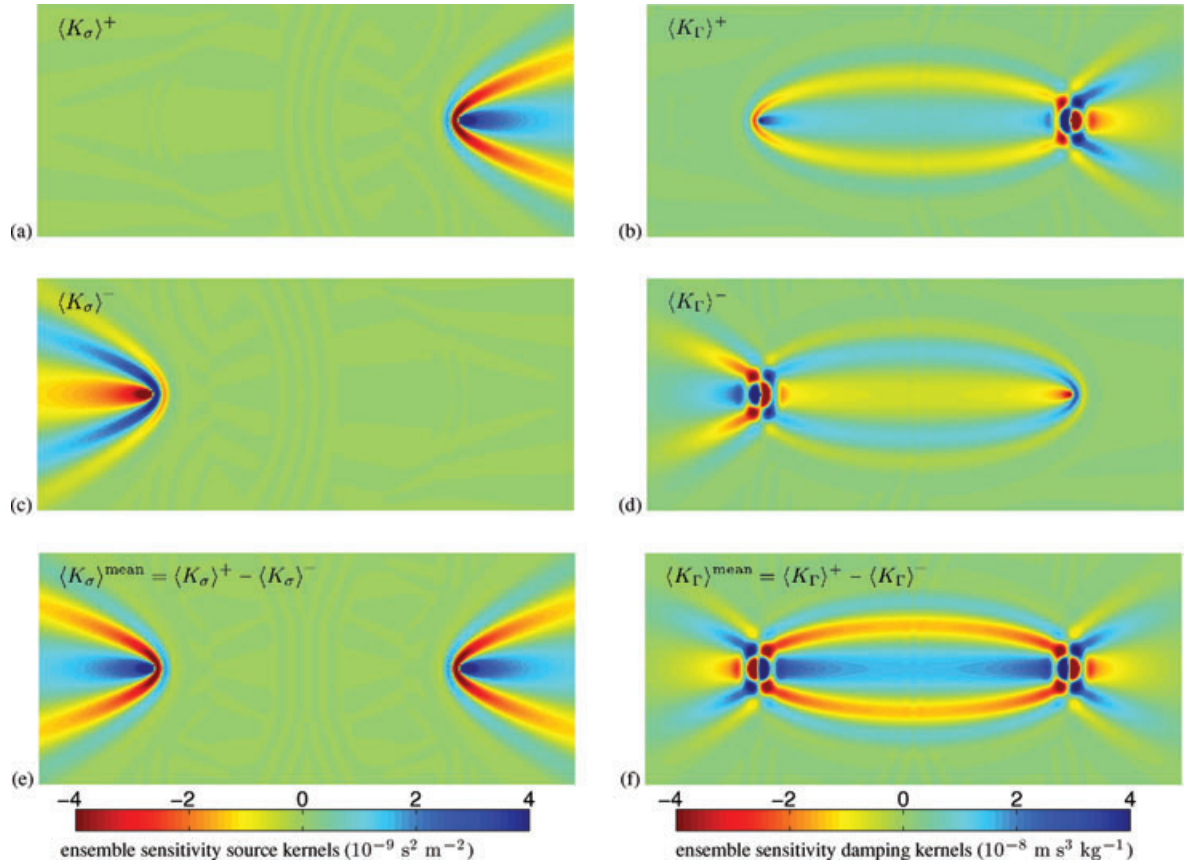

Figure 6. Ensemble source strength kernel $\langle K_\sigma \rangle$ given by (80), and ensemble damping kernel $\langle K_\Gamma \rangle$ given by (80). $\langle K_\sigma \rangle^+$ and $\langle K_\Gamma \rangle^+$ denote positive ensemble cross-correlation branch measurements, and $\langle K_\sigma \rangle^-$ and $\langle K_\Gamma \rangle^-$ denote negative branch measurements.

Fig. 6 illustrates these two types of kernels for positive and negative ensemble cross-correlation branch traveltime delay measurements. Unlike the ensemble density kernel, ensemble source strength and damping kernels exhibit strong hyperbolic 'jets'. Gizon & Birch (2002) also computed ensemble sensitivity kernels with respect to source strength and damping using analytical Green's functions for the Sun. Their results are in good qualitative agreement with ours.

7.2.6 Non-uniform noise distribution

In the 2-D examples considered thus far, we have assumed that the distribution of the noise is uniform. In this section we allow for a non-uniform distribution of the noise (see e.g. Hanasoge *et al.* 2008, for a discussion of non-uniform noise in helioseismology). Fig. 7 illustrates the process for the first contribution to the ensemble density kernel $\langle K_\rho^{\alpha\beta} \rangle^-$, associated with a measurement in the negative ensemble cross-correlaton branch of $\langle C^{\alpha\beta} \rangle$. As indicated in Fig. 7(a), the spectrum of the ensemble-averaged noise is the same as before, but it is restricted to the elliptical-shaped area outlined by the thin red contour in Fig. 7(e). Because the waves generated by the noise are pre-dominantly travelling from $\mathbf{x}^\alpha$ to $\mathbf{x}^\beta$, the ensemble cross correlation $\langle C^{\alpha\beta} \rangle$ is dominated by its negative branch, as shown in Fig. 7(b). We still have the symmetry $\langle C^{\alpha\beta} \rangle(t) = \langle C^{\beta\alpha} \rangle(-t)$, but because we no longer have a symmetric noise distribution, we no longer have the symmetry $\langle C^{\alpha\beta} \rangle(t) = \langle C^{\alpha\beta} \rangle(-t)$. A traveltime delay measurement in the negative branch of the ensemble cross correlation $\langle C^{\alpha\beta} \rangle$, as shown in Fig. 7(c), leads to the ensemble density kernel $\langle K_\rho^{\alpha\beta} \rangle^-$ via the process illustrated in Fig. 7(e).

Fig. 8 illustrates the process for the second contribution $\langle K_\rho^{\beta\alpha} \rangle^-$, which is associated with a measurement in the negative ensemble cross-correlation branch of $\langle C^{\alpha\beta} \rangle$. The final kernel $\langle K_\rho \rangle^- = \langle K_\rho^{\alpha\beta} \rangle^- + \langle K_\rho^{\beta\alpha} \rangle^-$ is shown in Fig. 9(a) and should be contrasted with the result for uniform noise shown in Fig. 5(d).

For a constant density perturbation in a homogeneous density and wave speed model, we may again use (77) to assess the accuracy of the ensemble density kernel $\langle K_\rho \rangle^-$, shown in Fig. 9(a). In this case we find that the left-hand side of (77) equals 16.1506 s, compared to an analytical value of 16.6667 s, a 3.1 per cent difference. This difference reflects the fact that the we do not know the proper analytical value of the ray theoretical traveltime when the noise is non-uniform, that is, the analytical expression becomes invalid.

Fig. 9(b) shows the ensemble source strength kernel $\langle K_\sigma \rangle^-$ and Fig. 9(c) shows the ensemble damping kernel $\langle K_\Gamma \rangle^-$ for non-uniform noise. These figures clearly illustrate that non-uniform noise has a profound impact on the interpretation of ensemble cross-correlation data in terms of structural variations.

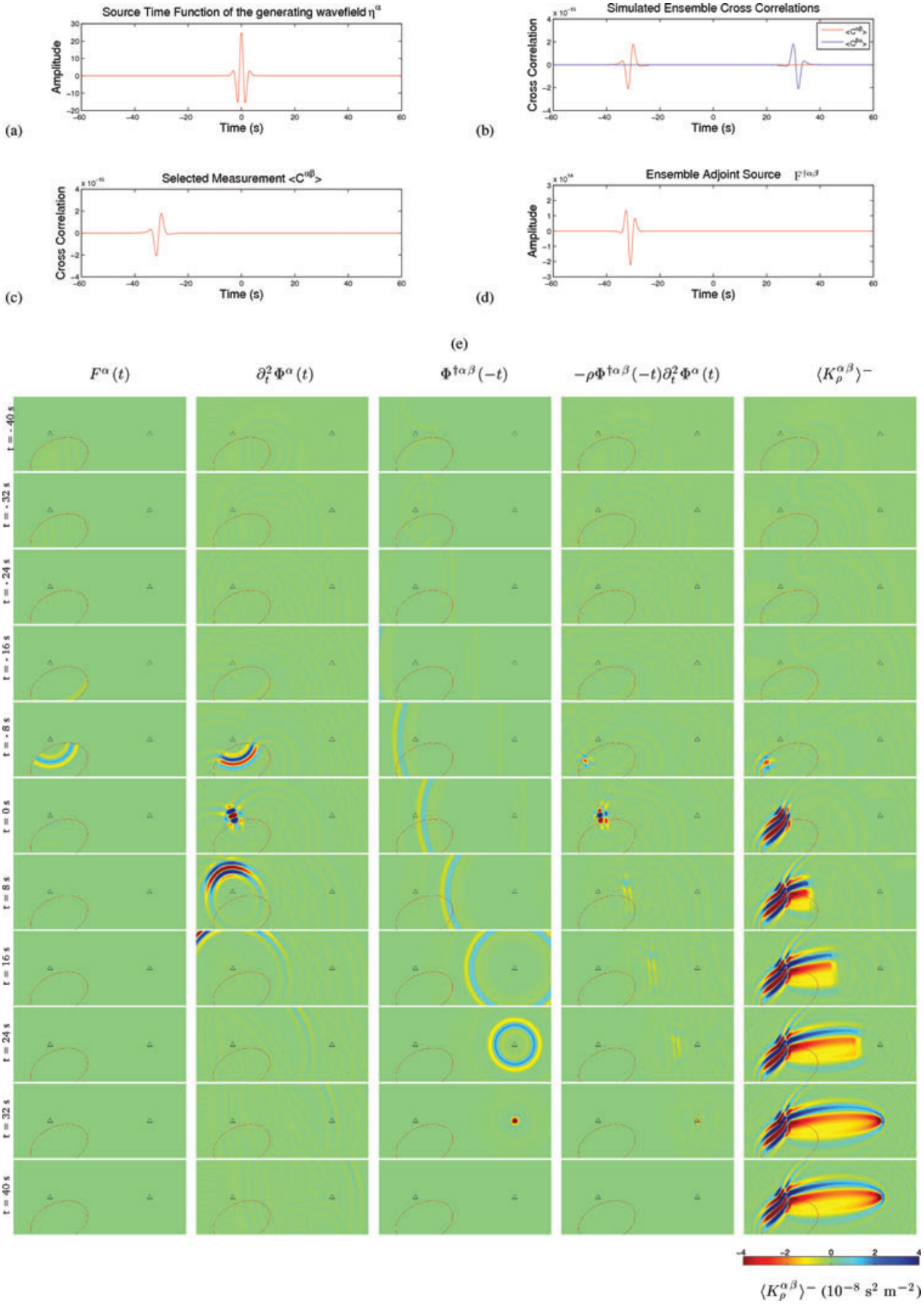

Figure 7. Similar to Fig. 2, but for non-uniform noise sources and a measurement in the negative branch of ensemble cross correlation $\langle C^{\alpha\beta}\rangle$. Construction of the contribution $\langle K_\rho^{\alpha\beta}\rangle^-$. The noise is restricted to the elliptical-shaped area outlined by the thin red contour.

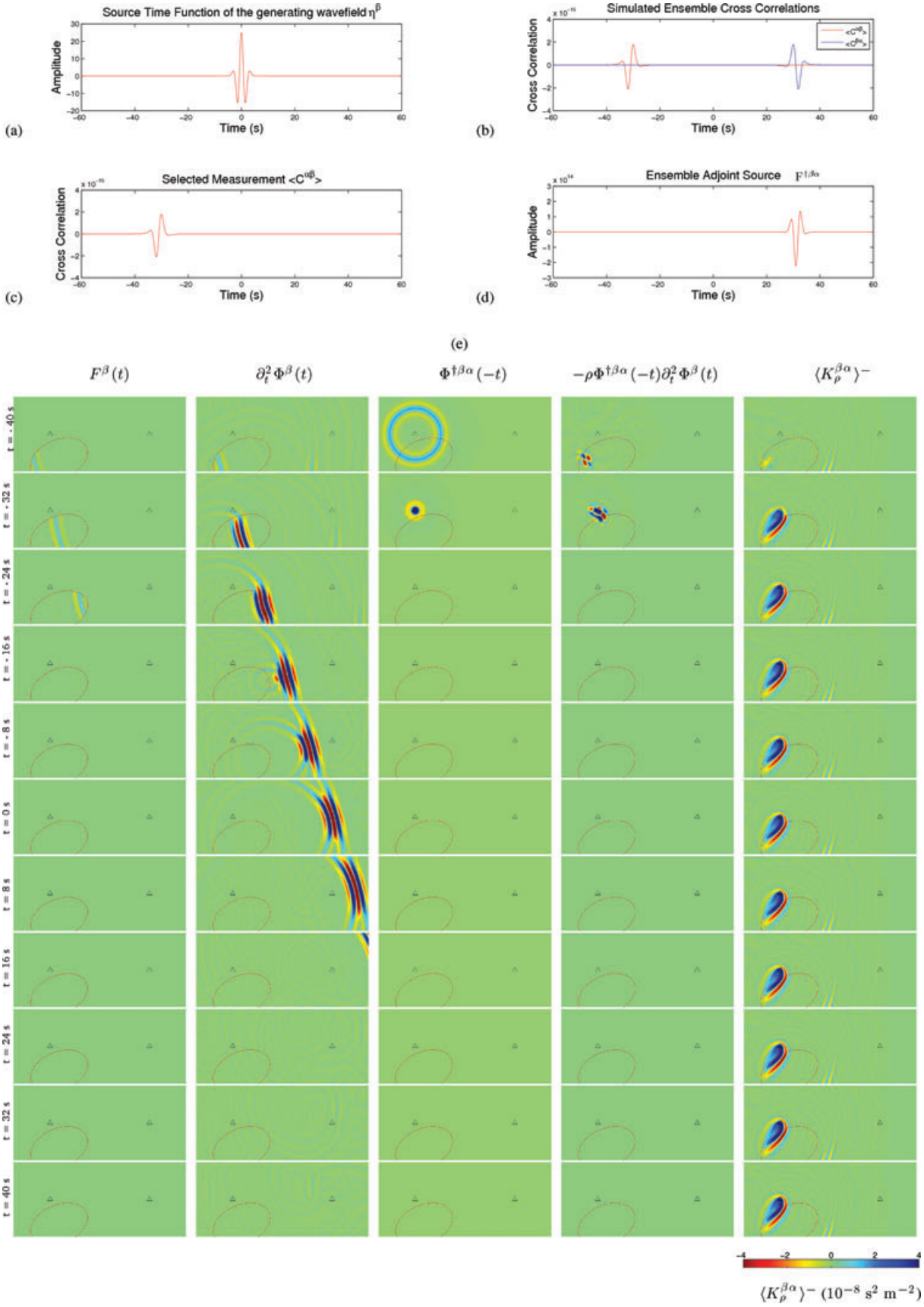

Figure 8. Similar to Fig. 3, but for non-uniform noise sources and a measurement in the negative branch of ensemble cross correlation $\langle C^{\alpha\beta}\rangle$. Construction of the contribution $\langle K_\rho^{\beta\alpha}\rangle^-$. The noise is restricted to the elliptical-shaped area outlined by the thin red contour in each panel in (e).

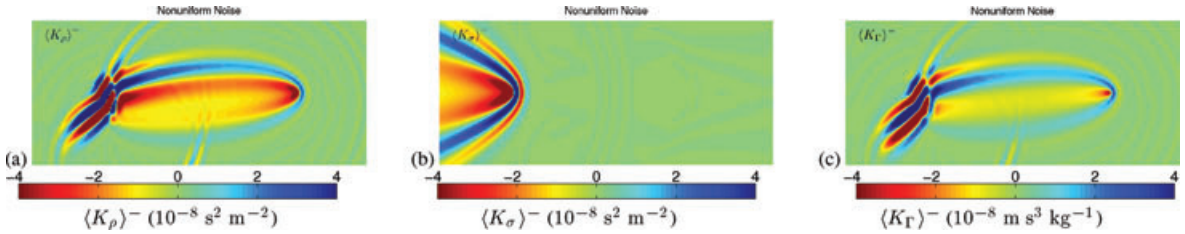

Figure 9. (a) $\langle K_\rho \rangle^- = \langle K_\rho^{\alpha\beta} \rangle^- + \langle K_\rho^{\beta\alpha} \rangle^-$ for non-uniform noise and a measurement in the negative branch of the ensemble cross correlation $\langle C^{\alpha\beta} \rangle$. It is obtained by adding $\langle K_\rho^{\alpha\beta} \rangle^-$ shown in the bottom right-hand side of Fig. 7(e) and $\langle K_\rho^{\beta\alpha} \rangle^-$ shown in the bottom right-hand side of Fig. 8(e). This ensemble kernel should be contrasted with the result for uniform noise shown in Fig. 5(d). (b) Ensemble source strength sensitivity kernel $\langle K_\sigma \rangle^-$ for non-uniform noise. (c) Ensemble damping sensitivity kernel $\langle K_\Gamma \rangle^-$ for non-uniform noise.

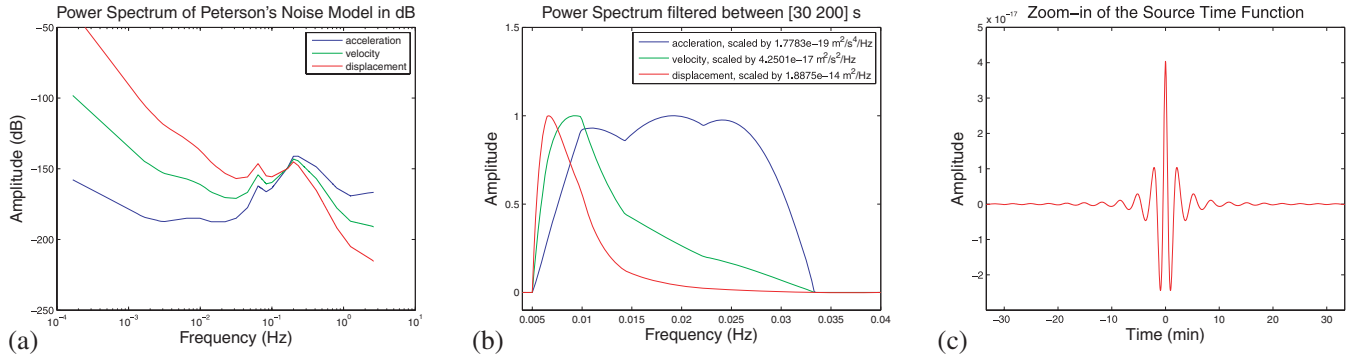

Figure 10. (a) Peterson's noise model (Peterson 1993). The spectrum of the ensemble-averaged noise in dB is shown between 0.1 s and 10 000 s. (b) Spectrum filtered between 30 and 200 s, reflecting the resolution of our simulations. (c) Source time function corresponding to the noise spectrum. The source for the ensemble forward simulation is localized around zero time.

8 3-D EXAMPLES

In this section, we illustrate 3-D noise cross-correlation sensitivity kernels. Only final ensemble sensitivity kernels are shown, since the constructions of such kernels is illustrated in the 2-D examples. We first consider regional problems, involving one $90^\circ \times 90^\circ$ 'chunk' of the cubed sphere (Komatitsch & Tromp 2002a,b). Next, to connect with classic ambient noise tomography, we explore the behaviour of ensemble sensitivity kernels in global simulations. As in the 2-D case, we consider two types of noise distributions: noise uniformly distributed across the surface of the model domain, and non-uniform noise restricted to a particular geographical area. We assume that the excitation is in the vertical direction, that is, that S_{jm} in (14) is of the form $S_{jm}(\mathbf{x}, \omega) = S(\mathbf{x}, \omega)\, \hat{n}_j(\mathbf{x})\, \hat{n}_m(\mathbf{x})$, where $\hat{\mathbf{n}}$ is the unit outward normal. This noise source generates Rayleigh surface waves; it is straightforward to consider non-vertical noise sources, which would generate Love and Rayleigh waves. The vertical component of ground motion recorded at receivers is used to compute cross correlations, following the usual approach in ambient noise tomography. Note, however, that one is free to choose which components to cross correlate, that is, $\hat{\mathbf{v}}^\alpha$ and $\hat{\mathbf{v}}^\beta$ in (8) may be chosen arbitrarily.

We use the Peterson noise model (Peterson 1993), depicted in Fig. 10(a), to generate the spectrum S of the ensemble-averaged noise. Since our simulations are band-limited, we filter the spectrum between 30 and 200 s, as shown in Fig. 10(b). The lower corner of the filter is controlled by the resolution of our numerical simulations, a period of approximately 30 s. For regional scale problems we may readily consider waves with a shortest period of approximately 2 s (e.g. Komatitsch *et al.* 2004; Tape *et al.* 2009).

All kernels involve spectral-element simulations of wave propagation in the Preliminary Reference Earth Model (PREM, Dziewonski & Anderson 1981), which preserves translational invariance of the Green's tensor. Note that the spectral-element solver SPECFEM3D_GLOBE (www.geodynamics.org) is capable of dealing with arbitrary 3-D earth models; hence our methods are not restricted by the choice of model.

The 3-D simulations require parallel processing based upon message passing. For regional (one chunk) simulations we use 100 cores, whereas for global (six chunk) simulations we use 600 cores. At the surface, the spectral-element mesh contains 160 elements along a side of a single chunk, for a total of 234 000 and 1 412 000 regional and global spectral elements, respectively. The simulations are accurate for periods of 27 s and longer, and require a time step of 0.19 s.

8.1 Regional simulations

The computational domain is one 'chunk' of the cubed sphere, with a free surface at the top and absorbing boundaries elsewhere.

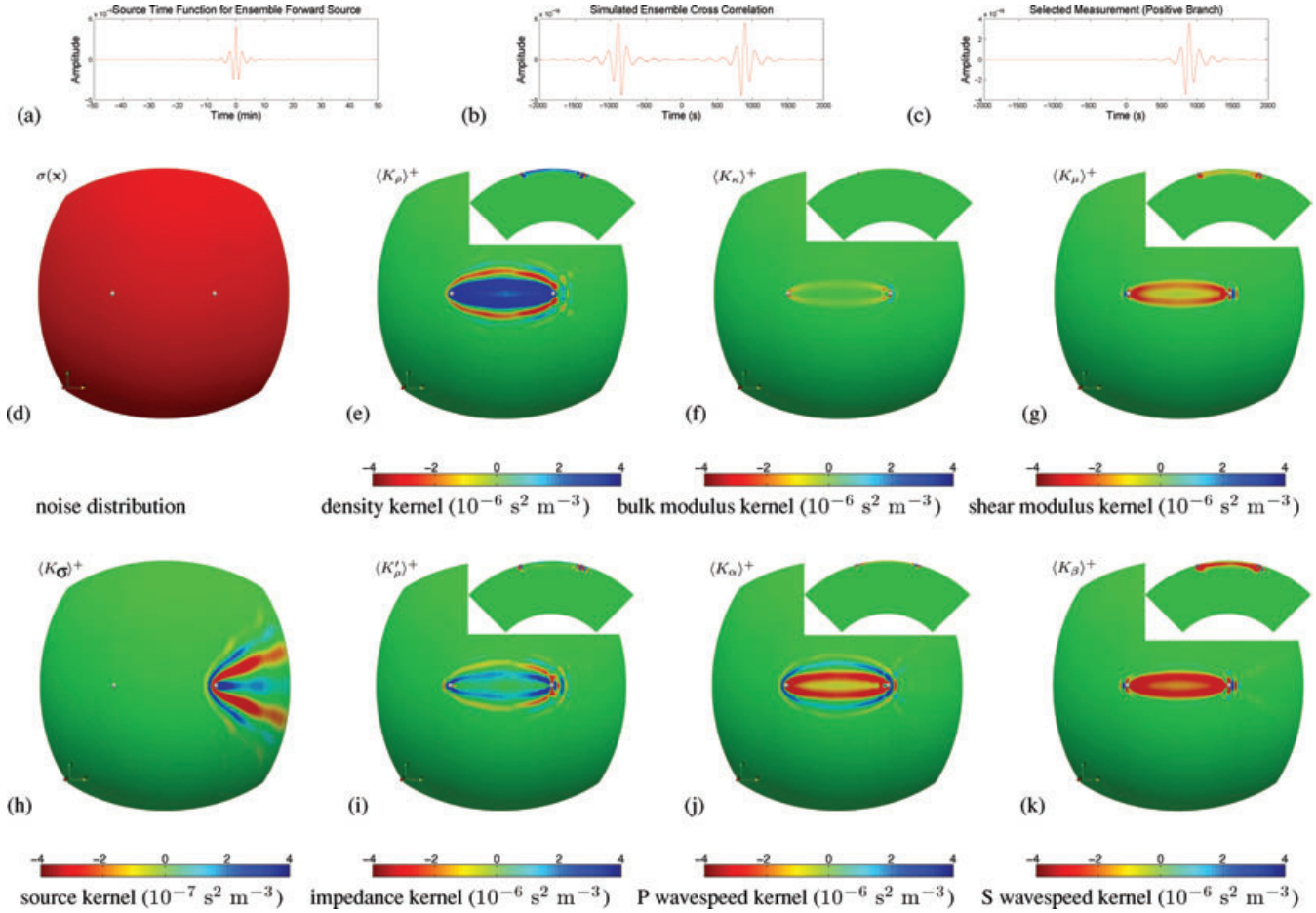

Figure 11. (a) Source time function used to generate the ensemble forward wavefield, adapted from Fig. 10(c). (b) Simulated ensemble cross correlation. (c) Measurement in the positive cross-correlation branch. (d) Map view of the evenly distributed noise (in red) and receiver geometry (white dots). (e)–(k) Ensemble sensitivity kernels. Kernels (e), (f) and (g) correspond to using the model parameters density ρ, bulk modulus κ and shear modulus μ, whereas kernels (i), (j) and (k) correspond to using density ρ, compressional wave speed α and shear wave speed β. The ensemble density kernel shown in (e) may be compared with the 2-D example shown in Fig. 5(c), whereas the ensemble source strength kernel shown in (h) is similar to the 2-D source strength kernel shown in Fig. 6(a). The small insets with cross-sections illustrate that noise cross-correlation kernels are similar to Rayleigh wave kernels, that is, they are mainly confined to the shallow upper mantle.

8.1.1 Uniform noise

Fig. 11 summarizes our results for noise uniformly distributed across the surface. Based upon the spectrum shown in Fig. 10, the source time function we use for calculating the ensemble forward wavefield is illustrated in Fig. 11(a). Following the same procedures as described in the 2-D examples, we obtain the simulated ensemble cross correlation by evaluating the ensemble forward wavefield at the receivers, as shown in Fig. 11(b). Next, we select an arrival in the branch of positive cross-correlation times, as shown in Fig. 11(c). Upon combining the two kernel contributions related to that one particular measurement (i.e. $\langle K^{\alpha\beta} \rangle$ and $\langle K^{\beta\alpha} \rangle$), we obtain the various kernels displayed in Figs 11(e)–(k). We see again that ensemble sensitivity kernels contain not only a classic banana-doughnut feature, but also weak hyperbolic fan tails. Sensitivity to the shear modulus is more pronounced than to the bulk modulus, reflecting the Rayleigh wave-dominated cross correlation. One may parametrize an isotropic earth model in terms of density, bulk modulus and shear modulus, involving kernels (27), (41) and (42) displayed in Figs 11(e), (f) and (g), or density, shear wave speed and compressional wave speed, involving kernels (45)–(47) displayed in Figs 11(i)–(k). We see that the Rayleigh wave has very little sensitivity to the bulk modulus, and thus compressional wave sensitivity is mainly through the shear modulus. We also note that the impedance kernel (45), shown in Fig. 11(i), is relatively weak, indicating that Rayleigh waves have limited sensitivity to density.

8.1.2 Non-uniform noise

3-D ensemble sensitivity kernels for an example of non-uniform noise are summarized in Fig. 12. Noise is restricted to the geographical area shown in red in Fig. 12(d), and hence the largest arrival in the ensemble cross correlation appears in the negative cross-correlation branch, as shown in Fig. 12(b). The reason for the opposite signs in Figs 11 and 12 is similar to the difference we see and discuss in Fig. 5.

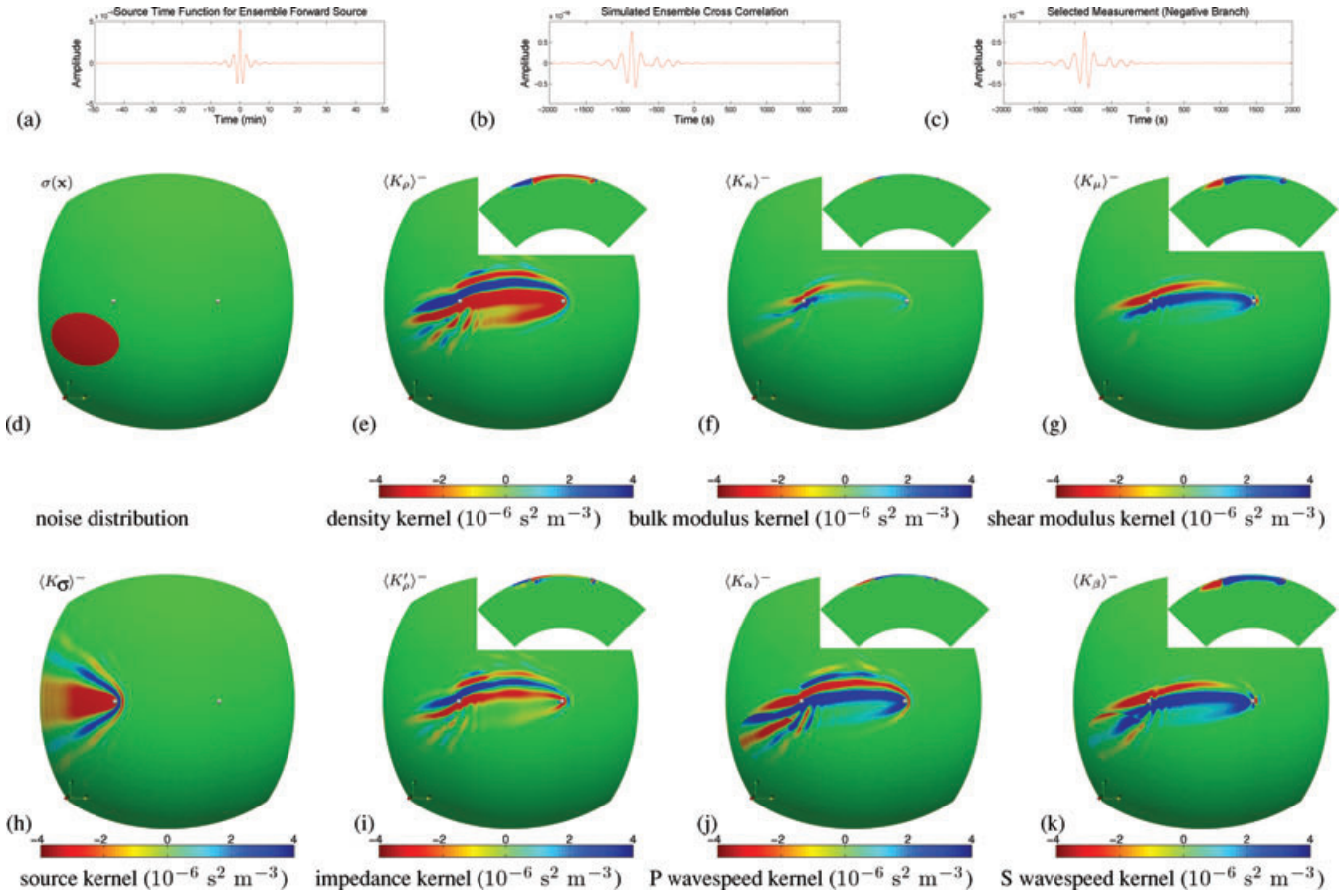

Figure 12. (a) Source time function used for the generation of the ensemble forward wavefield, adapted from Fig. 10(c). (b) Simulated ensemble cross correlation. (c) In this non-uniform noise example, waves mainly travel from the receiver on the left to the receiver on the right, restricting measurements to the negative cross-correlation branch. (d) Map view of the surface noise distribution (in red) and receiver geometry (two white dots). (e)–(k) Ensemble sensitivity kernels. Kernels (e), (f) and (g) correspond to using the model parameters density ρ, bulk modulus κ and shear modulus μ, whereas kernels (i), (j) and (k) correspond to using density ρ, compressional wave speed α and shear wave speed β. The ensemble density kernel shown in (e) may be compared with the 2-D example shown in Fig. 9(a), whereas the ensemble source strength kernel shown in (h) is similar to the 2-D source strength kernel shown in Fig. 9(b). Both map views and cross-sections indicate that ensemble cross-correlation measurements are not solely dependent on structure between the two receivers if the noise distribution is asymmetric.

8.2 Global simulations

Classic ambient noise tomography assumes that noise is uniformly distributed across the Earth's surface. In this section we explore ensemble sensitivity kernels under such ideal circumstances. We also investigate the nature of global ensemble kernels when noise is non-uniform.

Although we focus on cross-correlation sensitivity kernels, we begin with an illustration of simulated ensemble cross correlations as a function of epicentral distance. Using the data-processing workflow suggested by Bensen *et al.* (2007), Cupillard & Capdeville (2010) calculate simulated ensemble cross correlations by stacking more than 5000 synthetic cross correlations generated by random noise. Their results provide an opportunity to directly assess our calculation of ensemble cross correlations, discussed in Section 4, which does not involve stacking.

8.2.1 Comparison of simulated ensemble cross correlations

Following Cupillard & Capdeville (2010), we deploy a linear array of 12 receivers along the equator, with distances varying from 10° to 60°. To compare our simulated ensemble cross correlations with those of Cupillard & Capdeville (2010), we temporarily modify our spectrum of the ensemble-averaged noise to the whitened spectrum in their experiments. As described in the 2-D examples, ensemble cross correlations may be obtained based upon two simulations: one for the generating wavefield and the other for the ensemble forward wavefield. Using the relationships (38) and (39), evaluating the ensemble forward wavefield at one receiver provides the cross correlation between that receiver and the 'master' receiver (here, the first receiver). Since we may access all receivers at the same time, all cross correlations between the 'master' and other receivers may be obtained simultaneously, that is, based upon two simulations, regardless of the total number of receivers. The ensemble cross correlations due to the whitened spectrum are shown in Fig. 13, together with the cross correlations obtained by Cupillard & Capdeville (2010).

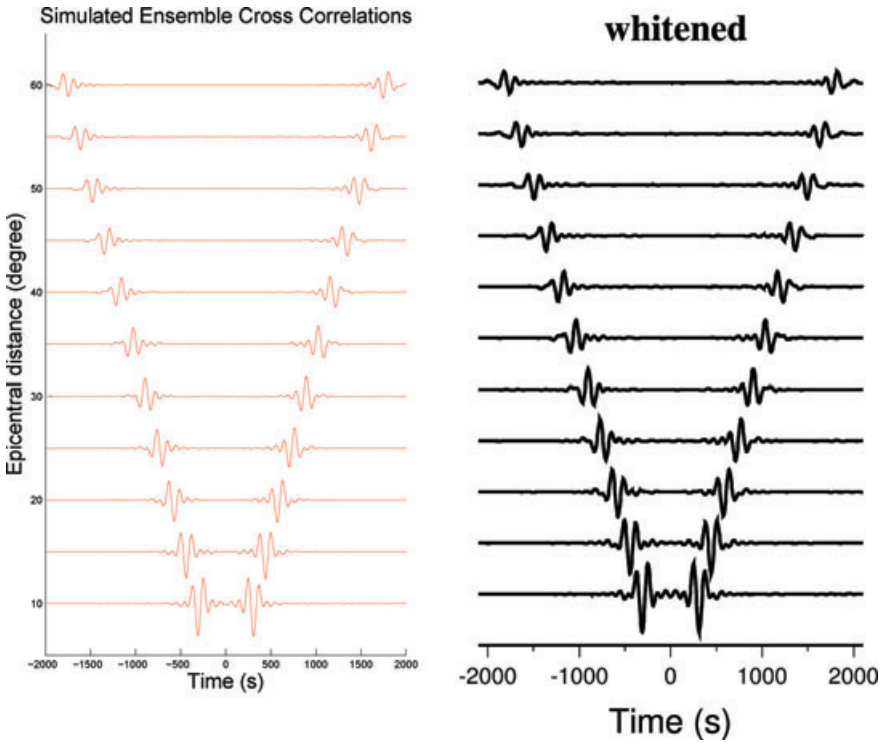

Figure 13. Record sections of cross correlations between the first and 11 other receivers in a linear array along the equator, with distances varying from 10° to 60°. The noise is assumed to be uniform, and the earth model is PREM. (a) Ensemble cross correlations obtained based on two numerical simulations of 7 hr of global seismic wave propagation, using the relationships (38) and (39). (b) Ensemble cross correlations obtained by stacking more than 5000 synthetic cross correlations from random noise. The two results are in good qualitative agreement, despite the two very different calculations of the ensemble cross correlation. Reproduced with permission from figure 2 of Cupillard & Capdeville (2010).

8.2.2 Sensitivity to shear wave speed

Ensemble cross correlations are dominated by Rayleigh surface waves, whose main sensitivity is to shear wave speed. In this section, we focus on such kernels in global earth models. In this case, ensemble cross correlations involve multi-orbit surface waves generated by noise. Therefore, we investigate minor and major arc sensitivity for uniform and non-uniform noise.

Fig. 14 illustrates the results for a global uniform noise distribution, that is, the ideal circumstance generally assumed in ambient noise tomography. Sensitivity along the minor arc is basically the same as in the regional case. Along the major arc, the two contributions $\langle K_\beta^{\alpha\beta}\rangle^+$ and $\langle K_\beta^{\beta\alpha}\rangle^+$ cancel, resulting in a vanishing ensemble shear wave speed sensitivity kernel $\langle K_\beta\rangle^+$ in the antipodal region. Thus, as long as the noise is uniform, the frequently made assumption that cross-correlation measurements are only sensitive to structure along the minor arc is valid.

Fig. 15 illustrates the results for non-uniform noise. The noise distribution is the same as in the regional case shown in Fig. 12(d). As in that case, we see that the 'jets' in the two kernel contributions, Figs 15(a) and (b), do not cancel each other, but, unlike the regional case, in the final kernel, Fig. 15(c), we see jets protruding from both receivers, reflecting the multi-orbit nature of the data: the noise generates waves travelling along both minor and major arcs. However, perhaps more disturbingly, there is significant sensitivity in the antipodal region, because in the non-uniform noise case the two kernel contributions do not cancel each other. Consequently, at periods long enough to generate surface waves that circumnavigate the globe, noise cross-correlation data due to non-uniform noise are sensitive to structure along the entire great circle. For short-period surface wave used in regional studies, (e.g. 3–30 s), this should not be cause for concern, since such waves are rapidly attenuated.

9 NOISE CROSS-CORRELATION TOMOGRAPHY

So far we have only considered 'one-to-one' cross-correlation measurements between two distinct locations. In this section, we consider the calculation of 'one-to-many' ensemble sensitivity kernels for noise cross-correlation tomography. This distinction is similar to the distinction between 'banana-doughnut' kernels and 'event' kernels in earthquake adjoint tomography (Tromp *et al.* 2005; Tape *et al.* 2007). Basically, the event kernel consists of a sum of weighted banana-doughnut kernels, and may be calculated based upon a single adjoint simulation. Our goal is to find a similar recipe for terrestrial and solar noise cross-correlation tomography.

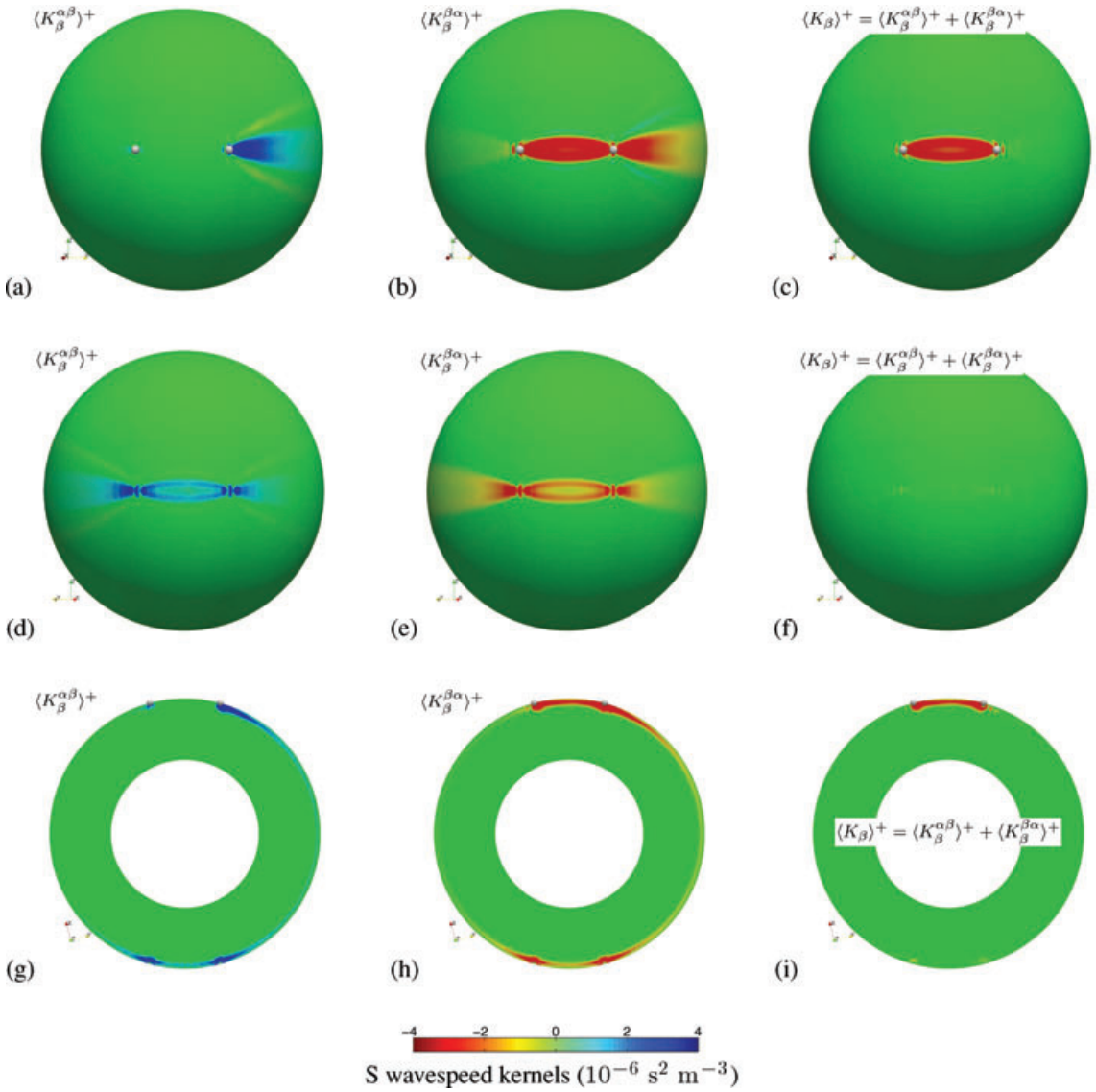

Figure 14. Ensemble shear wave speed kernels for uniform global noise. (a) The $\langle K_\beta^{\alpha\beta}\rangle^+$ contribution to the shear wave speed sensitivity kernel for a positive cross-correlation branch measurement. The two white dots denote the two receivers. (b) The $\langle K_\beta^{\beta\alpha}\rangle^+$ contribution to the shear wave speed sensitivity kernel. (c) The complete shear wave speed ensemble sensitivity kernel $\langle K_\beta\rangle^+ = \langle K_\beta^{\alpha\beta}\rangle^+ + \langle K_\beta^{\beta\alpha}\rangle^+$. The 'jets' in the two contributions have opposite signs and cancel each other, leaving a banana-donut kernel between the two receivers. (d)–(f) Same as (a)–(c) but viewed from the antipode, that is, looking at the major arc. The two contributions almost cancel, resulting in no sensitivity along the major arc. (g)–(i) Same as (a)–(c) but viewed in cross-section. Cancellation along the major arc is clearly visible.

9.1 'One-to-many' ensemble sensitivity kernels: terrestrial arrays

To move towards practical noise cross-correlation tomography, we need to slightly adjust and expand the notation. Let us redefine the 'one-to-one' misfit function (19) by adding two superscripts.

$$\chi^{\alpha\beta} = \frac{1}{2}\int_{-\infty}^{\infty}\left[\langle C^{\alpha\beta}\rangle_{\rm sim} - \langle C^{\alpha\beta}\rangle_{\rm obs}\right]^2 \,{\rm d}t, \tag{83}$$

where we have made the time integration limits explicit. Because of the symmetry (17), we have the identity

$$\chi^{\alpha\beta} = \chi^{\beta\alpha}. \tag{84}$$

Now let us consider the case of terrestrial noise cross-correlation tomography. Imagine we have an array with N seismographic stations from which we are cross correlating time-series. The overall misfit function is defined by

$$\chi = \frac{1}{2}\sum_{\alpha=1}^{N}\sum_{\beta\neq\alpha}\chi^{\alpha\beta}. \tag{85}$$

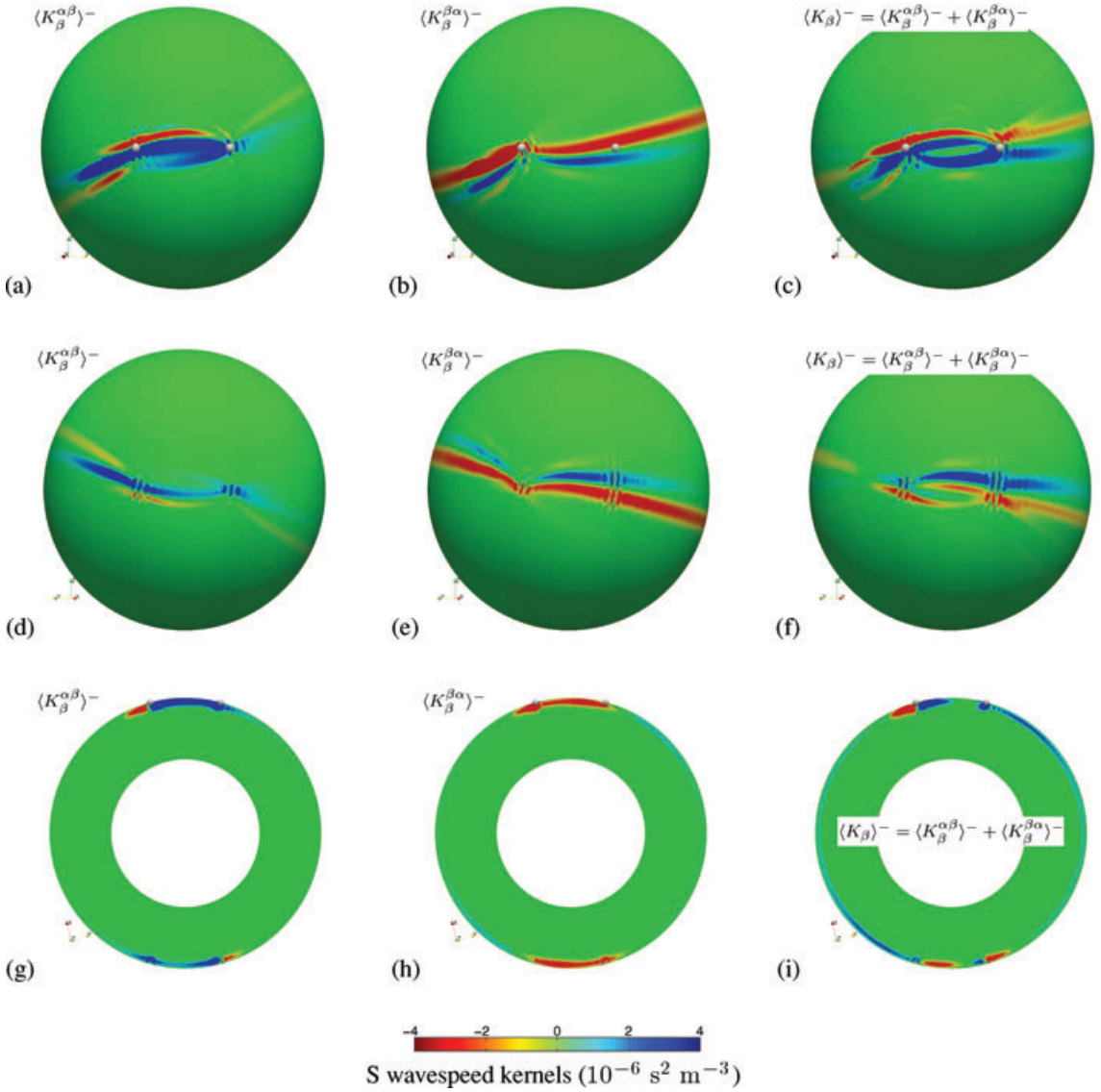

Figure 15. Ensemble shear wave speed kernels for non-uniform global noise. (a) The $\langle K_\beta^{\alpha\beta}\rangle^-$ contribution to the shear wave speed sensitivity kernel for a negative cross-correlation branch measurement. The two white dots denote the two receivers. (b) The $\langle K_\beta^{\beta\alpha}\rangle^-$ contribution to the shear wave speed sensitivity kernel. (c) The complete shear wave speed ensemble sensitivity kernel $\langle K_\beta\rangle^- = \langle K_\beta^{\alpha\beta}\rangle^- + \langle K_\beta^{\beta\alpha}\rangle^-$. The 'jets' in the two contributions do not cancel each other, and, unlike in the regional kernel shown in Fig. 12(k), jets protrude from each receiver. (d)–(f) Same as (a)–(c) but viewed from the antipode, that is, looking at the major arc. The two contributions do not cancel, resulting in significant sensitivity along the major arc. (g)–(i) Same as (a)–(c) but viewed in cross-section.

where $\chi^{\alpha\beta}$ is defined by (83). Taking the variation of this misfit leads to ensemble kernels of the form

$$\langle K\rangle = \frac{1}{2}\sum_{\alpha=1}^{N}\sum_{\beta\neq\alpha}(\langle K^{\alpha\beta}\rangle + \langle K^{\beta\alpha}\rangle) = \sum_{\alpha=1}^{N}\sum_{\beta\neq\alpha}\langle K^{\alpha\beta}\rangle, \tag{86}$$

where we have used the symmetry (84), and the fact that the sums over α and β are interchangeable. Now we introduce, for example, the 'ensemble density event kernel'

$$\langle K_\rho^\alpha\rangle \equiv \sum_{\beta\neq\alpha}\langle K_\rho^{\alpha\beta}\rangle = -\sum_{\beta\neq\alpha}\int \rho\,\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)\cdot\partial_t^2\mathbf{\Phi}^\alpha(t)\,\mathrm{d}t = -\int \rho\,\mathbf{\Phi}^{\dagger\,\alpha}(-t)\cdot\partial_t^2\mathbf{\Phi}^\alpha(t)\,\mathrm{d}t, \tag{87}$$

where we have defined

$$\mathbf{\Phi}^{\dagger\,\alpha} \equiv \sum_{\beta\neq\alpha}\mathbf{\Phi}^{\dagger\,\alpha\beta}. \tag{88}$$

This ensemble adjoint wavefield $\mathbf{\Phi}^{\dagger\,\alpha}$ is driven by the ensemble adjoint source

$$\mathbf{F}^{\dagger\,\alpha}(\mathbf{x},t) = \sum_{\beta\neq\alpha}\hat{\mathbf{v}}^\beta\,\Delta\langle C^{\alpha\beta}\rangle(t)\,\delta(\mathbf{x}-\mathbf{x}^\beta). \tag{89}$$

Basically, for a given location $\mathbf{x}^\alpha$, one simultaneously backprojects the differences between observed and simulated ensemble cross correlations from all other locations $\mathbf{x}^\beta$, $\beta \neq \alpha$. This defines the ensemble adjoint wavefield $\mathbf{\Phi}^{\dagger\,\alpha}$ associated with location $\mathbf{x}^\alpha$.

We conclude that the variation in the misfit function for noise cross-correlation tomography may be expressed as

$$\delta\chi = \sum_{\alpha=1}^{N} \int_\Omega \left(\delta \ln\rho \, \left\langle K_\rho^\alpha \right\rangle + \delta\mathbf{c} :: \langle \mathbf{K}_\mathbf{c}^\alpha \rangle + \langle \delta\Sigma^\alpha \rangle\right) \, \mathrm{d}^3\mathbf{x}, \tag{90}$$

that is, contributions from all points $\mathbf{x}^\alpha$, where we have introduced the ensemble event kernels

$$\left\langle K_\rho^\alpha \right\rangle = -\int \rho \, \mathbf{\Phi}^{\dagger\,\alpha}(-t) \cdot \partial_t^2 \mathbf{\Phi}^\alpha(t) \, \mathrm{d}t, \tag{91}$$

$$\left\langle \mathbf{K}_\mathbf{c}^\alpha \right\rangle = -\int \boldsymbol{\nabla}\mathbf{\Phi}^{\dagger\,\alpha}(-t) \, \boldsymbol{\nabla}\mathbf{\Phi}^\alpha(t) \, \mathrm{d}t \tag{92}$$

and

$$\langle \delta\Sigma^\alpha \rangle = \int \mathbf{\Phi}^{\dagger\,\alpha}(-t) \cdot \delta\mathbf{F}^\alpha(t) \, \mathrm{d}t. \tag{93}$$

The isotropic kernels are given by

$$\left\langle K_\mu^\alpha \right\rangle = -\int 2\mu \, \mathbf{D}^{\dagger\,\alpha}(-t) : \mathbf{D}^\alpha(t) \, \mathrm{d}t, \tag{94}$$

$$\langle K_\kappa^\alpha \rangle = -\int \kappa \, \boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\dagger\,\alpha}(-t) \, \boldsymbol{\nabla}\cdot\mathbf{\Phi}^\alpha(t) \, \mathrm{d}t, \tag{95}$$

where

$$\mathbf{D}^\alpha = \frac{1}{2}[\boldsymbol{\nabla}\mathbf{\Phi}^\alpha + (\boldsymbol{\nabla}\mathbf{\Phi}^\alpha)^T] - \frac{1}{3}(\boldsymbol{\nabla}\cdot\mathbf{\Phi}^\alpha)\,\mathbf{I}, \qquad \mathbf{D}^{\dagger\,\alpha} = \frac{1}{2}[\boldsymbol{\nabla}\mathbf{\Phi}^{\dagger\,\alpha} + (\boldsymbol{\nabla}\mathbf{\Phi}^{\dagger\,\alpha})^T] - \frac{1}{3}(\boldsymbol{\nabla}\cdot\mathbf{\Phi}^{\dagger\,\alpha})\,\mathbf{I}. \tag{96}$$

The results in this section are consistent with our earlier expressions for one-to-one kernels, because in that formulation there is one contribution for the first point $\mathbf{x}^\alpha$, and a second contribution for the second point $\mathbf{x}^\beta$, just as there is a sum of $N = 2$ contributions in (90). One is free to choose any measure of the difference between observed and simulated ensemble cross correlations for both positive and negative cross-correlation branches, for example, waveform differences, delay times or multitaper phase and amplitude measurements. If one makes measurements in both the positive and negative branches, the two measurements are simply combined in the appropriate form of the adjoint source (89). The entire procedure involves a number of forward simulations that scales linearly with the number of receivers.

9.2 'One-to-many' ensemble sensitivity kernels: helioseismology

Now let us consider the case of solar noise cross-correlation tomography. Here we have the luxury of cross correlating any pixel on the Sun's surface with any other pixel. One could imagine selecting M 'master' pixels whose time-series are cross correlated with N other pixels, where $N \gg M$. In this case the overall misfit function is defined by

$$\chi = \frac{1}{2} \sum_{\alpha=1}^{M} \sum_{\beta=1}^{N} \chi^{\alpha\beta}, \tag{97}$$

where the index α labels master pixels whereas the index β labels other pixels, and where $\chi^{\alpha\beta}$ is defined by (83). Taking the variation of this misfit leads to ensemble kernels of the form

$$\langle K \rangle = \frac{1}{2} \sum_{\alpha=1}^{M} \sum_{\beta=1}^{N} (\langle K^{\alpha\beta} \rangle + \langle K^{\beta\alpha} \rangle). \tag{98}$$

Now, working with a sum over β of eq. (C5), we introduce, for example, the 'ensemble density event kernel'

$$\begin{aligned} \left\langle K_\rho^\alpha \right\rangle &\equiv \frac{1}{2} \sum_{\beta=1}^{N} \left[\left\langle K_\rho^{\alpha\beta} \right\rangle + \left\langle K_\rho^{\beta\alpha} \right\rangle\right] \\ &= -\frac{1}{2} \int \rho \, \left[\mathbf{\Phi}^{\dagger\,\alpha}(-t) \cdot \partial_t^2 \mathbf{\Phi}^\alpha(t) + \overline{\mathbf{\Phi}}^{\dagger\,\alpha}(-t) \cdot \partial_t^2 \overline{\mathbf{\Phi}}^\alpha(t)\right] \, \mathrm{d}t. \end{aligned} \tag{99}$$

As usual, the M ensemble forward wavefields $\mathbf{\Phi}^\alpha(\mathbf{x}, \omega)$ are defined by the M ensemble forward sources (30). Similar to the approach in the previous section, the M wavefields $\mathbf{\Phi}^{\dagger\,\alpha}(\mathbf{x}, \omega)$ are defined by

$$\mathbf{\Phi}^{\dagger\,\alpha} \equiv \sum_{\beta=1}^{N} \mathbf{\Phi}^{\dagger\,\alpha\beta}, \tag{100}$$

and are generated by the M ensemble adjoint sources

$$\mathbf{F}^{\dagger\,\alpha}(\mathbf{x}, \omega) = \sum_{\beta=1}^{N} \hat{\mathbf{v}}^\beta \, \Delta\langle C^{\alpha\beta} \rangle(\omega) \, \delta(\mathbf{x} - \mathbf{x}^\beta), \tag{101}$$

that is, simultaneous backprojection from all pixels $\beta = 1, \ldots, N$. The M wavefields $\overline{\mathbf{\Phi}}^{\alpha}(\mathbf{x}, \omega)$ are driven by

$$\overline{\mathbf{F}}^{\alpha}(\mathbf{x}, \omega) = \boldsymbol{\nu}^{\alpha}\, \delta(\mathbf{x} - \mathbf{x}^{\alpha}). \tag{102}$$

Finally, the M wavefields $\overline{\mathbf{\Phi}}^{\dagger\,\alpha}(\mathbf{x}, \omega)$ are driven by

$$\overline{F}_i^{\dagger\,\alpha}(\mathbf{x}, \omega) = S_{ij}(\mathbf{x}, \omega) \sum_{\beta=1}^{N} \left[G_{jk}(\mathbf{x}, \mathbf{x}^{\beta}; \omega)\, \hat{\nu}_k^{\beta}\, \Delta\langle C^{\alpha\beta}\rangle(\omega) \right]^{*}. \tag{103}$$

It is now straightforward to show that the ensemble elastic tensor kernel may be written in the form

$$\langle \mathbf{K}_{\mathbf{c}}^{\alpha} \rangle = -\frac{1}{2} \int \left[\nabla \mathbf{\Phi}^{\dagger\,\alpha}(-t)\, \nabla \mathbf{\Phi}^{\alpha}(t) + \nabla \overline{\mathbf{\Phi}}^{\dagger\,\alpha}(-t)\, \nabla \overline{\mathbf{\Phi}}^{\alpha}(t) \right] \mathrm{d}t, \tag{104}$$

with similar expressions for isotropic model parameters. Finally, we have

$$\langle \delta \mathbf{\Sigma}^{\alpha} \rangle = \frac{1}{2} \int \left[\mathbf{\Phi}^{\dagger\,\alpha}(-t) \cdot \delta \mathbf{F}^{\alpha}(t) + \overline{\mathbf{\Phi}}^{\alpha}(t) \cdot \delta \overline{\mathbf{F}}^{\dagger\,\alpha}(-t) \right] \mathrm{d}t, \tag{105}$$

where

$$\delta \overline{F}_i^{\dagger\,\alpha}(\mathbf{x}, \omega) = \delta S_{ij}(\mathbf{x}, \omega) \sum_{\beta=1}^{N} \left[G_{jk}(\mathbf{x}, \mathbf{x}^{\beta}; \omega)\, \hat{\nu}_k^{\beta}\, \Delta\langle C^{\alpha\beta}\rangle(\omega) \right]^{*}. \tag{106}$$

Making ensemble cross-correlation measurements between all pixels $\alpha = 1, \ldots, M$ and $\beta = 1, \ldots, N$ requires $O(M)$ forward simulations, provided one takes advantage of relationship (39), which requires only $\mathbf{\Phi}^{\alpha}$. The ensemble kernel calculations require $O(M)$ forward simulations. Of course when $M = N$, the procedure outlined in this section is the same as discussed in the previous section.

9.3 Recipe

The following steps are involved in the construction of ensemble event kernels (91)–(95) for terrestrial noise cross-correlation tomography.

(i) Characterize the ensemble-averaged noise, namely $S_{ij}(\mathbf{x}, \omega)$ in (14).

(ii) For each point of interest $\mathbf{x}^{\alpha}$, calculate the generating wavefields $\boldsymbol{\eta}^{\alpha}(\mathbf{x}, t)$ based upon eq. (31), and store them as a 'movie' 'at locations where the noise is non-zero'.

(iii) Generate the source $\mathbf{F}^{\alpha}$ given by (32), which is needed for the calculation of the ensemble wavefield $\mathbf{\Phi}^{\alpha}$.

(iv) For all pairs $\mathbf{x}^{\alpha}$ and $\mathbf{x}^{\beta}$, calculate the theoretical ensemble cross correlation $\langle C^{\alpha\beta}\rangle(t)$ based upon expression (38).

(v) Determine a measure of the difference between observed and simulated ensemble cross correlations. This could be the waveform difference between simulated and measured positive/negative ensemble cross-correlation branches, traveltime delay measurements between simulated and measured ensemble cross correlations in both the positive and negative branches, etc. This forms the data set.

(vi) Construct the adjoint source $\mathbf{F}^{\dagger\,\alpha}$ given by (89), which is needed for the calculation of the ensemble adjoint wavefield $\mathbf{\Phi}^{\dagger\,\alpha}$.

(vii) Generate the ensemble event kernels (91)–(95) based upon a combined calculation of the wavefields $\mathbf{\Phi}^{\alpha}(\mathbf{x}, t)$ and $\mathbf{\Phi}^{\dagger\,\alpha}(\mathbf{x}, -t)$ (see Liu & Tromp 2006).

(viii) Add the ensemble event kernels to obtain the gradient of the misfit function.

Noise cross-correlation tomography on the Sun involves the modifications discussed in Section 9.2.

10 CONCLUSIONS

Under the assumption that noise is spatially uncorrelated, we efficiently calculate synthetic ensemble cross correlations which may be used as a reference in seismic interferometry. Our cross correlations compare favourably with results of Cupillard & Capdeville (2010) obtained following a standard stacking process.

Based upon a measure of misfit between simulated and observed noise cross correlations, we determine corresponding sensitivity kernels. Such kernels involve interactions between ensemble forward and adjoint wavefields, which may be readily calculated numerically in 3-D earth models, even for non-uniform noise. These 'ensemble kernels' form the basis of noise cross-correlation tomography, much like finite-frequency sensitivity kernels for earthquake data may be used in an iterative inversion procedure we refer to as 'adjoint tomography' (Tape *et al.* 2009, 2010).

In terrestrial seismology, the ensemble forward wavefield associated with a receiver of interest interacts with an ensemble adjoint wavefield generated by simultaneously backprojecting time-reversed differences between observed and simulated cross correlations from all the other receivers. Thus, the calculation of such kernels requires only three numerical simulations per receiver. In helioseismology, a limited number of 'master pixels' may be cross correlated with any number of other pixels, again simultaneously backprojecting data from all the other pixels, thereby providing tremendous coverage of the Sun. In this case, the number of simulations scales linearly with the number of master pixels.

Ensemble noise cross-correlation measurements typically involve stacks of individual seismograms with record lengths on the order of a day. It takes Rayleigh waves roughly 3 hr to circumnavigate the globe, and thus these measurements involve multiple orbits. The corresponding

sensitivity kernels reflect the nature of this measurement, exhibiting sensitivity to structure along the entire great circle that connects the two receivers of interest, much like kernels for multi-orbit surface waves (e.g. Zhou 2009). If the noise is uniform, sensitivity along the major arc practically vanishes, but for non-uniform noise there can be significant sensitivity along the entire great circle. This fact has important implications for regional noise cross-correlation studies based upon long-period data ($T > \sim$50 s), because tomographic inversions of regional noise data are affected by structure outside of the region of interest. For short-period noise cross-correlation data ($\sim$3–30 s) major arc sensitivity is less of a concern, because the related surface waves are scattered and attenuated relatively quickly.

ACKNOWLEDGMENTS

We thank Andreas Fichtner and Fan-Chi Liu for constructive reviews which helped to improve the manuscript. All 3-D simulations were performed on the 'Sesame' Dell cluster built and maintained by the Princeton Institute for Computational Science & Engineering. The open source spectral-element software packages SPECFEM2D and SPECFEM3D_GLOBE used for this paper are available for download via the Computational Infrastructure for Geodynamics (www.geodynamics.org). We acknowledge support by the National Science Foundation under grant EAR-0711177.

REFERENCES

Aki, K. & Richards, P.G., 1980, *Quantitative Seismology, Theory and Methods*, W.H. Freeman, San Francisco, California.

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.,* **169,** 1239–1260.

Bensen, G.D., Ritzwoller, M.H. & Yang, Y., 2009, A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise, *Geophys. J. Int.,* **177,** 1177–1196.

Brenguier, F., Shapiro, N.M., Campillo, M., Nercessian, A. & Ferrazzini, V., 2007, 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations, *Geophys. Res. Lett.,* **34,** 2305–2308.

Campillo, M. & Paul, A., 2003, Long-range correlations in the diffuse seismic coda, *Science,* **299,** 547–549.

Christensen-Dalsgaard, J., 2002, Helioseismology, *Rev. Mod. Phys.,* **74,** 1073–1129.

Cupillard, P. & Capdeville, Y., 2010, On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach, *Geophys. J. Int.,* **181,** 1687–1700.

Dahlen, F.A. & Tromp, J., 1998, *Theoretical Global Seismology*, Princeton University Press, New Jersey.

Duvall, Jr., T.L., Jefferies, S.M., Harvey, J.W. & Pomerantz, M.A., 1993, Time-distance helioseismology, *Nature,* **362,** 430–432.

Dziewonski, A. & Anderson, D., 1981, Preliminary reference Earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

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

Fan, Y. & Snieder, R., 2009, Required source distribution for interferometry of waves and diffusive fields, *Geophys. J. Int.,* **179,** 1232–1244.

Fichtner, A. & Igel, H., 2009, Sensitivity densities for rotational ground-motion measurements, *Bull. seism. Soc. Am.,* **99,** 1302–1314.

Fichtner, A., Kennett, B., Igel, H. & Bunge, H.-P., 2009, Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.,* **179,** 1703–1725.

Gee, L. & Jordan, T.H., 1992, Generalized seismological data functionals, *Geophys. J. Int.,* **111,** 363–390.

Gizon, L. & Birch, A.C., 2002, Time-distance helioseismology: the forward problem for random distributed sources, *Astrophys. J.,* **571,** 966–986.

Gizon, L. & Birch, A.C., 2005, Local helioseismology, *Living Rev. Sol. Phys.*, **2,** 6.

Gizon, L., Birch, A.C. & Spruit, H.C., 2010, Local helioseismology: three dimensional imaging of the solar interior, *Ann. Rev. Astron. Astrophys.,* in press (arXiv:1001.0930).

Hanasoge, S.M., Couvidat, S., Rajaguru, S.P. & Birch, A.C., 2008, Impact of locally suppressed wave sources on helioseismic traveltimes, *Mon. Not. R. astr. Soc.,* **391,** 1931–1939.

Kimman, W. & Trampert, J., 2010, Approximations in seismic interferometry and their effects on surface waves, *Geophys. J. Int.,* **182**, 461–476.

Komatitsch, D. & Tromp, J., 1999, Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.,* **139,** 806–822.

Komatitsch, D. & Tromp, J., 2002a, Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.,* **149,** 390–412.

Komatitsch, D. & Tromp, J., 2002b, Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.,* **150,** 308–318.

Komatitsch, D. & Vilotte, J.-P., 1998, The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.,* **88,** 368–392.

Komatitsch, D., Ritsema, J. & Tromp, J., 2002, The spectral-element method, Beowulf computing, and global seismology, *Science,* **298,** 1737–1742.

Komatitsch, D., Liu, Q., Tromp, J., Süss, P., Stidham, C. & Shaw, J.H., 2004, Simulations of ground motion in the Los Angeles basin based upon the spectral-element method, *Bull. seism. Soc. Am.,* **94,** 187–206.

Larose, E., Khan, A., Nakamura, Y. & Campillo, M., 2005, Lunar subsurface investigated from correlation of seismic noise, *Geophys. Res. Lett.*, **32,** 16 201–16 204.

Laske, G. & Masters, G., 1996, Constraints on global phase velocity maps from long-period polarization data, *J. geophys. Res.,* **101,** 16 059–16 075.

Liu, Q. & Tromp, J., 2006, Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.,* **96,** 2383–2397.

Liu, Q. & Tromp, J., 2008, Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.,* **174,** 265–286.

Luo, Y. & Schuster, G., 1990, Parsimonious staggered grid finite-differencing of the wave equation, *Geophys. Res. Lett.,* **17,** 155–158.

Peterson, J., 1993, Observations and modelling of background seismic noise, Tech. Rep., U.S. Geol. Surv.

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005, High-resolution surface-wave tomography from ambient seismic noise, *Science,* **307**(5715), 1615–1618.

Tanimoto, T., 2008, Normal-mode solution for the seismic noise cross-correlation method, *Geophys. J. Int.,* **175,** 1169–1175.

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

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009, Adjoint tomography of the Southern California crust, *Science,* **325,** 988–992.

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

Wapenaar, K., Slob, E. & Snieder, R., 2006, Unified Green's function retrieval by cross correlation, *Phys. Rev. Lett.,* **97**(23), 23 4301–23 4304.

Woodard, M.F., 1997, Implications of localized, acoustic absorption for heliotomographic analysis of sunspots, *Astrophys. J.,* **485,** 890–894.

Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007, Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.,* **168,** 259–274.

Zhou, Y., 2009, Multimode surface wave sensitivity kernels in radially anisotropic earth media, *Geophys. J. Int.,***176,** 865–888.

APPENDIX A: VARIATION OF THE ENSEMBLE CROSS CORRELATION

In this appendix, we determine the variation of the ensemble cross correlation $\langle C^{\alpha\beta}\rangle$. Upon taking the variation of the first term on the right-hand side of (19) we have

$$\int \Delta\langle C^{\alpha\beta}\rangle\,\delta\langle C^{\alpha\beta}\rangle\,\mathrm{d}t = \int \Delta\langle C^{\alpha\beta}\rangle\,\langle\delta C^{\alpha\beta}\rangle\,\mathrm{d}t = \left\langle \int \Delta\langle C^{\alpha\beta}\rangle\,\delta C^{\alpha\beta}\,\mathrm{d}t \right\rangle, \tag{A1}$$

where $\Delta\langle C^{\alpha\beta}\rangle = \langle C^{\alpha\beta}\rangle_{\rm sim} - \langle C^{\alpha\beta}\rangle_{\rm obs}$ denotes the ensemble-averaged cross-correlation difference. Upon taking the variation of (9) we have

$$\delta C^{\alpha\beta}(t) = \int [s^\alpha(t+\tau)\,\delta s^\beta(\tau) + \delta s^\alpha(t+\tau)\,s^\beta(\tau)]\,\mathrm{d}\tau. \tag{A2}$$

Thus, using (A1),

$$\int \Delta\langle C^{\alpha\beta}\rangle\,\delta\langle C^{\alpha\beta}\rangle\,\mathrm{d}t = \left\langle \int \Delta\langle C^{\alpha\beta}\rangle \int [s^\alpha(t+\tau)\,\delta s^\beta(\tau) + \delta s^\alpha(t+\tau)\,s^\beta(\tau)]\,\mathrm{d}\tau\,\mathrm{d}t \right\rangle. \tag{A3}$$

We need to extract the variation $\delta\mathbf{s}(\mathbf{x},t)$ from this expression to identify what will become the adjoint source in Appendix B. To this end, the first term on the right-hand side in (A3) may be expressed as

$$\iiint \Delta\langle C^{\alpha\beta}\rangle(\tau)\,s^\alpha(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\beta)\,\hat{\boldsymbol{\nu}}^\beta\cdot\delta\mathbf{s}(\mathbf{x},t)\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t, \tag{A4}$$

from which we recognize the contribution $\hat{\boldsymbol{\nu}}^\beta \int \Delta\langle C^{\alpha\beta}\rangle(\tau)\,s^\alpha(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\beta)$. The second term on the right-hand side in (A3) is

$$\iiint \Delta\langle C^{\alpha\beta}\rangle(t)\,s^\beta(\tau)\,\delta(\mathbf{x}-\mathbf{x}^\alpha)\,\hat{\boldsymbol{\nu}}^\alpha\cdot\delta\mathbf{s}(\mathbf{x},t+\tau)\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}\tau\,\mathrm{d}t. \tag{A5}$$

Define $t' = t+\tau$, that is, $\tau = t'-t$:

$$\begin{aligned}
&\iiint \Delta\langle C^{\alpha\beta}\rangle(t)\,s^\beta(t'-t)\,\mathrm{d}t\,\delta(\mathbf{x}-\mathbf{x}^\alpha)\,\hat{\boldsymbol{\nu}}^\alpha\cdot\delta\mathbf{s}(\mathbf{x},t')\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t' \\
&\quad= \iiint \Delta\langle C^{\alpha\beta}\rangle(\tau)\,s^\beta(t-\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\alpha)\,\hat{\boldsymbol{\nu}}^\alpha\cdot\delta\mathbf{s}(\mathbf{x},t)\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \\
&\quad= \iiint \Delta\langle C^{\alpha\beta}\rangle(-\tau)\,s^\beta(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\alpha)\,\hat{\boldsymbol{\nu}}^\alpha\cdot\delta\mathbf{s}(\mathbf{x},t)\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \\
&\quad= \iiint \Delta\langle C^{\beta\alpha}\rangle(\tau)\,s^\beta(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\alpha)\,\hat{\boldsymbol{\nu}}^\alpha\cdot\delta\mathbf{s}(\mathbf{x},t)\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t,
\end{aligned} \tag{A6}$$

where we have used (17) in the last equality. Thus we obtain the desired expression

$$\begin{aligned}
\int \Delta\langle C^{\alpha\beta}\rangle\,\delta\langle C^{\alpha\beta}\rangle\,\mathrm{d}t = \Big\langle \iiint \Big[&\hat{\boldsymbol{\nu}}^\beta \int \Delta\langle C^{\alpha\beta}\rangle(\tau)\,s^\alpha(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\beta) \\
&+ \hat{\boldsymbol{\nu}}^\alpha \int \Delta\langle C^{\beta\alpha}\rangle(\tau)\,s^\beta(t+\tau)\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^\alpha)\Big]\cdot\delta\mathbf{s}(\mathbf{x},t)\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \Big\rangle.
\end{aligned} \tag{A7}$$

Consequently, the right-hand side—like the left-hand side—is symmetric in α and β, as expected.

APPENDIX B: ADJOINT EQUATIONS AND KERNELS

Upon taking the variation of the action (19), using Hooke's law (2), we obtain

$$\begin{aligned}
\delta\chi = &\int \Delta\langle C^{\alpha\beta}\rangle\,\delta\langle C^{\alpha\beta}\rangle\,\mathrm{d}t - \left\langle \iiint_\Omega \boldsymbol{\lambda}\cdot\left[\delta\rho\,\partial_t^2\mathbf{s} - \boldsymbol{\nabla}\cdot(\delta\mathbf{c}:\boldsymbol{\nabla}\mathbf{s}) - \delta\mathbf{f}\right]\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \right\rangle \\
&- \left\langle \iiint_\Omega \boldsymbol{\lambda}\cdot\left[\rho\,\partial_t^2\delta\mathbf{s} - \boldsymbol{\nabla}\cdot(\mathbf{c}:\boldsymbol{\nabla}\delta\mathbf{s})\right]\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \right\rangle,
\end{aligned} \tag{B1}$$

where $\Delta\langle C^{\alpha\beta}\rangle = \langle C^{\alpha\beta}\rangle_{\rm sim} - \langle C^{\alpha\beta}\rangle_{\rm obs}$. Perturbing the free surface boundary condition (3) implies $\hat{\mathbf{n}}\cdot(\delta\mathbf{c}:\boldsymbol{\nabla}\mathbf{s} + \mathbf{c}:\boldsymbol{\nabla}\delta\mathbf{s}) = \mathbf{0}$ on $\partial\Omega$. Integrating the terms involving spatial and temporal derivatives of $\mathbf{s}$ and the variation $\delta\mathbf{s}$ by parts, we find

$$\begin{aligned}
\delta\chi = &\int \Delta\langle C^{\alpha\beta}\rangle\,\delta\langle C^{\alpha\beta}\rangle\,\mathrm{d}t - \left\langle \iiint_\Omega \left(\delta\rho\,\boldsymbol{\lambda}\cdot\partial_t^2\mathbf{s} + \boldsymbol{\nabla}\boldsymbol{\lambda}:\delta\mathbf{c}:\boldsymbol{\nabla}\mathbf{s} - \boldsymbol{\lambda}\cdot\delta\mathbf{f}\right)\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \right\rangle \\
&- \left\langle \iiint_\Omega \left[\rho\,\partial_t^2\boldsymbol{\lambda} - \boldsymbol{\nabla}\cdot(\mathbf{c}:\boldsymbol{\nabla}\boldsymbol{\lambda})\right]\cdot\delta\mathbf{s}\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}t \right\rangle - \left\langle \iint_{\partial\Omega} \hat{\mathbf{n}}\cdot(\mathbf{c}:\boldsymbol{\nabla}\boldsymbol{\lambda})\cdot\delta\mathbf{s}\,\mathrm{d}^2\mathbf{x}\,\mathrm{d}t \right\rangle.
\end{aligned} \tag{B2}$$

We demonstrate in Appendix A that the variation in the ensemble cross correlation is determined by (A7). Thus, in the absence of perturbations in the model parameters $\delta\rho$, $\delta\mathbf{c}$ and $\delta\mathbf{f}$, the variation in the action (B2) is stationary with respect to perturbations $\delta\mathbf{s}$ provided the Lagrange

multiplier $\boldsymbol{\lambda}$ satisfies the equation

$$\rho\,\partial_t^2\boldsymbol{\lambda} - \boldsymbol{\nabla}\cdot(\mathbf{c}:\boldsymbol{\nabla}\boldsymbol{\lambda}) = \hat{\boldsymbol{\nu}}^{\beta}\int \Delta\langle C^{\alpha\beta}\rangle(\tau)\,s^{\alpha}(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^{\beta}) \;+\hat{\boldsymbol{\nu}}^{\alpha}\int \Delta\langle C^{\beta\alpha}\rangle(\tau)\,s^{\beta}(t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^{\alpha}), \tag{B3}$$

subject to the free surface boundary condition

$$\hat{\mathbf{n}}\cdot(\mathbf{c}:\boldsymbol{\nabla}\boldsymbol{\lambda}) = \mathbf{0} \quad \text{on } \partial\Omega. \tag{B4}$$

To appreciate the nature of the Lagrange multiplier wavefield, let us define the 'adjoint wavefield' $\mathbf{s}^{\dagger}$ in terms of the Lagrange multiplier wavefield $\boldsymbol{\lambda}$ by

$$\mathbf{s}^{\dagger}(\mathbf{x},t) \equiv \boldsymbol{\lambda}(\mathbf{x},-t), \tag{B5}$$

that is, the adjoint wavefield is the time-reversed Lagrange multiplier wavefield $\boldsymbol{\lambda}$. Then the adjoint wavefield $\mathbf{s}^{\dagger}$ is determined by the set of equations

$$\rho\,\partial_t^2\mathbf{s}^{\dagger} - \boldsymbol{\nabla}\cdot\mathbf{T}^{\dagger} = \mathbf{f}^{\dagger}, \tag{B6}$$

where $\mathbf{f}^{\dagger}$ denotes the 'adjoint source'

$$\mathbf{f}^{\dagger}(\mathbf{x},t) = \hat{\boldsymbol{\nu}}^{\beta}\int \Delta\langle C^{\alpha\beta}\rangle(\tau)\,s^{\alpha}(-t+\tau)\,\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^{\beta}) + \hat{\boldsymbol{\nu}}^{\alpha}\int \Delta\langle C^{\beta\alpha}\rangle(\tau)\,s^{\beta}(-t+\tau)\mathrm{d}\tau\,\delta(\mathbf{x}-\mathbf{x}^{\alpha}), \tag{B7}$$

and where we have defined the *adjoint stress* in terms of the gradient of the adjoint displacement by

$$\mathbf{T}^{\dagger} \equiv \mathbf{c}:\boldsymbol{\nabla}\mathbf{s}^{\dagger}. \tag{B8}$$

The adjoint wave eq. (B6) is subject to the free surface boundary condition

$$\hat{\mathbf{n}}\cdot\mathbf{T}^{\dagger} = \mathbf{0} \quad \text{on } \partial\Omega. \tag{B9}$$

Upon comparing (B6)–(B9) with (1)–(3), we see that the adjoint wavefield $\mathbf{s}^{\dagger}$ is determined by exactly the same wave equation and boundary conditions as the regular wavefield, with the exception of the source term: the regular wavefield is determined by the distributed noise source $\mathbf{f}$, whereas the adjoint wavefield is generated at locations $\mathbf{x}^{\alpha}$ and $\mathbf{x}^{\beta}$ by the adjoint source (B7).

Provided the adjoint wavefield $\mathbf{s}^{\dagger}$ is determined by equations (B6)–(B9), the variation in the action (B2) reduces to

$$\delta\chi = \left\langle \int_{\Omega} (\delta\ln\rho\,K_{\rho} + \delta\mathbf{c}::\mathbf{K}_{\mathbf{c}} + \delta\Sigma)\,\mathrm{d}^3\mathbf{x} \right\rangle, \tag{B10}$$

where we use the notation $\delta\mathbf{c}::\mathbf{K}_{\mathbf{c}} = \delta c_{ijkl}\,K_{c_{ijkl}}$. The kernels K_{ρ} and $\mathbf{K}_{\mathbf{c}}$ are defined by

$$K_{\rho} \equiv -\int \rho\,\mathbf{s}^{\dagger}(-t)\cdot\partial_t^2\mathbf{s}(t)\,\mathrm{d}t, \tag{B11}$$

$$\mathbf{K}_{\mathbf{c}} \equiv -\int \boldsymbol{\nabla}\mathbf{s}^{\dagger}(-t)\,\boldsymbol{\nabla}\mathbf{s}(t)\,\mathrm{d}t \tag{B12}$$

and

$$\delta\Sigma = \int \mathbf{s}^{\dagger}\cdot\delta\mathbf{f}\,\mathrm{d}t. \tag{B13}$$

The fourth-order kernel $\mathbf{K}_{\mathbf{c}}$ has the same symmetries as the elastic tensor $\mathbf{c}$. These kernels tell us the change in the misfit function $\delta\chi$ due to changes in the model parameters $\delta\ln\rho$, $\delta\mathbf{c}$ and $\delta\mathbf{f}$ in terms of the original wavefield $\mathbf{s}$ determined by (1)–(3) and the adjoint wavefield $\mathbf{s}^{\dagger}$ governed by (B6)–(B9). Because (B10) still involves taking an ensemble average, the kernels (B11) and (B12) are not practical for noise cross-correlation tomography.

APPENDIX C: ENSEMBLE-AVERAGED ADJOINT EQUATIONS AND KERNELS

To take the ensemble average in (20), let us rewrite it as an expression in the frequency domain. Using Parseval's theorem, dropping the spatial and temporal dependencies to avoid clutter, we find

$$\delta\chi \;=\; -\frac{1}{2\pi}\left\langle \iint (-\omega^2\delta\rho\,\mathbf{s}^{\dagger}\cdot\mathbf{s} + \boldsymbol{\nabla}\mathbf{s}^{\dagger}:\delta\mathbf{c}:\boldsymbol{\nabla}\mathbf{s} - \mathbf{s}^{\dagger}\cdot\delta\mathbf{f})\,\mathrm{d}^3\mathbf{x}\,\mathrm{d}\omega \right\rangle. \tag{C1}$$

In the frequency domain, the adjoint source (33) becomes

$$\mathbf{f}^{\dagger}(\mathbf{x},\omega) = [\hat{\boldsymbol{\nu}}^{\beta}\,\Delta\langle C^{\alpha\beta}\rangle(\omega)\,s^{\alpha *}(\omega)\,\delta(\mathbf{x}-\mathbf{x}^{\beta}) + \hat{\boldsymbol{\nu}}^{\alpha}\,\Delta\langle C^{\beta\alpha}\rangle(\omega)\,s^{\beta *}(\omega)\,\delta(\mathbf{x}-\mathbf{x}^{\alpha})]. \tag{C2}$$

Thus the adjoint wavefield is

$$\begin{aligned}\mathbf{s}^{\dagger}(\mathbf{x},\omega) &= \int \mathbf{G}(\mathbf{x},\mathbf{x}';\omega)\cdot\mathbf{f}^{\dagger}(\mathbf{x}',\omega)\,\mathrm{d}^3\mathbf{x}' \\ &= [\Delta\langle C^{\alpha\beta}\rangle(\omega)\,\hat{\boldsymbol{\nu}}^{\alpha}\cdot\mathbf{s}^{*}(\mathbf{x}^{\alpha},\omega)\,\mathbf{G}(\mathbf{x},\mathbf{x}^{\beta};\omega)\cdot\hat{\boldsymbol{\nu}}^{\beta} + \Delta\langle C^{\beta\alpha}\rangle(\omega)\,\hat{\boldsymbol{\nu}}^{\beta}\cdot\mathbf{s}^{*}(\mathbf{x}^{\beta},\omega)\,\mathbf{G}(\mathbf{x},\mathbf{x}^{\alpha};\omega)\cdot\hat{\boldsymbol{\nu}}^{\alpha}\,].\end{aligned} \tag{C3}$$

Upon using expressions (7) and (C3) for the regular and adjoint wavefields in (C1), we find, switching to index notation and suppressing the angular frequency dependence, that

$$\begin{aligned}
\delta\chi &= -\frac{1}{2\pi}\left\langle \iint (-\omega^2\delta\rho\, s_i^\dagger\, s_i + \nabla_i s_j^\dagger\, \delta c_{ijk\ell}\, \nabla_k s_\ell - s_i^\dagger\, \delta f_i)\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}\omega \right\rangle \\
&= -\frac{1}{2\pi}\left\langle \iiiint \{-\omega^2\delta\rho \left[\Delta\langle C^{\alpha\beta}\rangle\, \hat{\nu}_m^\alpha\, G_{mp}^*(\mathbf{x}^\alpha,\mathbf{x}')\, G_{in}(\mathbf{x},\mathbf{x}^\beta)\, \hat{\nu}_n^\beta \right.\right. \\
&\qquad\qquad \left. + \Delta\langle C^{\beta\alpha}\rangle\, \hat{\nu}_m^\beta\, G_{mp}^*(\mathbf{x}^\beta,\mathbf{x}')\, G_{in}(\mathbf{x},\mathbf{x}^\alpha)\, \hat{\nu}_n^\alpha\right] G_{iq}(\mathbf{x},\mathbf{x}'') \\
&\qquad\qquad + \delta c_{ijk\ell}\left[\Delta\langle C^{\alpha\beta}\rangle\, \hat{\nu}_m^\alpha\, G_{mp}^*(\mathbf{x}^\alpha,\mathbf{x}')\, \nabla_i G_{jn}(\mathbf{x},\mathbf{x}^\beta)\, \hat{\nu}_n^\beta \right. \\
&\qquad\qquad \left.\left. + \Delta\langle C^{\beta\alpha}\rangle\, \hat{\nu}_m^\beta\, G_{mp}^*(\mathbf{x}^\beta,\mathbf{x}')\, \nabla_i G_{jn}(\mathbf{x},\mathbf{x}^\alpha)\, \hat{\nu}_n^\alpha\right] \nabla_k G_{\ell q}(\mathbf{x},\mathbf{x}'')\} \, f_p^*(\mathbf{x}')\, f_q(\mathbf{x}'')\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{x}'\, \mathrm{d}^3\mathbf{x}''\, \mathrm{d}\omega \right\rangle \\
&\quad + \frac{1}{2\pi}\left\langle \iiint \left[\Delta\langle C^{\alpha\beta}\rangle\, \hat{\nu}_m^\alpha\, G_{mp}^*(\mathbf{x}^\alpha,\mathbf{x}')\, G_{in}(\mathbf{x},\mathbf{x}^\beta)\, \hat{\nu}_n^\beta \right.\right. \\
&\qquad\qquad \left.\left. + \Delta\langle C^{\beta\alpha}\rangle\, \hat{\nu}_m^\beta\, G_{mp}^*(\mathbf{x}^\beta,\mathbf{x}')\, G_{in}(\mathbf{x},\mathbf{x}^\alpha)\, \hat{\nu}_n^\alpha\right] f_p^*(\mathbf{x}')\, \delta f_i(\mathbf{x})\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{x}'\, \mathrm{d}\omega \right\rangle.
\end{aligned} \tag{C4}$$

Now we can take the desired ensemble average to obtain

$$\begin{aligned}
\delta\chi &= -\frac{1}{2\pi}\iiint \{-\omega^2\delta\rho\big[\Delta\langle C^{\alpha\beta}\rangle\, \hat{\nu}_m^\alpha\, G_{mp}^*(\mathbf{x}^\alpha,\mathbf{x}')\, G_{in}(\mathbf{x},\mathbf{x}^\beta)\, \hat{\nu}_n^\beta \\
&\qquad\qquad + \Delta\langle C^{\beta\alpha}\rangle\, \hat{\nu}_m^\beta\, G_{mp}^*(\mathbf{x}^\beta,\mathbf{x}')\, G_{in}(\mathbf{x},\mathbf{x}^\alpha)\, \hat{\nu}_n^\alpha\big] G_{iq}(\mathbf{x},\mathbf{x}') \\
&\qquad\qquad + \delta c_{ijk\ell}\big[\Delta\langle C^{\alpha\beta}\rangle\, \hat{\nu}_m^\alpha\, G_{mp}^*(\mathbf{x}^\alpha,\mathbf{x}')\, \nabla_i G_{jn}(\mathbf{x},\mathbf{x}^\beta)\, \hat{\nu}_n^\beta \\
&\qquad\qquad + \Delta\langle C^{\beta\alpha}\rangle\, \hat{\nu}_m^\beta\, G_{mp}^*(\mathbf{x}^\beta,\mathbf{x}')\, \nabla_i G_{jn}(\mathbf{x},\mathbf{x}^\alpha)\, \hat{\nu}_n^\alpha\big] \nabla_k G_{\ell q}(\mathbf{x},\mathbf{x}')\} \, S_{pq}(\mathbf{x}',\omega)\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{x}'\, \mathrm{d}\omega \\
&\quad + \frac{1}{2\pi}\iint \left[\Delta\langle C^{\alpha\beta}\rangle\, \hat{\nu}_m^\alpha\, G_{mp}^*(\mathbf{x}^\alpha,\mathbf{x})\, G_{in}(\mathbf{x},\mathbf{x}^\beta)\, \hat{\nu}_n^\beta \right. \\
&\qquad\qquad \left. + \Delta\langle C^{\beta\alpha}\rangle\, \hat{\nu}_m^\beta\, G_{mp}^*(\mathbf{x}^\beta,\mathbf{x})\, G_{in}(\mathbf{x},\mathbf{x}^\alpha)\, \hat{\nu}_n^\alpha\right] \delta S_{pi}(\mathbf{x}',\omega)\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}\omega,
\end{aligned} \tag{C5}$$

where we have used (14) and

$$\langle f_j(\mathbf{x}',\omega)\, \delta f_m^*(\mathbf{x}'',\omega)\rangle = \delta S_{jm}(\mathbf{x}',\omega)\, \delta(\mathbf{x}'-\mathbf{x}''). \tag{C6}$$

To bring this into a more practical form, let us define the source

$$F_i^\alpha(\mathbf{x},\omega) \equiv G_{jk}^*(\mathbf{x},\mathbf{x}^\alpha;\omega)\, \nu_k^\alpha\, S_{ij}(\mathbf{x},\omega). \tag{C7}$$

The corresponding wavefield is

$$\mathbf{\Phi}^\alpha(\mathbf{x},\omega) \equiv \int_\Omega \mathbf{G}(\mathbf{x},\mathbf{x}';\omega)\cdot\mathbf{F}^\alpha(\mathbf{x}',\omega)\, \mathrm{d}^3\mathbf{x}'. \tag{C8}$$

Upon comparing (C8) with the expression for the cross correlation (16), it may be readily shown that

$$\langle C^{\alpha\beta}\rangle(t) = \hat{\boldsymbol{\nu}}^\alpha\cdot\mathbf{\Phi}^\beta(\mathbf{x}^\alpha,t). \tag{C9}$$

Next, we define the adjoint source

$$\mathbf{F}^{\dagger\,\alpha\beta}(\mathbf{x},\omega) \equiv \hat{\boldsymbol{\nu}}^\beta\, \Delta\langle C^{\alpha\beta}\rangle(\omega)\, \delta(\mathbf{x}-\mathbf{x}^\beta). \tag{C10}$$

The associated adjoint wavefield is

$$\mathbf{\Phi}^{\dagger\,\alpha\beta}(\mathbf{x},\omega) \equiv \int_\Omega \mathbf{G}(\mathbf{x},\mathbf{x}';\omega)\cdot\mathbf{F}^{\dagger\,\alpha\beta}(\mathbf{x}',\omega)\, \mathrm{d}^3\mathbf{x}'. \tag{C11}$$

Now (C5) may be rewritten as

$$\begin{aligned}
\delta\chi &= -\frac{1}{2\pi}\iint \left[-\omega^2\delta\rho\left(\Phi_i^{\dagger\,\alpha\beta}\Phi_i^\alpha + \Phi_i^{\dagger\,\beta\alpha}\Phi_i^\beta\right) + \delta c_{ijk\ell}\left(\nabla_i\Phi_j^{\dagger\,\alpha\beta}\nabla_k\Phi_\ell^\alpha + \nabla_i\Phi_j^{\dagger\,\beta\alpha}\nabla_k\Phi_\ell^\beta\right)\right] \mathrm{d}^3\mathbf{x}\, \mathrm{d}\omega \\
&\quad + \frac{1}{2\pi}\iint \left(\Phi_i^{\dagger\,\alpha\beta}\,\delta F_i^\alpha + \Phi_i^{\dagger\,\beta\alpha}\,\delta F_i^\beta\right) \mathrm{d}^3\mathbf{x}\, \mathrm{d}\omega,
\end{aligned} \tag{C12}$$

where

$$\delta F_i^\alpha(\mathbf{x},\omega) \equiv G_{jk}^*(\mathbf{x},\mathbf{x}^\alpha;\omega)\, \nu_k^\alpha\, \delta S_{ij}(\mathbf{x},\omega). \tag{C13}$$

Finally, again using Parseval's theorem, the change in the misfit may be written in the form

$$\delta\chi = \int_\Omega (\delta\ln\rho\, \langle K_\rho\rangle + \delta\mathbf{c} :: \langle\mathbf{K}_\mathbf{c}\rangle + \langle\delta\Sigma\rangle)\, \mathrm{d}^3\mathbf{x}, \tag{C14}$$

where we have defined the ensemble kernels

$$\langle K_\rho\rangle = -\int \rho\, \left[\mathbf{\Phi}^{\dagger\,\alpha\beta}(-t)\cdot\partial_t^2\mathbf{\Phi}^\alpha(t) + \mathbf{\Phi}^{\dagger\,\beta\alpha}(-t)\cdot\partial_t^2\mathbf{\Phi}^\beta(t)\right] \mathrm{d}t, \tag{C15}$$

$$\langle \mathbf{K}_{\mathbf{c}} \rangle = -\int [\boldsymbol{\nabla}\boldsymbol{\Phi}^{\dagger\,\alpha\beta}(-t)\,\boldsymbol{\nabla}\boldsymbol{\Phi}^{\alpha}(t) + \boldsymbol{\nabla}\boldsymbol{\Phi}^{\dagger\,\beta\alpha}(-t)\,\boldsymbol{\nabla}\boldsymbol{\Phi}^{\beta}(t)]\,\mathrm{d}t, \tag{C16}$$

and

$$\langle \delta\Sigma \rangle = \int [\boldsymbol{\Phi}^{\dagger\,\alpha\beta}(-t)\cdot\delta\mathbf{F}^{\alpha}(t) + \boldsymbol{\Phi}^{\dagger\,\beta\alpha}(-t)\cdot\delta\mathbf{F}^{\beta}(t)]\,\mathrm{d}t. \tag{C17}$$

Note that the symmetries of the elastic tensor $\mathbf{c}$ imply the same symmetries in the fourth-order kernel $\langle \mathbf{K}_{\mathbf{c}} \rangle$.