Multiple scattering of high-frequency seismic waves in the deep Earth: modelling and numerical examples.

Ludovic Margerin, G. Nolet

Multiple scattering of high-frequency seismic waves in the deep Earth: Modeling and numerical examples

Ludovic Margerin
Laboratoire de Géophysique Interne et Tectonophysique, CNRS, Grenoble, France

Guust Nolet
Department of Geosciences, Princeton University, Princeton, New Jersey, USA

Received 15 May 2002; revised 28 October 2002; accepted 29 January 2003; published 6 May 2003.

[1] We apply the modern theory of radiative transfer to the modeling of the global propagation of high-frequency seismic waves in the Earth. This theory stems from an exact statistical treatment of the wave equation and incorporates rigorously the effects of multiple scattering. The statistical mean time between scattering events (the mean free time) and the typical correlation length of the random fluctuations (the scale length) are introduced as the fundamental parameters of the theory. The integro-differential equation of transport describes statistically the propagation of energy in phase space and can be conveniently solved by means of Monte Carlo simulations. We provide a general description of the method, stressing the important modifications required to adapt it to global propagation. The theory is applied to the modeling of *PKP* precursors, probably the best documented examples of wave scattering at the global scale. Guided by recent results of *Hedlin et al.* [1997], we solve the transfer equation in a variety of Earth models presenting exponentially correlated fluctuations of elastic parameters superimposed upon PREM. The validity of Born approximation is tested in a series of random media with mean free time and scale length in the 100–3200 s and 4–24 km ranges, respectively. For errors in coda envelope amplitude bound by 20%, the Born approximation can be safely applied in media with mean free times larger than about 400 s, relatively independent of the scale length. This corresponds to rather moderate (<0.5% RMS) perturbations, thus severely limiting the range of validity of Born approximation. *INDEX TERMS:* 7203 Seismology: Body wave propagation; 7207 Seismology: Core and mantle; 7260 Seismology: Theory and modeling; *KEYWORDS:* multiple scattering, precursors, radiative transfer, Born approximation, Monte Carlo method, heterogeneity

Citation: Margerin, L., and G. Nolet, Multiple scattering of high-frequency seismic waves in the deep Earth: Modeling and numerical examples, *J. Geophys. Res.*, *108*(B5), 2234, doi:10.1029/2002JB001974, 2003.

1. Introduction

[2] In recent years, a lot of efforts have been invested in the development of accurate methods of generation of synthetic seismograms, in conjunction with the rapid growth of available computational power. For example, the spectral element method provides an innovative and powerful tool to generate full seismogram waveforms [e.g., *Komatitsch and Tromp*, 1999; *Chaljub*, 2000]. We anticipate that in the foreseeable future seismologists will be able to simulate the propagation of elastic waves in the whole range of the seismic spectrum. Nevertheless, the extraction of meaningful information from seismic data remains a key problem. It is noticable, for example, that tomography, probably one of the most successful achievements of seismology, relies solely on observations of travel times of isolated phases, thereby neglecting the information contained in wave amplitudes. Recently, *Nolet and Dahlen* [2000] have shown theoretically that the resolving power of seismic tomography was severely limited by the inability of ray theory to account for diffraction and other finite frequency phenomena. The current trend in modern tomography is to take into account scattering effects in the interpretation of travel times and waveforms [*Woodward*, 1993; *Li and Tanimoto*, 1993; *Zhao and Jordan*, 1998; *Dahlen et al.*, 2000; *Hung et al.*, 2000].

[3] Clearly, the role of scattered waves in the interpretation of direct arrivals is a topic of growing interest. The interpretation of the direct arrivals themselves is more difficult. At high frequencies (>1 Hz), the correlation length of amplitude and phase is very short, typically of the same order as the wavelength [*Aki*, 1973; *Flatté and Wu*, 1988], which implies rapid spatial variations of waveforms. How-

0148-0227/03/2002JB001974$09.00

ever, in this frequency range, seismogram envelopes constitute a robust observation that is increasingly being used to provide a description of the Earth in a statistical sense [e.g., *Sato and Fehler*, 1998]. The problem of imaging in complex media is common to many areas of physics like acoustics, astrophysics, atmospheric optics, to cite a few only. In all these fields a standard tool of interpretation of scattered waves is radiative transfer theory, also known as transport theory. First developed on a phenomenological basis by astrophysicists to model the propagation of light through clouds, transport theory now stands on its own firm theoretical foundations. In particular, multiple scattering theories borrowed from the field of quantum electrodynamics have established the connection between transport and wave equations. Mathematicians have also put forward the transport equation as an exact asymptotic consequence of the wave equation [*Ryzhik et al.*, 1996].

[4] In seismology, radiative transfer theory has been introduced by *Wu* [1985] to model the spatial distribution of energy in the coda of crustal earthquakes. The theory was applied to data by *Wu and Aki* [1988], who provided the first inferences on the mean free path and absorption length of waves in the lithosphere. An important step was taken by *Gusev and Abubakirov* [1987], who incorporated time dependence and scattering anisotropy in the modeling. They also introduced the now familiar Monte Carlo method to solve the radiative transfer equation. Recent years have seen the development of more and more sophisticated models including depth-dependent scattering and velocity structures [*Hoshiba*, 1994, 1997; *Margerin et al.*, 1998] and mode coupling [*Sato*, 1994; *Zeng*, 1993; *Margerin et al.*, 2000]. In this paper, transport theory is applied to the modeling of envelope records of global seismic data. In contrast with previous investigations, ray bending and spherical geometry will be taken into account. We first present the transport equation in a seismological context with an eye on the application to the modeling of *PKP* precursors. Because mode coupling can be neglected for this particular problem (as explained in section 4), we limit ourselves to a single mode formulation of transport theory. The evolution of polarization as well as mode conversions that are so important to explain equipartition of seismic waves in the coda have been considered elsewhere [*Margerin et al.*, 2000; *Hennino et al.*, 2001]. A Monte Carlo method of solution of transport equations is presented and several examples pertaining to the modeling of *PKP* precursors are examined. A section addresses the validity range of the well-known Born approximation in connection with recent results by *Hedlin et al.* [1997] and *Cormier* [1995, 1999].

2. Radiation Transport in Stratified Media

[5] The problem of seismic wave propagation in random media is vastly more complex than that of rays propagating in a smoothly varying 3-D Earth model, where rays are defined by their initial slowness vector and where ray bundles remain ordered. In a random medium the interaction with the scatterers is such as to perturb the ray slowness vector in a random fashion. In the limit of strong scattering, waves propagate uniformly in all directions and the wave field satisfies a diffusion equation. In this paper we are dealing with the far more complicated intermediate stage, where waves experience only a few scattering events between source and receiver, and can be analyzed using the theory of radiative transfer.

[6] In the theory of radiative transfer, we discard the phase information contained in individual rays; rather, we are interested in the energy transmitted, very much like an astronomer is interested in the light intensity obtained by summing photons. A considerable simplification is reached by formulating the problem in this way. The energy flux in such a medium is then a function of location, time and direction. It is important to realize that energy may flow in all directions from one point in space, unlike the situation we know so well from ray theory in a smooth Earth. In addition, because the probability of scattering is a function of the wavelength of the wave with respect to the size of the scatterer, the intensity is also a function of frequency. In this way we reduce the number of variables for a scalar wave field to seven: three for position, two for direction, frequency and time. To conserve the phase information, we would have to formulate a finite difference or finite element scheme with billions of variables to be handled in a computer at the same time. The price we pay for our simplification is that we can only construct the envelope of the seismic signal.

[7] In this section, we derive from a phenomenological viewpoint the transport equation for acoustic waves taking into account important specificities of seismology like reflection/transmission at major discontinuities inside the Earth and the continuous refraction of rays in smooth parts of the reference medium. This level of presentation will be sufficient for our purposes. Only when necessary, we will borrow some results from the complete elastic theory as exposed by *Weaver* [1990] and *Ryzhik et al.* [1996].

2.1. Physical Assumptions

[8] Before going to a detailed presentation, it is important to outline the physical assumptions underlying the approach. We assume that small fluctuations of the elastic parameters are superimposed on a piecewise smooth background medium. For example, in an acoustic medium, the velocity v will be decomposed as $v(\mathbf{R}) = v_r(\mathbf{R})(1 + \alpha(\mathbf{R}))$, where $\mathbf{R}$ is the position vector, $v_r(\mathbf{R})$ is the wave speed in a reference (unperturbed) background medium like PREM [*Dziewonski and Anderson*, 1981] and $\alpha(\mathbf{R})$ represents a fluctuation with zero statistical mean. Central to the understanding of transport theory is the idea of separation of scales. On the one hand, we assume that the typical length scale over which the field v_r varies significantly is much larger than the wavelength. On the other hand, the typical scale over which the field α varies can be of the same order as the wavelength, which entails scattering of the waves by the inhomogeneities. The fact that the wavelength is considered otherwise as a small quantity allows us to use ray theory to predict the evolution of the wave field between scattering events. For a precise mathematical formulation of the scaling relations, we refer to *Ryzhik et al.* [1996]. The fluctuating part α will be described by a correlation function $\Gamma(\mathbf{R}, \mathbf{r}) = \langle \alpha(\mathbf{R} - \mathbf{r}/2)\alpha(\mathbf{R} + \mathbf{r}/2) \rangle$, where angle brackets denote an ensemble or statistical average. This function can be expanded in a Fourier integral:

$$\Gamma(\mathbf{R}, \mathbf{r}) = \int \Phi_\alpha(\mathbf{R}, \mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3\mathbf{k}. \quad (1)$$

Provided the field fluctuations are quasi-homogeneous, i.e., exhibit fast variations in the **r** variable as compared to the **R** variable, Φ_α can be interpreted as the local power spectrum of inhomogeneities. Thus it is possible to allow for slow variations of the correlation function along the rays to model possible statistical inhomogeneity of the medium.

2.2. Specific Intensity and Related Quantities

[9] The fundamental quantity describing the radiation field is the so-called specific intensity or radiance defined as

$$I(\omega, t, \mathbf{R}, \mathbf{n}) = \frac{dE}{d\Omega(\mathbf{n}) d\omega |\mathbf{n} \cdot d\mathbf{S}| dt}, \tag{2}$$

where dE is the amount of energy in the solid angle $d\Omega(\mathbf{n})$ around space direction **n** within the frequency range $[\omega, \omega + d\omega]$ flowing through the surface dS during the time interval dt. It is an angularly resolved flux whose direction of propagation can be tracked through the medium. Specific intensity is not directly accessible to seismologists but rather the local energy density defined as

$$\rho(t, \mathbf{R}) = \int_{\Delta\omega} d\omega \int_{4\pi} \frac{1}{v_g(\omega, \mathbf{R})} I(\omega, t, \mathbf{R}, \mathbf{n}) d\Omega(\mathbf{n}), \tag{3}$$

where $\Delta\omega$ is the frequency band of the radiation, v_g is the group speed of a wave packet with central frequency ω and 4π denotes the whole sphere of space directions. Note that in this definition, the group velocity is independent of the propagation direction of the waves. This implies that the background medium as well as the statistical properties of the fluctuations are isotropic, which will suffice for our applications. In absence of fluctuations ($\alpha \sim 0$) and of temporal dispersion v_r and v_g obviously coincide. Another quantity of interest is the energy current vector defined as

$$\mathbf{J}(t, \mathbf{R}) = \int_{\Delta\omega} d\omega \int_{4\pi} I(\omega, t, \mathbf{R}, \mathbf{n}) \mathbf{n} d\Omega(\mathbf{n}), \tag{4}$$

so that the net rate of energy flow through an arbitrary surface $d\mathbf{S}$ is given by $\mathbf{J} \cdot d\mathbf{S}$.

2.3. Variation of Intensity in Free Space

[10] The first point we wish to examine is the variation of intensity induced by the propagation in the background medium. Thus we leave aside for the moment the effects of scattering ($\alpha \sim 0$) to concentrate on the effects of ray bending caused by v_r. To address this point, it is interesting to note that the specific intensity at a point can also be viewed as a set of quasi-plane waves having random phases and propagating independently. To study the variation of intensity due solely to the velocity gradients in the reference medium, we assume that each quasi-plane wave propagates according to the laws of geometrical optics. Suppose we want to follow a beam of energy defined by its extension ($\Delta\mathbf{R}_0$, $\Delta\omega_0$, $\Delta\Omega_0$) in phase space at $t = t^0$. At time t it occupies another volume of phase space ($\Delta\mathbf{R}(t)$, $\Delta\omega(t)$, $\Delta\Omega(t)$). When absorption is absent, the total energy E carried by the beam is conserved as it evolves through the medium and should therefore verify $dE/dt = 0$. Using the definitions of specific intensity and energy density, this can be rewritten as

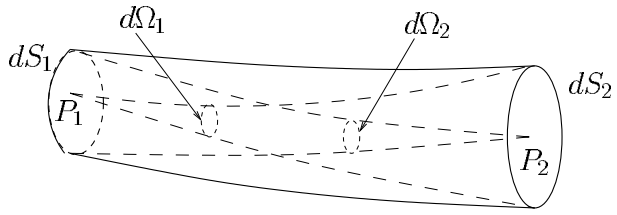

Figure 1. Schematic view of the evolution of a beam of energy in a smooth medium. The energy flowing through the elementary surface dS_1 (respectively dS_2) consists of a set of quasi-plane waves propagating in the small solid angle $d\Omega_2$ (respectively $d\Omega_1$) subtended by dS_1 (respectively dS_2) at P_2 (respectively P_1).

$$\frac{d}{dt} \int_{\Delta\omega(t)} d\omega \int_{\Delta\mathbf{R}(t)} d^3\mathbf{R} \int_{\Delta\Omega(t)} \frac{1}{v_g(\omega, \mathbf{R})} I \cdot (\omega, t, \mathbf{R}, \mathbf{n}) d\Omega(\mathbf{n}) = 0. \tag{5}$$

In Appendix A, we show that in a stationary, isotropic medium this implies the following:

$$\frac{dI}{dt} + \frac{d \log(v_r^2)}{dt} I = 0, \tag{6}$$

and accordingly, the generalized intensity $\mathcal{I} = v_r^2 I$ is conserved along the ray path. We can give this result a simple physical interpretation. Consider a beam traveling through two surfaces S_1 and S_2 perpendicular to the ray at locations P_1 and P_2, respectively (see Figure 1). For the sake of simplicity we have dropped the time and frequency dependence in this illustration. By definition of a beam, the energy flowing through the two surfaces should be equal. Now the specific intensities at P_1 and P_2 are

$$I_1 = \frac{dE}{dS_1 d\Omega_1}, \quad I_2 = \frac{dE}{dS_2 d\Omega_2}, \tag{7}$$

respectively. In equation (7), $d\Omega_{1(2)}$ denotes the solid angle subtended by $dS_{1(2)}$ at $P_{1(2)}$ (see Figure 1). By definition, the geometrical spreading $R_{i \to j}$ along the ray connecting P_i to P_j satisfies the following relation:

$$\frac{dS_1 d\Omega_1}{dS_2 d\Omega_2} = \frac{R_{2\to1}^2}{R_{1\to2}^2} \tag{8}$$

From ray theory, we know that the geometrical spreading in 3-D obeys the reciprocity relation $v_1 R_{1\to2} = v_2 R_{2\to1}$ [*Dahlen and Tromp*, 1998; *Snieder and Chapman*, 1998] which imposes

$$v_1^2 I_1 = v_2^2 I_2, \tag{9}$$

where v_0 and v_1 are the velocities at P_1 and P_2, respectively. Hence equation (6) simply reflects the geometrical effects of focusing/defocusing. In the case of a homogeneous medium, the specific intensity itself is conserved along the ray which makes it a fundamental quantity in optical measurements.

2.4. The Transfer Equation

[11] From a phenomenological viewpoint, the transfer equation results from a detailed balance of energy as we

follow a wave beam propagating in direction **n** along an elementary ray path. Because of the interaction with the scatterers, the beam distributes energy over all space directions and this entails the following change of intensity δI over the time interval δt:

$$\delta I = -\frac{\delta t}{\tau} I, \tag{10}$$

where the scattering mean free time τ depends on the wavelength, the correlation length of the inhomogeneities and the RMS perturbations as explained in section 2.6. The scattering mean free time can be viewed as the typical timescale over which the coherent part of the wave field attenuates or alternatively as the statistical mean time between two scattering events. We also have to consider that, upon scattering, the wave beams propagating locally in space direction $\mathbf{n}'$ can transfer energy to direction **n**, which results in a gain term:

$$\delta I = \frac{\delta t}{\tau} \int_{4\pi} p(\mathbf{n}, \mathbf{n}') I(\omega, t, \mathbf{R}, \mathbf{n}') d\Omega(\mathbf{n}'). \tag{11}$$

The prefactor $\delta t/\tau$ represents the relative loss of each beam during propagation. The normalized phase function $p(\mathbf{n}, \mathbf{n}')$ describes the angular anisotropy of the scattering and can be interpreted as the probability for a local plane wave propagating in direction $\mathbf{n}'$ to be scattered into direction **n**. As will be seen in section 2.6, the phase function is completely determined by the choice of correlation function. Considering the possible presence of sources of intensity $\mathcal{S}$ (ω, t, **R**, **n**) and collecting the results of equations (6), (10), and (11), we obtain the transport equation

$$\left(\frac{d}{dt} + \frac{d \log v_r^2}{dt} + \frac{1}{\tau}\right) I(\omega, t, \mathbf{R}, \mathbf{n}) = \mathcal{S}(\omega, t, \mathbf{R}, \mathbf{n}) + \frac{1}{\tau} \int_{4\pi} p(\mathbf{n}, \mathbf{n}') I(\omega, t, \mathbf{R}, \mathbf{n}') d\Omega(\mathbf{n}'). \tag{12}$$

It is seen that the different contributions are simply added which reflects the additivity property of intensities in random media. We emphasize that the derivative d/dt in equation (12) has to be interpreted as a total or "material derivative" and can be expanded as

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \nabla_{\mathbf{R}} + \frac{d\mathbf{n}}{dt} \cdot \nabla_{\mathbf{n}}, \tag{13}$$

where we have taken into account the conservation of frequency along the ray. The derivatives $d\mathbf{R}/dt$, $d\mathbf{n}/dt$ can be further explicated with the aid of the ray equations (see Appendix A). The d/dt operator is similar to the "material derivative" of fluid mechanics which physically means that we follow the beam of intensity along the ray path.

2.5. Boundary Conditions at Interfaces

[12] As a wave beam encounters an interface, the energy of each plane wave is partitioned into reflected and transmitted waves. In addition, we have to consider that the beam as a whole undergoes a change of solid angle. In order to derive the equations connecting the incoming and outgoing intensities, we write a balance of energy on a small portion $d\mathbf{S}$ of the interface (see Figure 2). The incoming wave is incident from the upper part of the medium with velocity V_1 and is transmitted into the lower part of the medium with velocity V_2. In the general case, the interface may be curved and will be approximated locally by its tangent plane. Once again, the laws of geometrical optics are used to deal with each quasi-plane wave composing the beam. Henceforth, the different physical quantities are accompanied with subscripts i, r, t to identify the incident, reflected and transmitted waves, respectively. In order to verify the conservation of energy at the boundary, the sum of the fluxes of the outgoing waves must be balanced by the flux of the incoming waves:

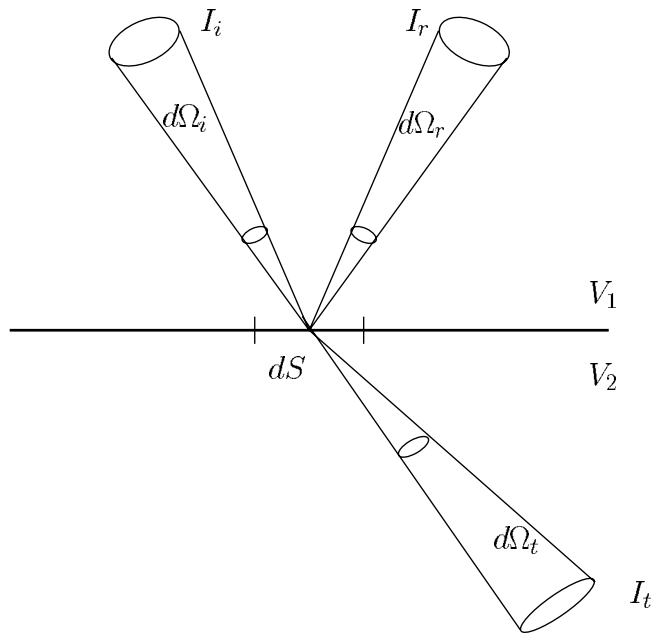

Figure 2. Reflection and transmission of specific intensity at a boundary between media presenting a wave speed contrast V_1/V_2. The beam is incident from the upper medium with velocity V_1 and is partially reflected and transmitted. The incident, reflected, and transmitted waves are labeled with subscripts i, r, and t, respectively. I and $d\Omega$ denote the specific intensities and associated solid angles.

$$I_i |\mathbf{n}_i \cdot d\mathbf{S}| d\Omega(\mathbf{n}_i) = I_r |\mathbf{n}_r \cdot d\mathbf{S}| d\Omega(\mathbf{n}_r) + I_t |\mathbf{n}_t \cdot d\mathbf{S}| d\Omega(\mathbf{n}_t). \tag{14}$$

Neglecting to first order the effects of surface curvature, we have $d\Omega(\mathbf{n}_r) = d\Omega(\mathbf{n}_i)$. Using Snell's law, the change of solid angle upon transmission can be written as

$$\frac{d\Omega(\mathbf{n}_t)}{d\Omega(\mathbf{n}_i)} = \frac{\sin\theta_t V_2 \cos\theta_i}{\sin\theta_i V_1 \cos\theta_t}, \tag{15}$$

where θ denotes the angle **n** makes with $d\mathbf{S}$. Injecting the last relation into equation (16) and expressing the flux ratios through the energy reflection and transmission coefficients

$\mathcal{R}$ and $\mathcal{T}$ [see, e.g., *Ben-Menahem and Singh*, 2001], we obtain the desired boundary conditions:

$$I_t = \frac{V_1^2}{V_2^2}\mathcal{T} I_i \tag{16}$$

$$I_r = \mathcal{R} I_i. \tag{17}$$

Equation (16) can be given a simple interpretation: the ratio of velocities models the focusing/defocusing of the beam as a whole, while the transmission coefficient accounts for the change of energy flux of each individual plane wave upon propagation through the interface.

2.6. Wave Content of the Transfer Equation

[13] So far, we have based our reasoning on phenomenological concepts, disregarding important problems such as the connection between the specific intensity and the field itself. However, it is important to note that the derivation of the transfer equation can be made completely rigorous using multiple-scattering theory [*Rytov et al.*, 1989; *Lagendijk and Van Tiggelen*, 1996; *Apresyan and Kravtsov*, 1996]. Moreover, the theory can be extended to elastic wave propagation incorporating rigorously the effects of polarization and mode coupling [*Weaver*, 1990; *Ryzhik et al.*, 1996]. In particular, they show that the radiance is essentially the Fourier transform of the coherence function of the wave field.

[14] For our seismological applications, it is important to relate the parameters of the transfer equation to the statistical description of the medium fluctuations. Provided the perturbations are weak enough, it is possible to show [*Rytov et al.*, 1989] that for scalar waves with central frequency ω_0, the mean free time and phase function or scattering radiation pattern can be written as

$$\frac{1}{\tau(\mathbf{R})} = v_g(\omega_0, \mathbf{R}) \int_{4\pi} 2\pi k_0^4(\mathbf{R}) \Phi_\alpha(\mathbf{R}, k_1(\mathbf{n} - \mathbf{n}')) d\Omega(\mathbf{n}), \tag{18}$$

$$p(\mathbf{R}, \mathbf{n}, \mathbf{n}') = \frac{\Phi_\alpha(\mathbf{R}, k_1(\mathbf{n} - \mathbf{n}'))}{\displaystyle\int_{4\pi} \Phi_\alpha(\mathbf{R}, k_1(\mathbf{n} - \mathbf{n}') d\Omega(\mathbf{n})}, \tag{19}$$

respectively. Because the medium is assumed statistically isotropic, neither τ nor p depends explicitly on $\mathbf{n}'$. Note that a subtle difference exists between k_0 and k_1. The former is the wave number at frequency ω_0 in the reference medium while the latter is the real part of the renormalized wave number at the same frequency in the random medium [*Sheng*, 1995; *Rytov et al.*, 1989]. When perturbations are weak, the two quantities differ very little. Both τ and p may depend slowly on $\mathbf{R}$ either due to a change of the wave number or to the statistical inhomogeneity of the medium. Because the mean free time represents the relative loss of energy per unit time of propagating waves, it measures in a meaningful way the scattering strength of the medium. We shall therefore prefer this parameter to the often used RMS perturbations to qualify the scattering medium.

[15] The reader may have recognized that the integral in equation (18) is nothing more than the scattering cross section per unit volume obtained in the Born approximation, provided the slight difference between k_0 and k_1 is neglected [*Rytov et al.*, 1989]. The phase function is also seen to be proportional to the power spectrum of inhomogeneities evaluated at $k_1(\mathbf{n} - \mathbf{n}')$, the exchange vector between the incoming and outgoing propagation directions. This coincides again with the predictions of the Born approximation, with the same comment as above. These results may be generalized to elastic waves provided one takes care of all the possible correlations among the different elastic parameters and mode conversions. In practice, this means that the solution of the Born approximation [e.g., *Wu and Aki*, 1985] suffices to compute all the parameters of the transport equation. For completeness, we present in Appendix B expressions for the mean free time of elastic waves in media with exponential correlations and velocity-type perturbations, as defined by *Wu and Aki* [1985]. It is worth mentioning that as a consequence of equations (18) and (19), the transport equation reduces to the Born approximation when the propagation time t is much smaller than the mean free time. The requirement $t/\tau \ll 1$ ensures that the losses of the primary field are small and that multiple scattering can be neglected [*Ishimaru*, 1978].

3. Monte Carlo Solution of Transport Equations

[16] In this section, we briefly describe the numerical scheme adopted to solve the transport equation. As is well known, Monte Carlo simulations constitute a standard method of solution of complex transport problems involving a large phase space. In the mathematical literature it is possible to find rigorous justifications of the method although the proofs are rather involved [*Bal et al.*, 2000; *Lapeyre et al.*, 1998]. Heuristic arguments can be found in the physical literature [*Lux and Koblinger*, 1991], and in what follows we shall adhere to this level of rigor.

3.1. Outline of the Simulation

[17] A natural and transparent way of understanding Monte Carlo simulations relies on the observation that the transport equation is completely analogous to the Boltzmann equation of the kinetic theory of gases. Therefore, by simulating the trajectories of a large number of particles it is possible to estimate average values of various physical quantities of interest such as flux or energy density. This approach has been adopted in a recent paper by *Yoshimoto* [2000]. Below, we briefly describe the logical organization of the simulation of the space-time evolution of the energy in the scattering medium.

[18] Each particle is described by a set of variables that completely specify its state, that is, the position vector $\mathbf{R}$, the propagation direction $\mathbf{n}$, the propagation time t, and a weight w. This weight enables the modeling of absorption due to anelasticity, or reflection/transmission at boundaries. We shall henceforth assume that the waves have central frequency ω_0 and will drop the frequency dependence. The simulation of one particle's history is composed of several steps, as described below:

[19] 1. The particle is launched at the source region, which consists in randomly selecting its takeoff time t_0,

its initial position $\mathbf{R}_0$, and its initial propagation direction $\mathbf{n}_0$, according to the definition of the source function.

[20] 2. The flight time T_1 to the first collision is determined according to the definition of the mean free time. The ray equations are integrated from time $t = t_0$ up to time $t_1 = t_0 + T_1$. The state variables of the particle assume the new values $\mathbf{R}_1$, $\mathbf{n}_1^-$, t_1.

[21] 3. At this stage, a scattering event occurs and the propagation direction jumps to the new value $\mathbf{n}_1^+$ according to the definition of the differential scattering cross section.

[22] 4. The process is repeated as follows: a propagation time T_i from the $(i-1)$th to the ith collision is selected. The integration of the ray equations provides the evolution of the particle from $\mathbf{R}_{i-1}$, $\mathbf{n}_{i-1}^+$, t_{i-1} to $\mathbf{R}_i$, $\mathbf{n}_i^-$, t_i, where the particle jumps from $\mathbf{n}_i^-$ to $\mathbf{n}_i^+$ because of scattering, and so on, until the particle escapes the space-time domain where we wish to compute the solution of the transport equation.

[23] 5. The particle is monitored during propagation and when it reaches some region of interest its contribution to the flux or energy density is calculated.

[24] 6. After simulation of a large number of independent such random walks the contributions of each particle are averaged to obtain an estimate of energy flux.

3.2. Simulation of Random Vectors

[25] After this brief introduction, we examine at greater depth each step of the simulation. The modeling of random variables is at the heart of any Monte Carlo algorithm,and we will examine this point in a general context. Suppose we wish to model a random vector X_i, $1 \le i \le N$, with joint probability density $w_n(x_1; \cdots ;x_N)$, where the subscript denotes the number of components. In the general case where the different variables X_i are not independent, it is always possible to use an elementary decomposition of w_N into conditional probabilities. Denoting by x_i, the value taken by the random variable X_i, the probability of the event $\cap_{1\le i\le N}\,(x_i \le X_i < x_i + dx_i)$ is written as

$$P(x_1 \le X_1 < x_1 + dx_1; \cdots ; x_N \le X_N < x_N + dx_N) = w_1(x_1)w_2(x_1|x_2)\cdots w_n(x_1; \cdots ; x_{N-1}|x_N)dx_1 \cdots dx_N \qquad (20)$$

where $w_k(x_1; \cdots ;x_{k-1}\,|x_k)$ is the conditional probability of X_k given $X_1, \cdots X_{k-1}$. This decomposition offers a natural way of simulating a complex PDF by reading the right-hand side from left to right. We proceed first with the simulation of X_1, then with the simulation of X_2 conditioned on X_1, then with the simulation of X_3 given X_1, X_2, and so on, up to X_N. Hence the simulation of a random vector can always be simplified into the much simpler repeated selection of random variables. A number of clever algorithms have been devised to generate random numbers distributed according to some prescribed PDF and we refer the reader to *Lux and Koblinger* [1991] for extensive references on the topic. Below, we provide the probabilistic interpretation of the transport process from a physical point of view. We will not dwell on a detailed description of the algorithm itself since this has already been presented a number of times in the seismological literature [*Gusev and Abubakirov*, 1987; *Hoshiba*, 1994, 1997; *Yoshimoto*, 2000; *Margerin et al.*, 2000].

3.3. Simulation of the Transport Process

3.3.1. Source

[26] In transport theory, the source is modeled by a function $\mathcal{S}\,(t, \mathbf{R}, \mathbf{n})$ prescribing the flux of particles leaving position $\mathbf{R}$ at time t in direction $\mathbf{n}$. For convenience, we demand that S be normalized as follows:

$$\int dt \int d^3\mathbf{R} \int S(t, \mathbf{R}, \mathbf{n}) d\Omega(\mathbf{n}) = 1. \qquad (21)$$

S can thus be interpreted as the joint probability density function of the variables $\mathbf{R}$, $\mathbf{n}$ and t. In order to launch a particle, we require the simulation of a random vector ($\mathbf{R}_0$, $\mathbf{n}_0$, t_0) according to the PDF S. In the numerical examples presented below, the source is assumed to be isotropic and point-like, thus $S(t, \mathbf{R}, \mathbf{n}) = 1/4\pi\delta(\mathbf{R} - \mathbf{R}_0)s(t)$, with $\int_{-\infty}^{+\infty} s(t)dt = 1$. In this case, the simulation greatly simplifies since the variables are independent and the problem becomes axisymmetric. The assumption is not too restrictive for the *PKP* precursor modeling, since these waves correspond to rays that leave the source in a tiny solid angle. In addition, the data are generally averaged over different paths, which tends to wash out the details of the radiation pattern.

3.3.2. Propagation and Scattering

[27] After a particle has been launched or scattered, we need to select its flight time T_i to the next scattering position. In what follows, we assume that the ray equations can be numerically solved in the background medium. In the simple case where the mean free time τ is constant, the intensity I decays exponentially along a ray according to $e^{-t/\tau}$. It is straightforward to generalize this result to the case where the mean free time is a function of the position along the ray $\mathbf{R}$, by simple integration of equation (10). The probability of scattering in the time interval $[T, T + dT]$ thus reads

$$P(T \le T_N < T + dT) = \frac{e^{-\int_0^T \frac{dt}{\tau[\mathbf{R}(t)]}}}{\tau[\mathbf{R}(T)]} dT, \qquad (22)$$

where the integral is taken along the ray path in the background medium.

[28] When the particle encounters an interface between media with different wave speeds, the boundary conditions (16)–(17) have to be implemented. The change of solid angle due to the deterministic refraction is modeled by imposing Snell's law. The partition of energy at the boundary is treated as a Bernouilli process with probability of reflection $\mathcal{R}$ and probability of transmission $\mathcal{T} = 1 - \mathcal{R}$.

[29] From equation (19), we interpret $p(\mathbf{n}, \mathbf{n}')$ as the probability for a particle propagating in direction $\mathbf{n}'$ to be scattered into direction $\mathbf{n}$. We note that for acoustic or P to P scattering in a statistically isotropic medium, the phase function depends on the cosine of the scattering angle $\mathbf{n} \cdot \mathbf{n}'$ only, independent of the type of perturbations. As usual, the new propagation direction $\mathbf{n}_i^+$ is first randomly selected in a local orthogonal frame with basis vectors ($\mathbf{x}$, $\mathbf{y}$, $\mathbf{n}_i^-$) (note that the choice of $\mathbf{x}$ and $\mathbf{y}$ is arbitrary). The coordinates of $\mathbf{n}_i^+$ are then rotated to keep track of the propagation vector in a global frame.

3.3.3. Detection

[30] We monitor the evolution of the particle in the medium and calculate its contribution to the energy flux at the surface. This is simply done by dividing the surface of the Earth into patches of elementary area ΔA and by dividing the time into intervals of constant width Δt, small compared to the mean free time. Each time a particle impinges on the Earth surface at θ, ϕ and time t, its weight $w_i^{(\theta,\phi)}$ is added to the corresponding space-time bin and the average local flux J is estimated as

$$J(\theta,\phi) = \frac{1}{M}\sum_{i=1}^{M}\frac{w_i^{(\theta,\phi)}}{\Delta A}, \tag{23}$$

where M is the total number of simulated particles.

3.3.4. Born Approximation

[31] Monte Carlo simulations of the transport equation enable the computation of the Born approximation in a flexible way. Because Born theory is a single-scattering approximation, the particle is allowed to be scattered only once. Doing so, we obtain the first order term of a multiple-scattering series which does not coincide with the Born approximation unless the mean free time is very large ($\tau \rightarrow \infty$). The precise reason for this difference stems from the fact that the Born approximation neglects the energy losses due to scattering along the wave path. The scattering loss L_s is simply computed as the integral of equation (10) along the ray path. By weighting the particles with $1/L_s$, we exactly compensate for the scattering attenuation and therefore obtain the Born approximation. Simple consistency checks of the method are obtained by remarking (1) that the Born solution scales linearly with the perturbations squared and (2) that the unscattered energy is completely independent of the perturbations and should match the predictions of ray theory. We have verified that our Born solution behaves so.

4. Application to *PKP* Precursors

[32] Probably the best documented observation of wave scattering in the deep Earth are the so-called precursors of the *PKP* phase. In recent work [*Hedlin et al.*, 1997; *Hedlin and Shearer*, 2000], Hedlin and his coworkers have performed a thorough analysis of these precursors and we will summarize their findings here. The precursors are scattered wave trains preceding the *df* branch of *PKP* that are usually observed in the 124–142° epicentral distance range. They constitute a unique example of scattering because in this case the diffracted energy comes in before the main phase of the seismogram and is not blurred by the coda of any other phase. Seismologists now agree that the precursors stem from the scattering of the PKP_{ab} and PKP_{bc} phases by inhomogeneities in the lower mantle. More precisely, the onset time of the precursors usually coincides fairly well with the predicted minimum arrival time for scattered PKP_{ab} and PKP_{bc} waves, if one assumes that scattering begins just above the core-mantle boundary (CMB).

[33] A major and original contribution of *Hedlin et al.*'s [1997] work is their quantitative analysis of the space-time dependence of the energy of the precursors. First, they measured the root mean square amplitudes of the precursors as a function of time and epicentral distance by averaging the energy envelopes of a large collection of precursor's records. Using the Born approximation, they were able to predict theoretically the energy distribution of the precursors assuming that heterogeneities with exponential or Gaussian correlation function are superimposed upon PREM. They reached the thought-provoking conclusion that the often invoked D″ layer is unable to explain the observed duration of the precursors. They proposed that small-scale heterogeneities with a correlation length of about 8 km and fluctuations of about 1% would be present at all depths in the lower mantle. It is interesting to note that the geometry of the wave paths does not allow a complete reconstruction of the perturbations that give rise to scattering. The precursors correspond to wave that are mostly deflected in a small cone around the forward direction. Hence we will be insensitive to impedance-type perturbations that give rise to backward scattering as described by *Wu and Aki* [1985]. For this particular reason, only velocity-type perturbations will be investigated in this paper.

[34] Of course, if scatterers are present throughout the mantle, one might question the validity of single-scattering approximations. As explained before, the transport equation is an exact consequence of the wave equation that incorporates the effects of multiple scattering. The Monte Carlo simulations will thus provide a reference solution against which the Born approximation can be tested. As already outlined before, the Born approximation is recovered from transport theory in the limit of weak scattering ($\tau \rightarrow \infty$). Before going to a detailed comparison, we will first present some simplifications of the physics that were introduced in order to simplify the numerical work.

4.1. Model Assumptions

[35] We wish to solve the transport equation in an Earth model that closely resembles that proposed by *Hedlin at al.* [1997]. However some difficulties arise in doing so because the continuous change of velocity in the Earth results in depth dependent mean free times and phase functions. In Figure 3, we show the exact depth dependence of the mean free time calculated at a central frequency of 1.3 Hz in a lower mantle containing heterogeneities with 4, 5, 8, or 16 km length scale and 1% RMS perturbations. The mean free time thus decreases slightly with the radius within the Earth. For heterogeneities with 8 km length scale, the ratios between the mean free time at the CMB, and the mean free time 1200 km and 2000 km above the CMB, are approximately 1.08 and 1.25 respectively. It seems therefore reasonable to approximate the exact radial dependence with a whole mantle mean value. We also assume like *Hedlin et al.* [1997] that scattering occurs on receiver side (or source side) only.

[36] Another difficulty arises from the continuous change of the angular dependence of the scattering caused by the variation of the product wave number times length scale along a ray. In order to quantify in a simple way the change of scattering anisotropy with depth, we introduce the mean cosine of the scattering angle defined as

$$g(\mathbf{R}) = \int_{-1}^{1} p(\mathbf{R},\mu)\mu \, d\mu, \tag{24}$$

where μ is the cosine of the scattering angle. The dependence of the phase function on the cosine of the

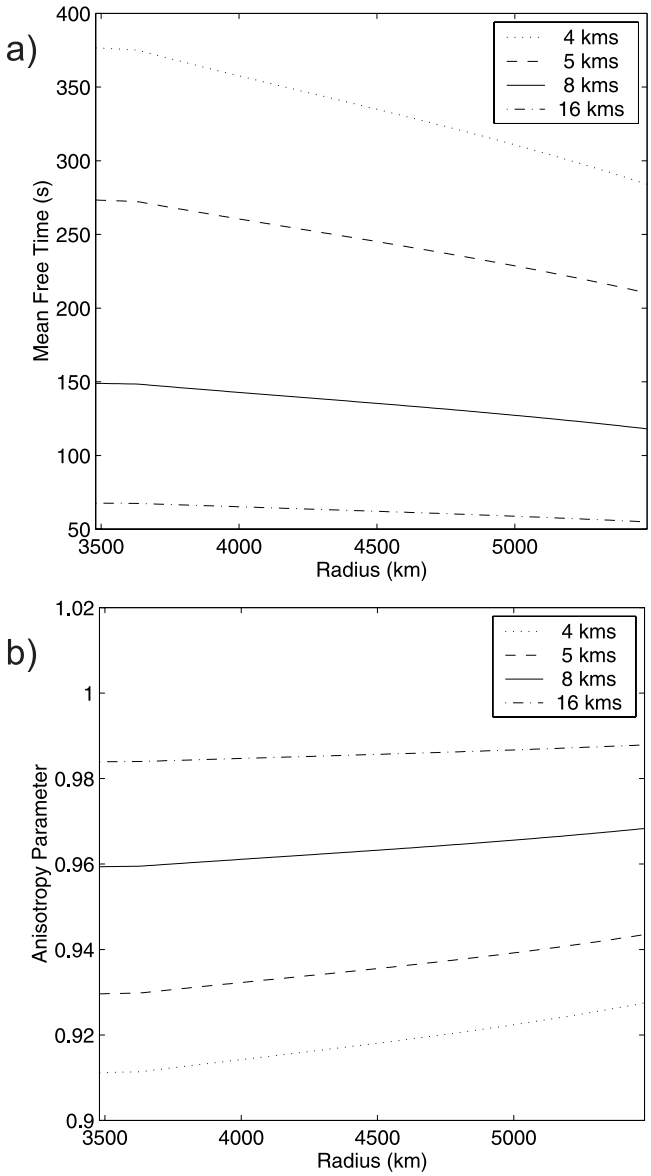

Figure 3. Multiple scattering parameters of Hedlin-like models. (a) Mean free time. (b) Mean cosine of scattering angle.

scattering angle solely has been made explicit in equation (24). Obviously, g lies in the interval] - 1, 1 [. When g is positive, scattering is predominantly in the forward direction. In Figure 3, it is seen (1) that the depth dependence of g in Hedlin like models is rather weak and (2) that g is very close to 1 indicating a strong tendency to forward scattering. This implies (1) that a mean correlation length can be defined for the whole mantle and (2) that the probability of backward scattering is extremely weak. Wave propagation in the mantle can therefore me modeled as a (possibly multiple) forward scattering process. This last assumption has also been implicitly used by *Hedlin et al.* [1997].

[37] Until now, the elastic nature of the medium has somewhat been overlooked in our presentation. In order to discuss the importance of P to S mode coupling, we show in Figure 4 the P to P, P to S and total mean free time as functions of the correlation length. For length scales larger than 4 km, P to P scattering largely dominates over P to S mode conversions. This justifies the acoustic treatment of the scattering used in this study. However, the discontinuous change of elastic properties at major discontinuities inside the Earth imposes strong mode conversions. In order to take these effects into account, we have made use of the exact elastic reflection and transmission coefficients in the modeling. Since S to P conversions are extremely weak, the part of the energy that is converted into shear energy at a boundary is considered as irreversibly lost. In the Monte Carlo simulation, this is appropriately accounted for by ascribing to the particle a weight which is proportional to the P to P transmission coefficient.

[38] To summarize, we model the energy transport in the mantle as a multiple forward scattering process, taking accurately into account the effects of ray curvature and mode conversions at interfaces, but neglecting the effect of the velocity gradients on the scattering parameters. As explained above, this is not too strong a simplification because the changes of mean free time and phase function with depth are rather moderate in the lower mantle.

4.2. Some Numerical Examples

[39] The reader may have noticed that the Monte Carlo method never asks for the computation of dynamic properties such as geometrical spreading. This constitutes a big advantage for the problem of *PKP* precursors where the incident wave field has a singularity at the b caustic. Of course this singularity is integrable but it requires a careful scheme of numerical integration. To demonstrate on a simple example that the Monte Carlo simulations incorporate in a natural way the strong effects of focusing/defocusing, we have computed the decay of energy flux of the *ab* and *bc* branches of *PKP* from the caustic to 160° epicentral distance in a perfectly transparent Earth. In this case, no scattering occurs and the solution of the transport equation should then reduce to the predictions of ray theory. In order to speed up the Monte Carlo ray-tracing procedure, the PREM model has been approximated with 134 layers of constant velocity gradient. In this case, analytical formulas for ray tracing can be derived [e.g., *Nolet*, 1981], which avoids the costly numerical integration of the ray equations. We have developed independently a dynamic ray-tracing code that provides a simple "exact" reference solution in PREM against which the Monte Carlo simulation can be tested. Because of its inherently discrete character, the simulation enables the calculation of the flux at a discrete set of epicentral distances only. The comparison for an isotropic point source with unit total energy is shown in Figure 5. The agreement between the two approaches is seen to be excellent. The Monte Carlo method catches rather well the divergence of the flux in the neighborhood of the caustic, in spite of the crude approximation of the PREM model. We emphasize that this comparison is one-to-one with no adjustable parameter.

[40] Another noticable advantage of our method is its ability to provide the full seismogram envelope in a single step. Thus contrary to Hedlin's approach, we do not compute (1) the precursor amplitudes using Born approximation and (2) the $cd + df$ amplitudes using ray theory.

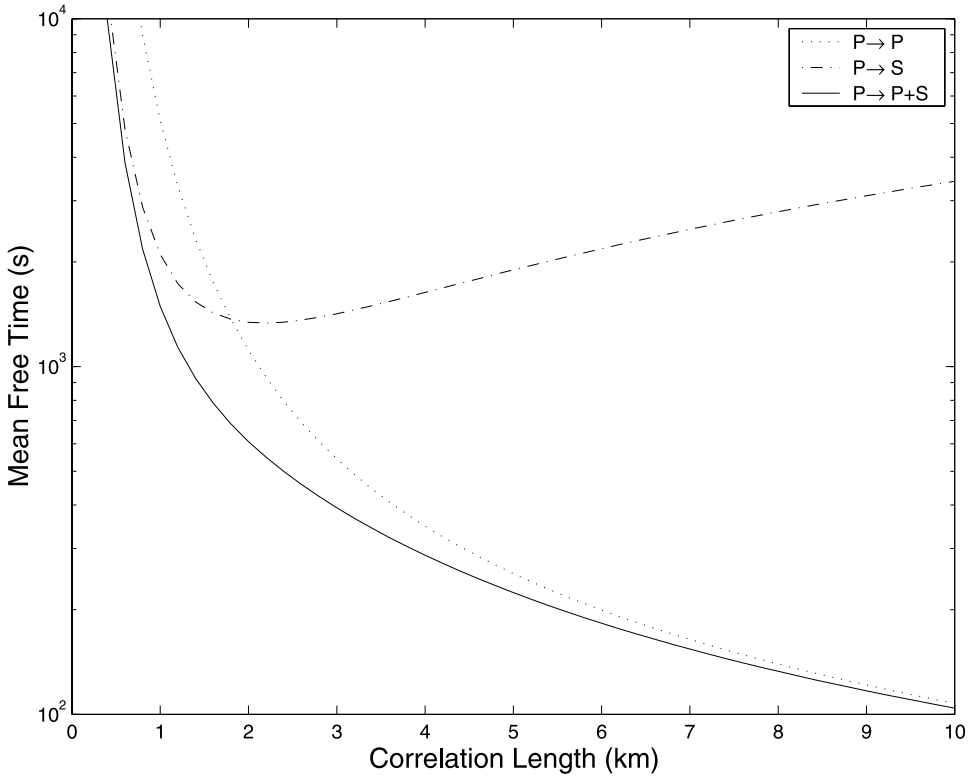

Figure 4. Mode conversions as a function of length scale in a Hedlin-like model. Dotted curve *P-P* scattering; dash-dotted curve, *P-S* scattering; solid curve, total. The effect of mode coupling is seen to become negligible for correlation lengths larger than 4 km.

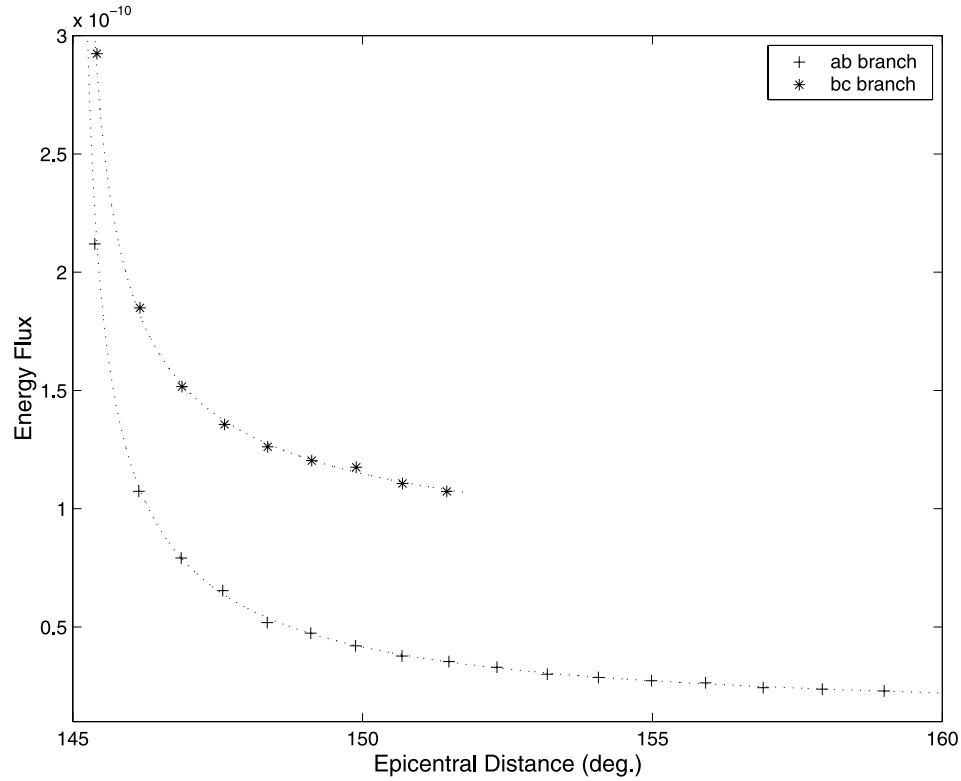

Figure 5. Energy flux of PKP_{ab} and PKP_{bc} as a function of epicentral distance in a perfectly elastic PREM model for an isotropic, point-like source with unit total energy. Symbols, Monte Carlo simulations; dashed line, dynamic ray tracing.

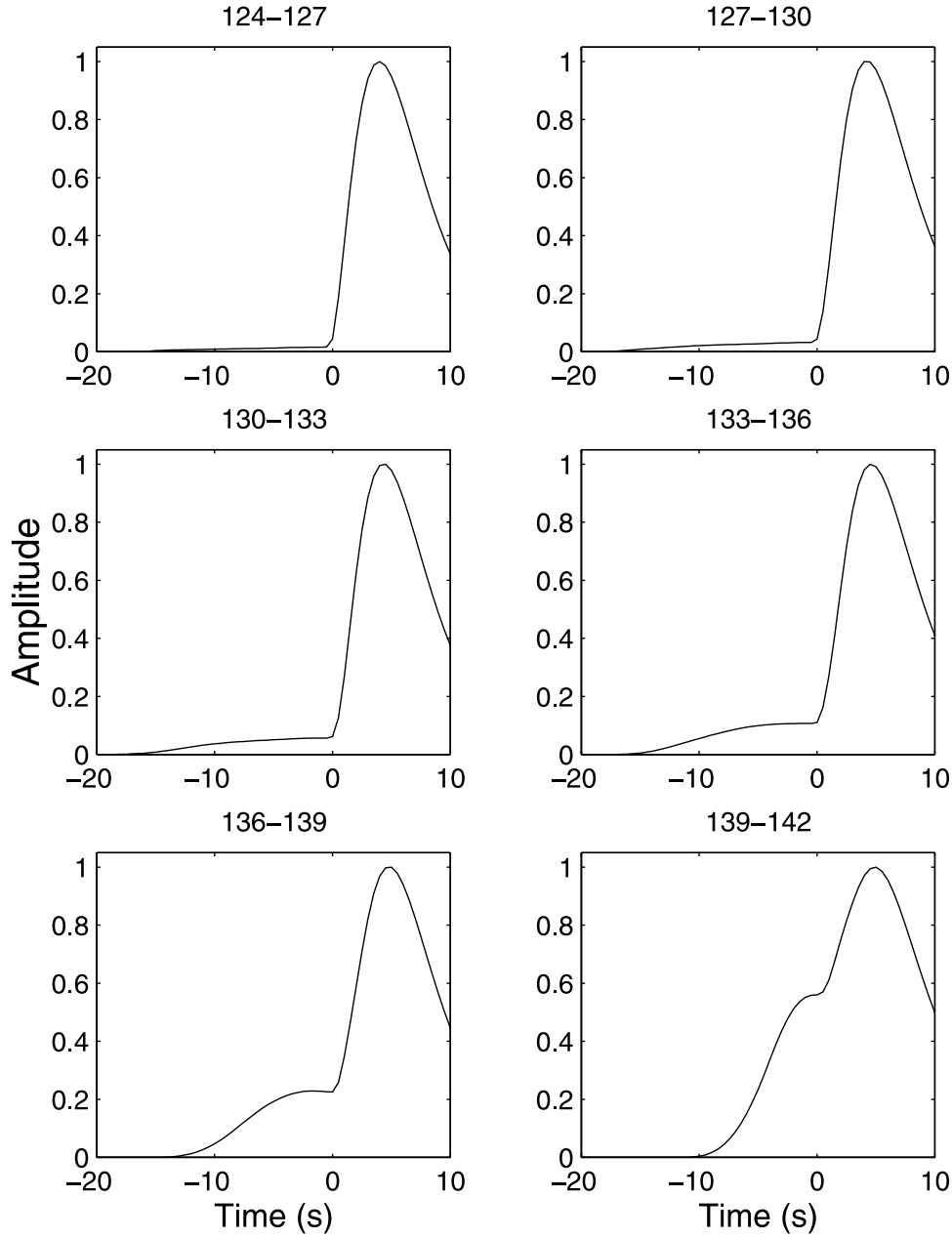

Figure 6. Example of full seismogram envelopes computed near the PKP_{df} arrival at various epicentral distances (indicated on top of each plot). The lower mantle contains heterogeneities with 16 km length scale and 0.6% RMS perturbations.

Instead, we obtain a complete solution that incorporates precursors, direct waves and coda. The normalization of the precursors amplitudes with respect to the peak amplitudes of the core phases is therefore made simple. In order to ensure rapid statistical convergence, we divide the Earth's surface into annuli of 3° epicentral distance. The seismogram envelopes are computed as the RMS energy flux, taking as reference time the arrival time of the *df* branch. This is in agreement with the processing of data. As will be explained in a forthcoming paper, the data are first binned into six epicentral distance ranges, next they are aligned with respect to the *df* phase, then their RMS amplitude are evaluated. An example of full seismogram envelope computed in a heterogeneous mantle with weak (0.6%) small-scale (16 km) perturbations is shown in Figure 6. As often observed, the energy of the precursors is seen to decay very rapidly from the caustic to the shadow zone.

5. Born Versus Radiative Transfer

[41] In the rest of the paper we present a series of tests of Born approximation against transport theory. A general theoretical study of the limitations of the Born approximation can be found in the work of *Hudson and Heritage* [1981]. In order to determine the mean free time and phase function from the length scale and RMS perturbations of inhomogeneities, we assume a wave speed of 13.65 km/s and a central frequency of the waves of 1.3 Hz. Although still actively

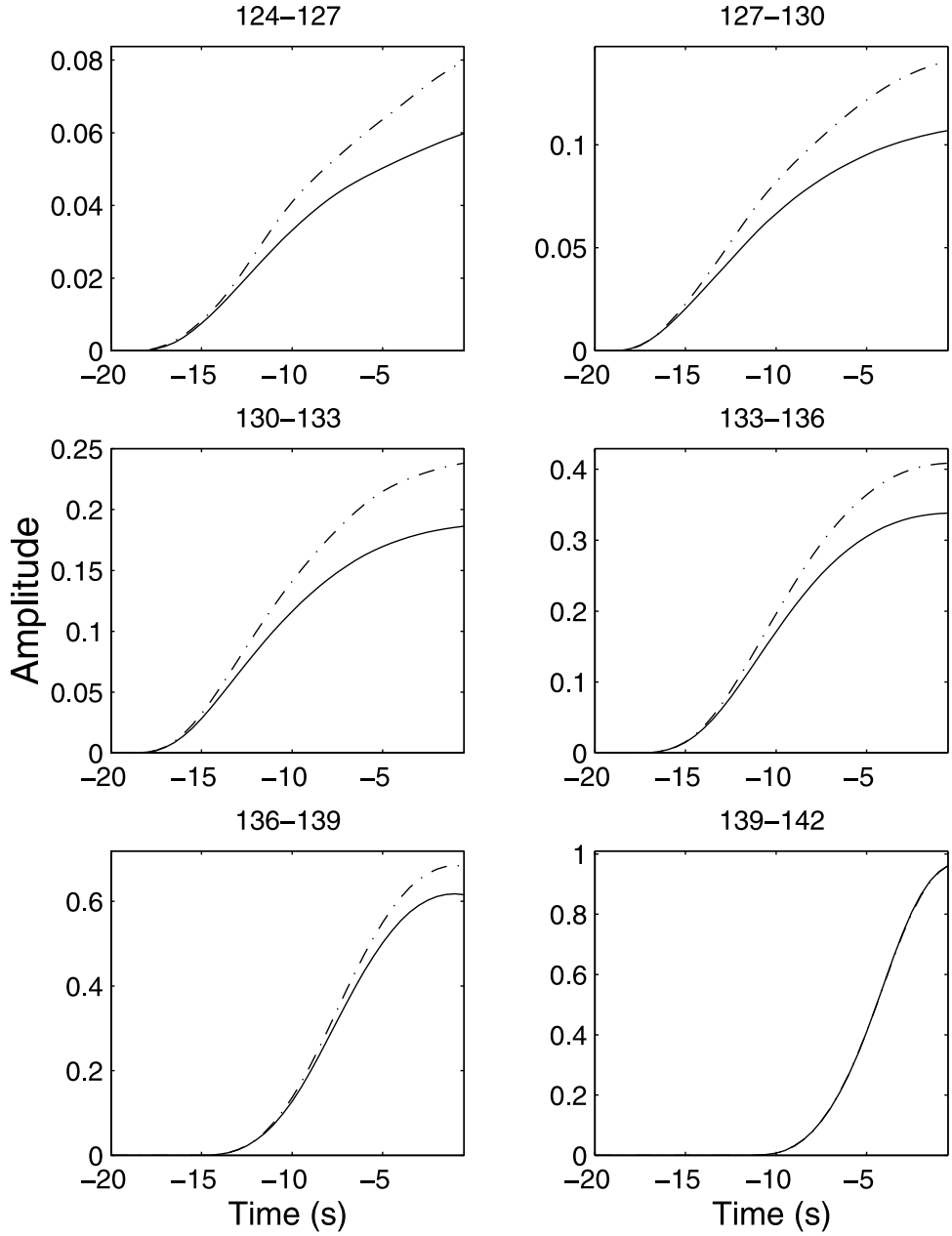

Figure 7. Comparison between Born approximation (solid line) and transport theory (dashed line) as a function of epicentral distance (indicated on top of each plot). The lower mantle has 8 km length scale heterogeneities and a mean free time of 200 s (equivalent to 0.8% RMS perturbations). The source is isotropic with a duration of about 8 s. Amplitudes have been normalized with respect to the peak *cd* + *df* amplitudes.

debated, the quality factor of the inner core has been assigned a PREM-like value of 400. Energy is released at a point-like isotropic source located at the surface of the Earth, with a simple source time function of about 8 s duration.

[42] We first examine the effects of strength of the scattering and heterogeneous layer thickness at fixed length scale. The example shown in Figure 7 corresponds to an Earth model with 8 km length scale heterogeneities distributed throughout the lower mantle, with a mean free time of 200 s (RMS perturbations of about 0.75%), which is similar to the preferred whole mantle scattering model of Hedlin. The precursors mean amplitude is plotted as a function of time before the *df* arrival. The peak *cd* + *df* amplitude has been normalized to 1 at all epicentral distances. Not surprisingly, Born approximation underestimates the precursor amplitudes and it is noticable that the discrepancy with transport theory clearly depends on the epicentral distance. In the vicinity of the *b* caustic Born approximation matches the complete solution extremely well, while as one moves deeper and deeper into the shadow zone its accuracy deteriorates, suggesting the prominence of multiple scattering in this region.

[43] We pursue our investigation on the validity of Born approximation by examining in Figure 8 the effects of

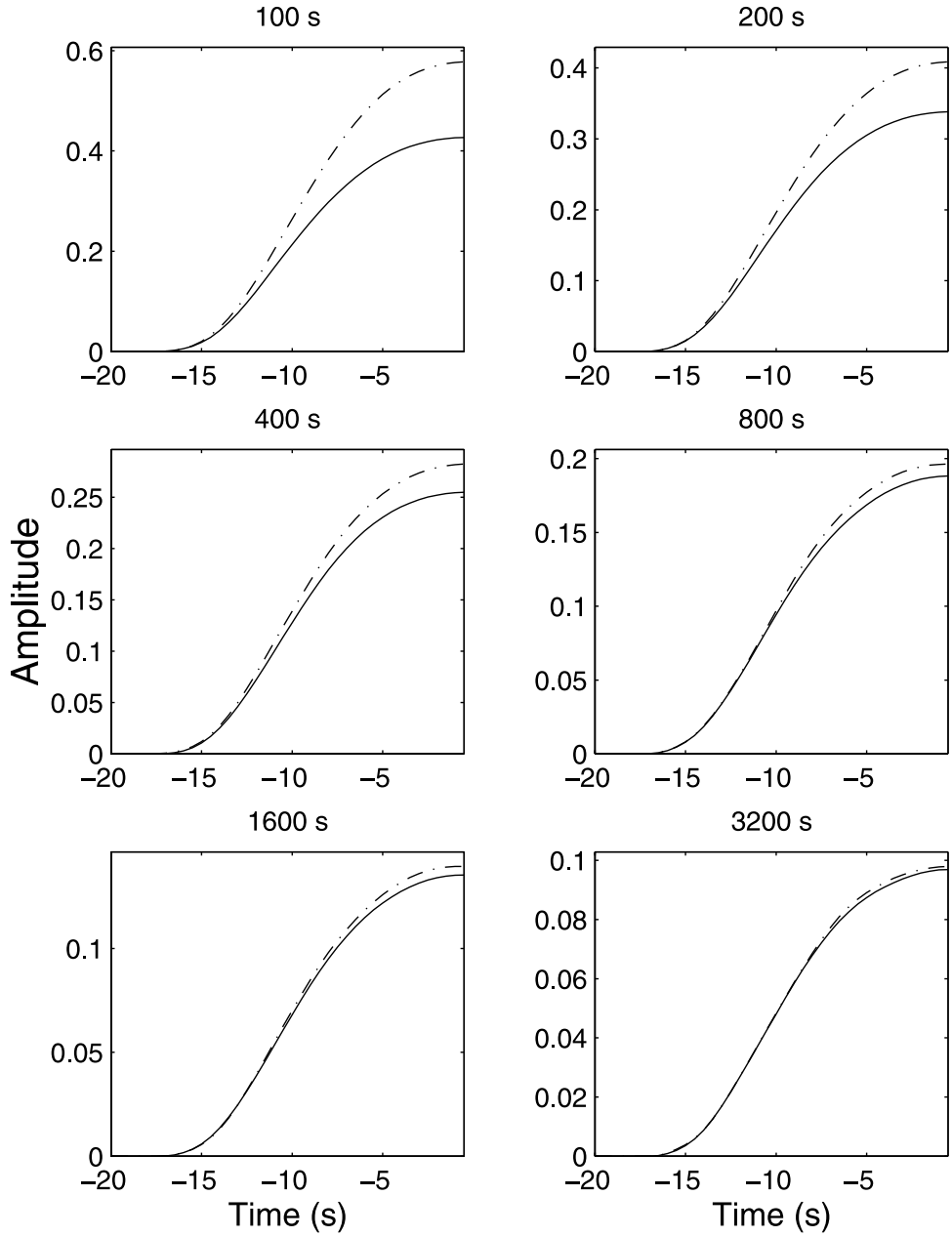

Figure 8. Comparison between Born approximation (solid line) and transport theory (dashed line) in the 133–136 epicentral distance range for a whole mantle scattering model with fixed length scale (8 km) and increasing values of the mean free time (indicated on top of each plot).

scattering strength in a medium with 8 km length scale. Figure 8 serves to illustrate the convergence of single and multiple scattering theories in the limit of large mean free times (small perturbations). When the mean free time is of the order 10 times the transit time of the waves through the lower mantle, the Born approximation closely approximates the full solution. If we translate this "rule of thumb" into perturbations, it implies that in a whole mantle scattering model with 8 km length scale, perturbations should be <0.5% for the Born approximation to be valid.

[44] Obviously, if we shrink the scattering layer, the constraint becomes less severe. This is examined in greater details in Figure 9, where the upper limit of scattering is located successively at 150 km, 500 km, and 2000 km above the core-mantle boundary. The comparison of the results for mean free times of 400 s (weak perturbations) and 100 s (strong perturbations) illustrates the cumulative effects of multiple scattering. We relate this phenomenon to the fact that in a medium with constant mean free time τ, the distribution of the number of scatterings within a time interval T follows a Poisson distribution with parameter $\lambda = T/\tau$. As is well known, the Poisson distribution is characterized by its long tail, which tends to favor high-order scatterings. It is also noticable that the thickness of the layer of scatterers plays a major role in the overall shape of the envelopes. If scattering is restricted to D″ the amplitudes of the precursors start to decay about 5 s before the PKP_{df} comes in, independent of the mean free time.

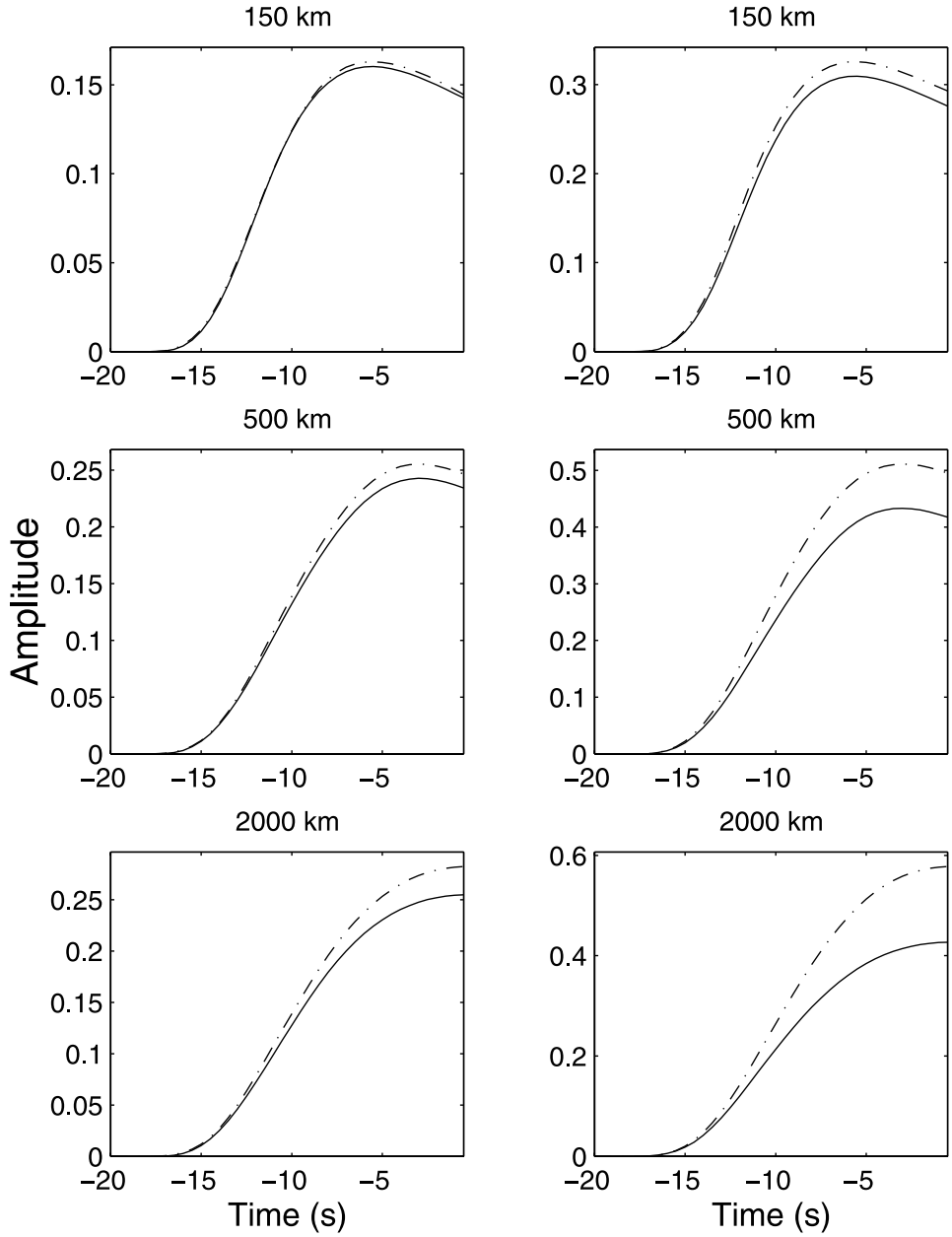

Figure 9. Comparison between Born approximation (solid line) and transport theory (dashed line) in the 133–136 epicentral distance for various thickness of the scattering layer. The mean free time is 400 s on the left and 100 s on the right.

[45] After focusing on the role of thickness and mean free time, we show that the correlation length of the scatterers also strongly influences the amplitude decay of the precursors in the shadow zone. As illustrated in Figure 10, the scattered energy decreases by a factor 2 or more as the length scale increases by a factor 4. The discrepancy between Born approximation and transport theory clearly increases with the length scale, an effect that is amplified by the decrease of the mean free time. Thus multiple scattering strongly enhances the diffusion of energy off of the forward direction. An important implication of this result pertaining to data interpretation follows: because Born approximation overpredicts the amplitude decay of the precursors in the shadow zone, it may lead to a systematic underestimation of the length scales of lower mantle heterogeneities.

[46] In order to quantify in a simple and convenient way the accuracy of Born approximation we introduce the total (time integral) energy of the precursors predicted by Born and transport theories, denoted by E^b and E^t, respectively. In Figure 11, a scatterplot of the ratio $|E^t - E^b|/E^t$ as a function of length scale and mean free time illustrates the limits of accuracy of Born approximation in the 133–136 epicentral distance range. Plots at other epicentral distance range exhibit essentially the same feature. A noticable exception is the 139–142 epicentral distance range where E^b differs from E^t by no more than 10%. From Figure 11, we infer that

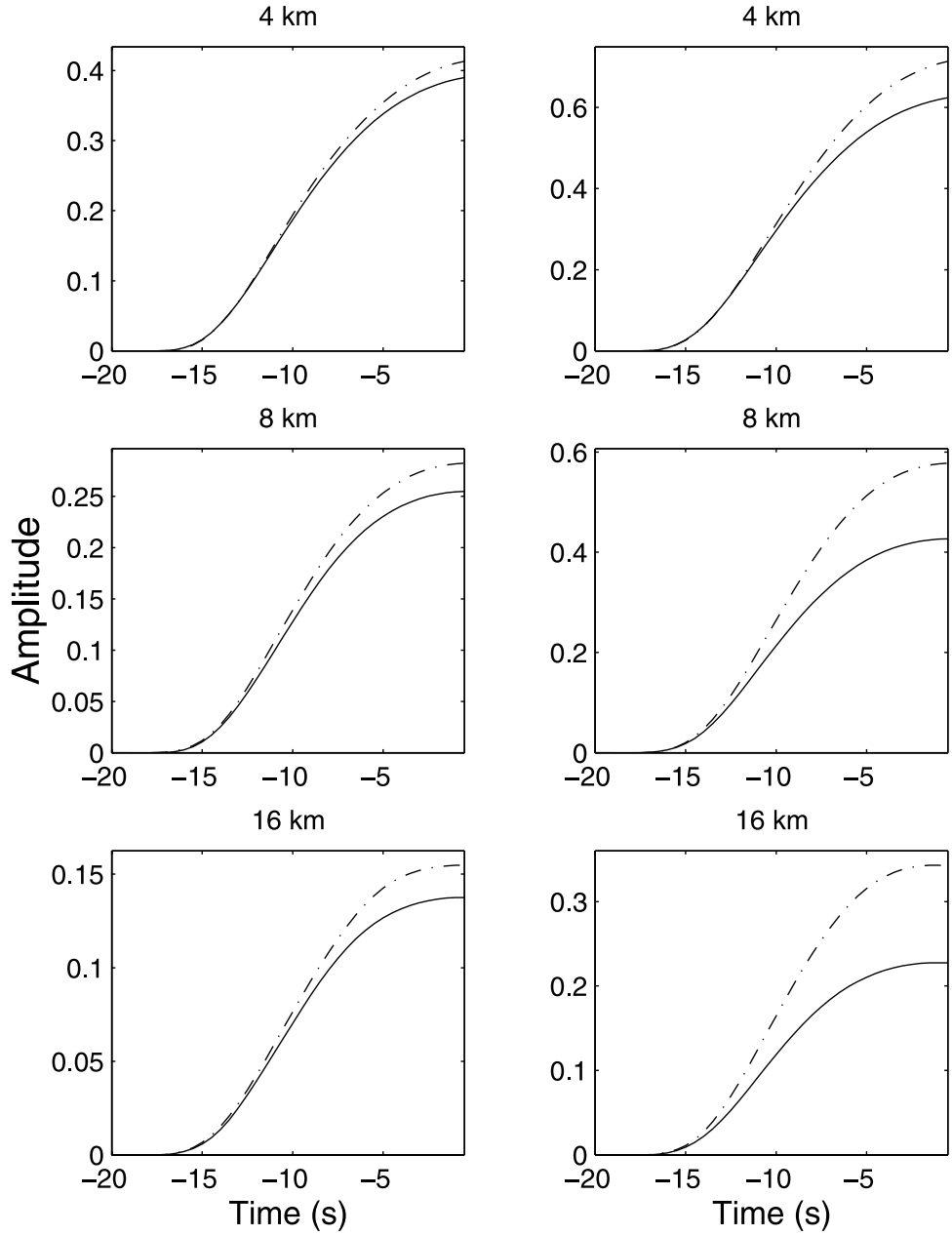

Figure 10. Comparison between Born approximation (solid line) and transport theory (dashed line) in the 133–136 epicentral distance range in whole mantle scattering models with increasing values of the correlation length (indicated on top of each plot). The mean free time is 400 s on the left and 100 s on the right.

the limit of applicability of Born approximation lies between 400 and 800 s mean free times, relatively independent of the length scale. We recall that for length scales of 8 km, this corresponds to about 0.3–0.5% perturbations, which is exceeded by the scatterers proposed by *Hedlin et al.* [1997] to explain precursors observations. A reevaluation of their results in the context of transport theory appears therefore as an important topic for future investigations.

6. Conclusion

[47] Transport theory emerges as a potentially useful tool to interpret average amplitudes of high-frequency global seismic data. Its primary domain of application is the inversion of small-scale statistical fluctuations in the deep Earth. Because the transport equation incorporates in a rigorous way the effects of multiple scattering, it is restricted neither to weak fluctuations nor to scattering media with limited spatial extension. We have specifically developed a Monte Carlo code that enables the forward modeling of high-frequency *PKP* precursors. Our numerical investigations show that the often used Born approximation suffers from severe limitations and may model inaccurately the seismograms envelopes. In a companion paper, we will apply our method to the inversion of a large data set of *PKP* precursors and show the feasibility of a

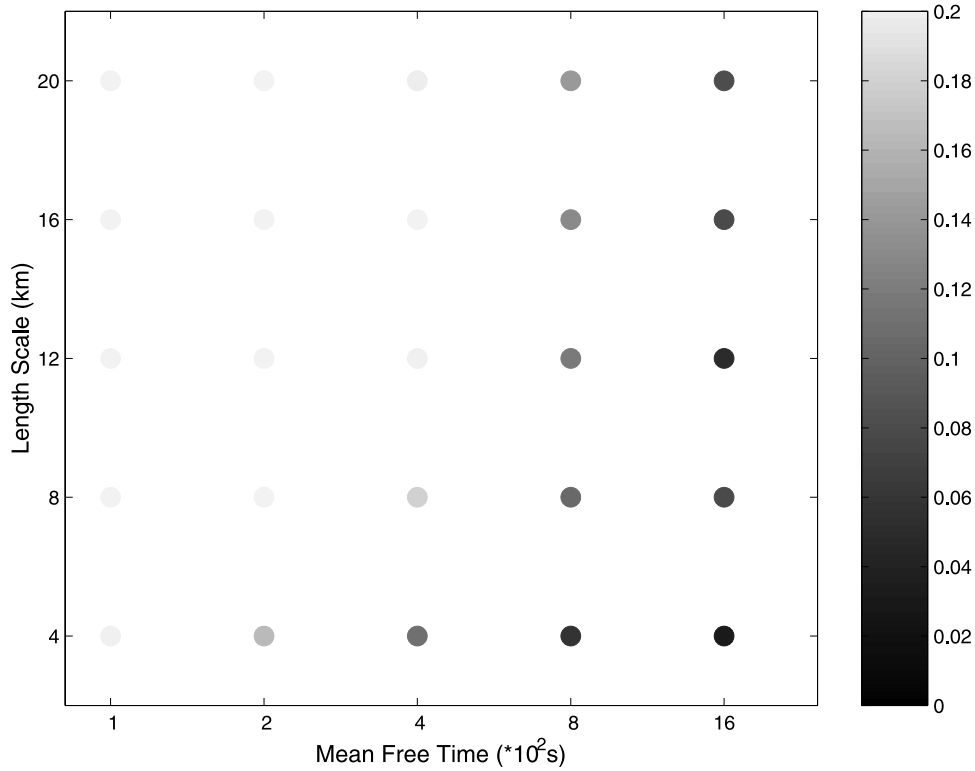

Figure 11. Scatterplot of the ratio $|E^b - E^t|/E^t$ as a function of the mean free time in seconds (horizontal axis) and length scale in km (vertical axis) in the 133–136 epicentral distance range for whole mantle scattering. E^b (respectively E^t) denotes the integrated energy of the precursors predicted by Born approximation (respectively transport theory). When the mean free time is of the order 400 s or less, the difference between the predictions of the two theories exceeds 20%. This provides an overall view of the validity of the Born approximation in the length scale-mean free time parameter space. See color version of this figure at back of this issue.

complete inversion of average high-frequency seismic amplitudes.

Appendix A: Ray Equations

[48] As is customary [*Landau and Lifchitz*, 1982], we can define the frequency ω as the Hamiltonian H for rays:

$$H = \omega(t, \mathbf{R}, \mathbf{k}), \tag{A1}$$

and the ray equations in space-time therefore take the canonical from

$$\frac{d\mathbf{R}}{dt} = \frac{\partial \omega}{\partial \mathbf{k}}, \tag{A2}$$

$$\frac{d\mathbf{k}}{dt} = -\frac{\partial \omega}{\partial \mathbf{r}}, \tag{A3}$$

$$\frac{d\omega}{dt} = \frac{\partial \omega}{\partial t}. \tag{A4}$$

In what follows, the medium will be deemed stationary and therefore we can drop the time dependence from ω, which implies that the frequency is conserved along a space-time ray. $\mathbf{R}$ and $\mathbf{k}$ are the conjugate variables defining the phase space for the ray analysis and can be viewed as functions of their initial values $\mathbf{R} = \mathbf{R}^0$, $\mathbf{k} = \mathbf{k}^0$ at time $t = t^0$, so that

$$\mathbf{R} = \mathbf{R}(\mathbf{R}^0, \mathbf{k}^0, t) \quad \text{and} \quad \mathbf{k} = \mathbf{k}(\mathbf{R}^0, \mathbf{k}^0, t). \tag{A5}$$

We now introduce the new variable $\mathbf{k}$ in lieu of ω and $\mathbf{n}$ in equation (5) which now reads

$$\frac{dE}{dt} = \frac{d}{dt} \int_{\Delta \mathbf{k}(t)} \int_{\Delta \mathbf{R}(t)} \frac{I(\omega, t, \mathbf{R}, \mathbf{n})}{v_g} J_1 d^3\mathbf{k} d^3\mathbf{R}, \tag{A6}$$

where J_1 is the Jacobian of the coordinate change, which can be written symbolically as

$$J_1 = \left| \frac{d\Omega(\mathbf{n}) d\omega}{d^3\mathbf{k}} \right|. \tag{A7}$$

ω and $\mathbf{n}$ must now be viewed as functions of $\mathbf{k}$. The next step to compute the total derivative is to introduce as new coordinates ($\mathbf{k}^0$, $\mathbf{R}^0$). The method is in complete analogy with a change from Eulerian to Lagrangian coordinates in continuum mechanics. Introducing the Jacobian of the transformation in compact notations as

$$J_2 = \left| \frac{d^3\mathbf{R} d^3\mathbf{k}}{d^3\mathbf{R}^0 d^3\mathbf{k}^0} \right|, \tag{A8}$$

we obtain

$$\frac{dE}{dt} = \int\limits_{\Delta \mathbf{k}^0} \int\limits_{\Delta \mathbf{R}^0} \frac{d}{dt}\left(\frac{I(\omega, t, \mathbf{R}, \mathbf{n}) J_1 J_2}{v_g}\right) d^3\mathbf{R}^0 d^3\mathbf{k}^0. \tag{A9}$$

After straightforward manipulations, and reverting to the original variables, we find

$$\int\limits_{\Delta\Omega} d\Omega(\mathbf{n}) \int\limits_{\Delta r} d^3\mathbf{R} \int\limits_{\Delta\omega} d\omega \frac{1}{v_g}\left(\frac{d}{dt} + \frac{d}{dt}\log\left|\frac{J_1 J_2}{v_g}\right|\right) I(\omega, t, \mathbf{R}, \mathbf{n}) = 0. \tag{A10}$$

Because of the arbitrary choice of the beam extension in phase space one concludes

$$\frac{dI}{dt} + I\frac{d}{dt}\log\left|\frac{J_1 J_2}{v_g}\right| = 0. \tag{A11}$$

From Liouville theorem we know that a volume of phase space is invariant and thus $J_2 = 1$. In an isotropic medium, the Jacobian J_1 can be expressed as

$$J_1 = \left|\frac{\sin\theta d\theta d\phi d\omega}{\sin\theta d\theta d\phi k^2 dk}\right|, \tag{A12}$$

or

$$J_1 = \frac{v_g \omega^2}{v_0^2}. \tag{A13}$$

In a stationary medium the frequency along the space-time ray is conserved: $d\omega/dt = 0$; hence, equation (6) is established.

Appendix B: Scattering Parameters in Elastic Media

[49] For completeness, in this appendix we provide closed formulas from which all the scattering parameters of the transport equation can be calculated. We focus on the case of exponential correlations in media with velocity-type perturbations as defined by *Wu and Aki* [1985]. For convenience, let us introduce the differential scattering cross sections [*Sato and Fehler*, 1998], defined as the ratio of the flux of energy scattered into direction $\mathbf{n}$ per unit volume and unit solid angle, to the flux of energy carried by the incident wave propagating in direction $\mathbf{n}'$. Introducing the cosine of the scattering angle $\mu = \mathbf{n} \cdot \mathbf{n}'$, one obtains [*Wu and Aki*, 1985]

$$\frac{d\sigma^{pp}}{d\Omega}(\mathbf{n}, \mathbf{n}') = \frac{(k^p a)^4}{2\pi a}\langle\epsilon^2\rangle \frac{\left(\frac{1}{3} + \mu + \frac{2}{3}\mu^2\right)^2}{\left[1 + 2(k^p a)^2(1-\mu)\right]^2}, \tag{B1}$$

for the P to P differential scattering cross section and

$$\frac{d\sigma^{ps}}{d\Omega}(\mathbf{n}, \mathbf{n}') = \frac{(k^s a)^4 \beta}{2\pi a \alpha}\langle\epsilon^2\rangle \frac{\left(\sqrt{1-\mu^2} + \frac{2\beta}{\alpha}\mu\sqrt{1-\mu^2}\right)^2}{\left[1 + (k^s a)^2\left(1 + \frac{\beta^2}{\alpha^2} - 2\frac{\beta}{\alpha}\mu\right)\right]^2}, \tag{B2}$$

for the P to S differential scattering cross section. In equations (B1)–(B2), the following notations have been introduced: α, local P wave speed; β, local S wave speed; k^p, local P wave number; k^s, local S wave number; $\langle\ \epsilon^2\ \rangle$, mean square perturbations; $d\Omega$, elementary solid angle;a, correlation length of the fluctuations. Upon integration of relations (B1)–(B2) over the whole solid angle, one obtains the total scattering cross sections (per unit volume) Σ^{pp} – Σ^{ps} for P to P and P to S scattering, respectively. The corresponding P to P and P to S scattering mean free times are defined as

$$\tau^{pp} = \frac{1}{\alpha\Sigma^{pp}}, \tag{B3}$$

$$\tau^{ps} = \frac{1}{\alpha\Sigma^{ps}}. \tag{B4}$$

Exact expressions for τ^{pp} and τ^{ps} can be derived using the results of *Wu and Aki* [1985]. However, the final formulas are so complicated that we rather provide the reader with asymptotic expressions, which capture the essential behavior of the scattering mean free times. The following formulas are obtained at low frequencies ($k^p a$, $k^s a \ll 1$):

$$\tau^{pp} \approx \frac{135a}{184\langle\epsilon^2\rangle\alpha(k^p a)^4}, \tag{B5}$$

$$\tau^{ps} \approx \frac{45a}{19\sqrt{3}\langle\epsilon^2\rangle\alpha(k^s a)^4}, \tag{B6}$$

while at high frequencies ($k^p a$, $k^s a \gg 1$), we find

$$\tau^{pp} \approx \frac{a}{2\langle\epsilon^2\rangle\alpha(k^p a)^2}, \tag{B7}$$

$$\tau^{ps} \approx \frac{2a}{\sqrt{3}\nu\langle\epsilon^2\rangle\alpha}, \tag{B8}$$

where ν is an irrational constant close to 2.70194. Note that our asymptotic results for P to S mode conversions assume the Earth to be approximately a Poisson solid. In the low frequency regime, one find the classical ω^{-4} frequency dependence for both τ^{pp} and τ^{ps}. At high frequencies, τ^{pp} decreases as ω^{-2}, while τ^{ps} tends to a constant.

[50] **Acknowledgments.** The authors would like to thank F. A. Dahlen for many stimulating discussions and encouragements in the course of this research. Discussions with S.-H. Hung, M. Hedlin, P. Earle, M. Campillo, and N. M. Shapiro were also greatly appreciated. This research was partially supported by NSF (EAR-0003348).

References

Aki, K., Scattering of *P* waves under the Montana Lasa, *J. Geophys. Res.*, *78*, 1334–1346, 1973.

Apresyan, L. A., and Y. A. Kravtsov, *Radiation Transfer: Statistical and Wave Aspects*, Gordon and Breach, Newark, N. J., 1996.

Bal, G., G. Papanicolaou, and L. Ryzhik, Probabilistic theory of transport processes with polarization, *SIAM J. Appl. Math.*, *60*, 1639–1666, 2000.

Ben-Menahem, A., and S. J. Singh, *Seismic Waves and Sources*, Dover, Mineola, N. Y., 2001.

Chaljub, E., Modélisation numérique de la propagation d'ondes sismiques à l'échelle du globe, Ph.D. thesis, Univ. Paris 7, Paris, 2000.

Cormier, V. F., Time-domain modelling of *PKIKP* precursors for constraints on the heterogeneity in the lowermost mantle, *Geophys. J. Int.*, *121*, 725–736, 1995.

Cormier, V. F., Anisotropy of heterogeneity scale lengths in the lower mantle from *PKIKP* precursors, *Geophys. J. Int.*, *136*, 373–384, 1999.

Dahlen, F. A., and J. Tromp, *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, N. J., 1998.

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

Dziewonski, A., and A. M. Anderson, Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*, 297–356, 1981.

Flatté, S. M., and R. S. Wu, Small-scale structure un the lithosphere and asthenosphere deduced from arrival time and amplitude fluctuations at NORSAR, *J. Geophys. Res.*, *93*, 6601–6614, 1988.

Gusev, A. A., and I. R. Abubakirov, Monte-Carlo simulation of record envelope of a near earthquake, *Phys. Earth Planet. Inter.*, *49*, 30–36, 1987.

Hedlin, M. A. H., and P. Shearer, An analysis of large-scale variations in small-scale mantle heterogeneity using Global Seismographic Network recordings of precursors to *PKP*, *J. Geophys. Res.*, *105*, 13,655–13,673, 2000.

Hedlin, M. A. H., P. Shearer, and P. S. Earle, Seismic evidence for small-scale heterogeneity throughout the Earth's mantle, *Nature*, *387*, 145–150, 1997.

Hennino, R., N. Tregourès, N. M. Shapiro, L. Margerin, M. Campillo, B. A. van Tiggelen, and R. L. Weaver, Observation of equipartition of seismic waves, *Phys. Rev. lett.*, *86*, 3447–3450, 2001.

Hoshiba, M., Simulation of coda wave envelope in depth dependent scattering and absorption structure, *Geophys. Res. Lett.*, *21*, 2853–2856, 1994.

Hoshiba, M., Seismic coda wave envelope in depth dependent *S*-wave velocity structure, *Phys. Earth planet. Inter.*, *21*, 15–22, 1997.

Hudson, J. A., and J. R. Heritage, The use of the Born approximation in seismic scattering problems, *Geophys. J. R. Astron. Soc.*, *66*, 221–240, 1981.

Hung, S.-H., F. A. Dahlen, and G. Nolet, Frechet kernels for finite-frequency traveltimes: II. Examples, *Geophys. J. Int.*, *141*, 175–203, 2000.

Ishimaru, A., *Wave Propagation and Scattering in Random Media*, vols. I–II, Academic, San Diego, Calif., 1978.

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

Lagendijk, A., and B. A. Van Tiggelen, Resonant multiple scattering of light, *Phys. Rep.*, *270*, 143–215, 1996.

Landau, L., and E. Lifchitz, *Théorie des Champs*, Mir, Moscow, 1982.

Lapeyre, B., E. Pardoux, and R. Sentis, *Méthodes de Monte-Carlo pour les Équations de Transport et de Diffusion*, Springer-Verlag, New York, 1998.

Li, X.-D., and T. Tanimoto, Waveforms of long period body waves in a slightly aspherical Earth model, *Geophys. J. Int.*, *112*, 92–102, 1993.

Lux, I., and L. Koblinger, *Monte Carlo Particle Transport Methods: Neutron and Photon Calculations*, CRC Press, Boca Raton, Fla., 1991.

Margerin, L., M. Campillo, and B. A. van Tiggelen, Radiative transfer and diffusion of waves in a layered medium: New insight into coda, *Q. Geophys. J. Int.*, *134*, 596–612, 1998.

Margerin, L., M. Campillo, and B. A. van Tiggelen, Monte Carlo simulation of multiple scattering of elastic waves, *J. Geophys. Res.*, *105*, 7873–7892, 2000.

Nolet, G., Linearized inversion of (teleseismic) data, in *The Solution of the Inverse Problem in Geophysical Interpretation*, edited by R. Cassinis, pp. 9–31, Plenum, New York, 1981.

Nolet, G., and F. A. Dahlen, Wave front healing and the evolution of seismic delay times, *J. Geophys. Res.*, *105*, 19,043–19,054, 2000.

Rytov, S. M., Y. A. Kravtsov, and V. I. Tatarskii, *Principles of Statistical Radiophysics*, vol. 4, *Wave Propagation Through Random Media*, Springer-Verlag, New York, 1989.

Ryzhik, L. V., G. C. Papanicolaou, and J. B. Keller, Transport equations for elastic and other waves in random media, *Wave Motion*, *24*, 327–370, 1996.

Sato, H., Multiple isotropic scattering model including *P*-*S* conversions for the seismogram envelope formation, *Geophys. J. Int.*, *117*, 487–494, 1994.

Sato, H., and M. Fehler, *Wave Propagation and Scattering in the Heterogeneous Earth*, Springer-Verlag, New York, 1998.

Sheng, P., *Introduction to Wave Scattering Localization and Mesoscopic Phenomena*, Academic, San Diego, Calif., 1995.

Snieder, R., and C. Chapman, The reciprocity properties of geometrical spreading, *Geophys. J. Int.*, *132*, 89–95, 1998.

Weaver, R. L., Diffusivity of ultrasound in polycrystals, *J. Mech. Phys. Solids*, *38*, 55–86, 1990.

Woodward, M. J., Wave-equation tomography, *Geophysics*, *57*, 15–26, 1993.

Wu, R. S., Multiple scattering and energy transfer of seismic waves-separation of scattering effect from intrinsic attenuation: I. Theoretical modeling, *Geophys. J. R. Astron. Soc.*, *82*, 57–80, 1985.

Wu, R. S., and K. Aki, Elastic wave scattering by a random medium and the small-scale inhomogeneities in the lithosphere, *J. Geophys. Res.*, *90*, 10,261–10,273, 1985.

Wu, R. S., and K. Aki, Multiple scattering and energy transfer of seismic waves, separation of scattering effect from intrinsic attenuation. II. Application of the theory to Hindu Kush region, *Pure Appl. Geophys.*, *128*, 49–80, 1988.

Yoshimoto, K., Monte-Carlo simulation of seismogram envelope in scattering media, *J. Geophys. Res.*, *105*, 6153–6161, 2000.

Zeng, Y., Theory of scattered *P*-wave and *S*-wave energy in a random isotropic scattering medium, *Bull. Seismol. Soc. Am.*, *83*, 1264–1275, 1993.

Zhao, L., and T. H. Jordan, Sensitivity of frequency-dependent traveltimes to laterally heterogeneous, isotropic Earth structure, *Geophys J. Int.*, *133*, 683–704, 1998.

L. Margerin, Laboratoire de Géophysique Interne, Observatoire de Grenoble, Université Joseph Fourier, BP 53X, 38041 Grenoble Cedex, France. (Ludovic.Margerin@lgit.obs.ujf-grenoble.fr)

G. Nolet, Department of Geosciences, Princeton University, Guyot Hall, Princeton, NJ 08544, USA.

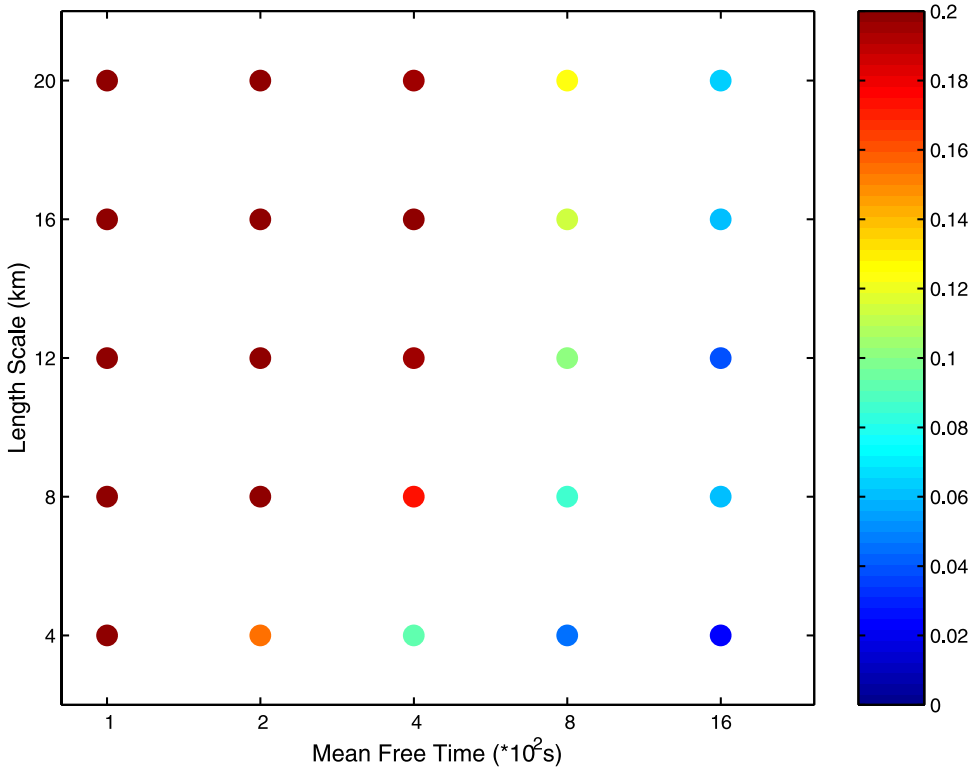

Figure 11. Scatterplot of the ratio $|E^b - E^t|/E^t$ as a function of the mean free time in seconds (horizontal axis) and length scale in km (vertical axis) in the 133–136 epicentral distance range for whole mantle scattering. E^b (respectively E^t) denotes the integrated energy of the precursors predicted by Born approximation (respectively transport theory). When the mean free time is of the order 400 s or less, the difference between the predictions of the two theories exceeds 20%. This provides an overall view of the validity of the Born approximation in the length scale-mean free time parameter space.