Statistically stable velocity macro-model estimation

Eric Dussaud

ABSTRACT

Velocity analysis resolves relatively long scales of earth structure, on the order of 1 km. Migration produces images with length scales (wavelengths) on the order of 10's of m. In between these two scale regimes lies another, corresponding roughly to structures between 60 to 300m in extent, in which the resolution of velocity analysis is uncertain and the energy of images is small to non-existent. The proposed thesis aims at assessing the impact on velocity analysis and imaging of uncertainty at these intermediate length scales, using ideas on time reversal and imaging in randomly inhomogeneous media developed by G. Papanicolaou and colleagues, in combination with velocity estimation methods of differential semblance type.

INTRODUCTION

Motivation

The Earth's crust is extremely heterogeneous, with structures occurring over a broad range of scales, from the micrometer-millimeter grain and pore scale, to the multi-kilometer sedimentary basin scale (where hydrocarbons are typically found). These heterogeneities include variations in lithology, porosity, permeability, pore fluid properties, and conditions of pore pressure, temperature, and stress, among others.

Exploration seismology aims at determining the structure of the near-subsurface (usually 0-10km) of the Earth from observations of reflected seismic waves, generally for the purpose of locating hydrocarbons. Because the mechanical properties of the heterogeneous near-surface span a wide range of scales, seismic waves propagate in a variety of complex nonlinear ways. In order to understand seismic time responses in the earth, the effects of such complexity on seismic wave fields need to be resolved. To model the propagation of waves, the subsurface is generally viewed as an acoustic medium. Although this model does not accommodate many of the phenomena associated with seismic wave propagation, it is the simplest one in which to carry out the analysis and is widely used in the seismic processing industry. Such a medium is generally described by the following two parameters: the acoustic wave

speed and the mass density. These parameters occur as coefficients in the wave equation which describes wave propagation in an acoustic medium. It was established long ago that reflected signals are caused by localized, rapid changes in rock properties. Inspection of direct measurements (well logs) often shows that these reflection zones exhibit oscillatory or abrupt changes in mechanical properties. Because the density plays a relatively minor role and changes less dramatically than the velocity in the upper crust of the Earth, we may view it as a constant and disregard it. Accordingly, the significant properties of an arbitrary Earth model are a smooth velocity macro-model plus a rough reflectivity micro-model. For this reason, seismic data are commonly modeled by a high-frequency single scattering approximation [6], the so-called 'ray+Born' approximation. This amounts to a linearization of the wave equation in the medium coefficients about a smooth background. That is, the medium coefficient is written as a sum of a smooth background component and an oscillatory perturbation (the reflectivity), assumed to be small enough that linearization is accurate. This *scale dichotomy* between long scale and short scale components reflects the common conceptual division: velocity is responsible for kinematics, while reflectivity is responsible for dynamics. The separation of scales also allows for great simplification in the description of solutions of the solutions of the wave equation: in particular, waves propagate according to geometrical optics in the background medium, and (partially) reflect at the singularities of the reflectivity.

Both the smooth background and the perturbation are in general unknown and have to be reconstructed jointly. Seismic imaging techniques typically resolve the long-scale component of velocity via a process called *velocity analysis*, whereas *migration* (or *linearized inversion*) methods commonly resolve the reflectivity. However, the standard techniques do **not** estimate the intermediate scale structure. In fact, there is a widespread belief that seismic data simply do not contain any reliable information on this intermediate scale structure. Indeed, Jannine, Tarantola et al. [28] showed that while seismic data contain information on both long wavelengths ($\lambda \geq 300$ m) and short wavelengths of velocity ($\lambda \leq 60$ m), they do not contain information on the middle wavelengths ($60 \text{ m} \leq \lambda \leq 300$ m). This is the well-known *scale gap*. This is corroborated by Claerbout [19] who described the "information" gap by the following plot (Imaging the Earth's interior, page 47):

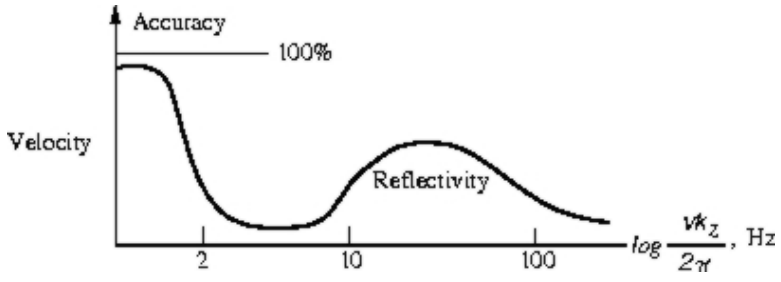

FIG. 1. Reliability of information obtained from surface seismic measurements

The proposed work aims at providing a new way to look at this familiar "fact". Although the scale gap indicates that seismic data provide little intrinsic information about the middle scales of velocity, it is not clear that this intermediate structure scale does not influence the resolution of the long (background velocity) and short (image) scales. In fact, because the seismic problem is nonlinear (these are components of the velocity), one would expect "energy" or (lack of) "information" to cascade between scales. That is, it seems very reasonable to think that the uncertainty at one scale would pollute the estimation of one (velocity analysis) and the quality of the other (migration). The choice of a random field to model the intermediate scale structure follows naturally from the above discussion: what the random process represents is precisely the uncertainty and the lack of information inherent to these medium scales. The equation of linear acoustics wave propagation becomes therefore a stochastic partial differential equation, and the underlying theory becomes that of wave propagation in stochastic media.

Three-scale asymptotics

In order to study the effect of the random fluctuations of the medium scale velocity component, we will consider a model of the Earth with three distinct scales of velocity:

- "Deterministic" reflectors (the zones of rapid velocity changes) are structures on wavelength scale λ, corresponding to the short-scale component of velocity.

- the background velocity "macro-model" (the component which may be estimated via velocity analysis) varies on the scale L, which will also be the typical propagation distance.

- The medium scale velocity will be regarded as a randomly fluctuating field on the scale l.

This setting corresponds to the *high-frequency regime* in which $\lambda \ll l \ll L$. Under this asymptotic assumption, waves propagate over many correlation lengths so *multiple scattering* is significant. Furthermore, random fluctuations are slowly varying on the wavelength scale, so that scattered waves concentrate within a narrow cone and propagate essentially in the same direction as the primary wave. Therefore multiple scattering occurs essentially in the form of *multipathing* (i.e. forward scattering). In particular, the wave fields can be approximated within the framework of geometrical optics, which accounts for multipathing. However, this approximation method completely ignores diffraction effects which become important when the distance L covered by the wave field in the heterogeneous medium is sufficiently large ($L \gg l/\lambda$). We will use the *parabolic equation approximation* [4; 3; 54], an extension of the geometrical optics approximation which takes diffraction effects into account.

Wave propagation in randomly inhomogeneous media

The phenomena associated with high-frequency wave propagation in randomly inhomogeneous media and the methods of their analysis are considerably varied. Numerous theoretical, computational, and experimental studies have dealt with wave propagation in stochastic media, both in 1-D layered media [33] and in 2-D and 3-D random media, e.g. [18; 57; 32; 27; 42; 40; 58]. More recently, the scattering phenomena associated with stochastic wave propagation have been put to contribution in the context of time-reversed acoustics. In time reversal experiments, a signal emitted by a localized source is recorded by an array of transducers, and then re-emitted into the medium time-reversed, that is, the tail of the signal is sent first. In the absence of absorption, the re-emitted signal propagates back toward the source and focuses approximately on it. In a homogeneous medium, the size of the refocused spot is approximately $\lambda L/a$, where L is the distance of propagation, λ is the wavelength, and a is the size of the transducer array. There are two striking features associated with time reversal in randomly inhomogeneous media. The first one is the *super-resolution* phenomenon which arises as a by-product of multipathing: the transducer array captures waves that were initially moving away from it but get scattered onto it by the inhomogeneities. As the result, the array captures a larger part of the total wave field emanating from the point source and thus appears to be larger than its physical size. The refocused spot is now $\lambda L/a_e$, where $a_e > a$ is the *effective* size of the array. The second feature is that the time-reversed field is *self-averaging*, so that the refocusing of the signal is in essence *statistically stable*, i.e. it does **not** depend on the particular realization of the random medium. This process has had numerous applications such as in ultrasound medical imaging, foliage and ground penetrating radar, mine detection in the ocean, non-destructive testing (e.g. identifying defects in materials), and underwater acoustics (e.g. wireless and secure communications), among others. It has been extensively studied, both experimentally, numerically and theoretically [20; 38; 37; 36; 24; 29; 44; 21; 22; 41; 2; 9; 35; 34]. More recently, time reversal has been the subject of active mathematical research in the context of imaging in randomly inhomogeneous media [5; 13; 12; 10]. In particular, these papers show that it is possible to construct self-averaging functionals of the wave field which can then be used not only to "image" point scatterers but also to quantify the influence of the randomly inhomogeneous medium. Although these ideas have been developed with the seismic imaging as a potential application, a thorough examination of the applicability of these techniques to the seismic inverse problem has yet to be made.

Applications to seismic imaging

The proposed work aims precisely at doing so, in the following contexts. First, we will study the influence of the random fluctuations on the estimation of the background velocity. Estimation of a sufficiently accurate background velocity model in areas where the geology is complex is one of the core challenges in seismic imaging. Since the data can be predicted from the velocity by means of the linear acoustics model, the determination of the latter

can be formulated as an inverse problem, i.e. the background velocity can be adjusted so that the model predicts the observed data as well as possible. A popular formulation of this inverse problem is the output least-squares, in which case the data are to be fitted in the $\mathcal{L}^2$ sense. Differential semblance optimization [14; 49; 50; 43; 25; 17; 31; 46; 47] is one of many formulations of velocity analysis as an optimization problem. In the form presented in [51; 52; 47], the differential semblance objective is the only quadratic functional with smooth (C^∞) dependence on the background medium [47]. This result is supported by several numerical examples [16]. Moreover, for laterally homogeneous medium (i.e. the velocity is a function of depth only) and with the travel times replaced by their hyperbolic moveout approximations, the stationary points of the differential semblance functional are actually asymptotic global minima [52]. The same result was achieved in [46], with a laterally homogeneous medium and a single reflector (it can be extended to accommodate many reflectors) but with the correct (ray-theory) travel times, provided the medium is non constant above the reflector. In this thesis, we propose to construct statistically stable functionals of the differential semblance type, i.e. functionals which essentially vanish at the correct background velocity. The second part of the proposed work will focus on the influence of the cross-scale interactions on the quality of migration. Migration is the central step in seismic processing because it actually produces detailed map ("image") of the structure of the near-subsurface of the Earth. As it was demonstrated in [30; 6; 48; 39; 55], migration is a technique of imaging discontinuities of parameters describing the medium and its accuracy relies on a correct background velocity model. The proposed work will show that the effect of the random fluctuations in the medium scales of velocity can be explicitly quantified, using the techniques discussed previously.

Contents of this proposal

This proposal is organized as follows. Sections 1 and 2 are of an introductory nature. Section 1 gives an overview of the seismic inverse problem, with an emphasis on seismic migration and velocity analysis using differential semblance optimization. Section 2 contains an overview of the properties of time-reversed acoustic fields, including a detailed mathematical analysis of the phenomena of super-resolution and self-averaging of certain functionals of the wave fields. Section 3 represents the core of this thesis proposal. We show there how the tools developed by G. Papanicolaou and his group might be used in the context of seismic imaging.

THE SEISMIC INVERSE PROBLEM

In this section, we give a brief overview of the seismic inverse problem. The focus will be mostly on the aspects that are of primary interest in the proposed thesis, namely migration (imaging) and velocity analysis via differential semblance optimization. The presentation is based on the notes of Professor Symes [48] as well as the paper by Stolk and Symes [47].

Modeling and migration

We assume that the medium occupies an open part $X \subset \mathbb{R}^n$, with boundary $\partial X \subset \mathbb{R}^{n-1}$. It is often the case that the medium is actually confined to a half-space by writing $X = \{x \in \mathbb{R}^n |\ x_n > 0\}$. Sources and receivers are distributed over a source and receiver manifold Σ_s, Σ_r that are assumed to be open parts of the boundary ∂X. Depending on the experiment configuration, Σ_s and Σ_r may partially coincide, or be totally disjoint (note that source and receiver locations, although discrete in reality, are commonly idealized as continuous). The measurements are made during the time interval $I_t = [0, T]$. The medium is described solely by the acoustic speed of sound $c(\mathbf{x})$. The constant density linear acoustic model of small amplitude wave propagation in a fluid is given by:

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} - \nabla^2 p = g, \tag{1}$$

where g is a body force divergence (the source), and where p represents the pressure field. Assuming that the source term g is known, the inverse problem is to infer the velocity c from a sampling of p at receiver locations. By Duhamel's principle, a causal solution for the inhomogeneous equation (1) is constructed as a superposition of Green's functions

$$p(\mathbf{x}, t) = \int_0^t \int G(\mathbf{x}, \mathbf{y}, t - \tau) f(\mathbf{y}, \tau) d\mathbf{y} d\tau$$

Because the inverse problem as formulated is intractable (the relation between c and p is highly nonlinear), a common practice is to linearize the wave equation, thereby writing the medium coefficient as the sum of a long scale component c_0 and a small scale perturbation δc. This is the *Born approximation*. The perturbation contains the singularities and is assumed to be small.

We further assume that the source has point support, i.e. $g(\mathbf{x}, t) = f(t)\delta(\mathbf{x} - \mathbf{x}_s)$. In the high-frequency single scattering approximation, the data are modeled as the trace of the first-order perturbation to the Green's function on the hypersurface $\{x_n = 0\}$:

$$d(\mathbf{x}_s, \mathbf{x}_r, t) = f(t) *_t \delta G(\mathbf{x}_s, \mathbf{x}_r, t) = f(t) *_t \int \left[\frac{2\delta c(\mathbf{x})}{c^3(\mathbf{x})} \int G(\mathbf{x}, \mathbf{x}_r, t - \tau) \frac{\partial^2 G}{\partial t^2}(\mathbf{x}_s, \mathbf{x}, \tau) d\tau \right] d\mathbf{x}. \tag{2}$$

We now define the linearized forward operator $F = F[c_0]$ to be the (linear) operator mapping the reflectivity $r = 2\delta c / c^3$ to the data d, i.e.

$$d(\mathbf{x}_s, \mathbf{x}_r, t) = F[c_0] r(\mathbf{x})$$

The next approximation consists the geometrical optics approximation to the Green's functions on the right-hand-side of (2). This approximation is accomplished via the *progressing wave expansion* (it amounts in keeping the first term in the progressing wave expansion - or

WKB expansion - of G, hence the 'ray+Born' term mentioned in the introduction). If $n = 3$, we obtain the so-called *Kirchhoff* modeling operator:

$$d(\mathbf{x}_s, \mathbf{x}_r, t) \simeq f'''(t) *_t \int a(\mathbf{x}, \mathbf{x}_s, \mathbf{x}_r) \delta \left(t - T(\mathbf{x}, \mathbf{x}_r) - T(\mathbf{x}, \mathbf{x}_s)\right) \frac{2\delta c(\mathbf{x})}{c^3(\mathbf{x})} d\mathbf{x}, \tag{3}$$

where $a(\mathbf{x}, \mathbf{x}_s, \mathbf{x}_r)$ is just an amplitude factor.

It is well known [39; 56; 48] that the linear modeling operator F, under certain assumptions, is a Fourier integral operator, i.e. it has a kernel of the form

$$K(\mathbf{x}_s, \mathbf{x}_r, t; \mathbf{x}) = \int A(\mathbf{x}, \mathbf{x}_s, \mathbf{x}_r, \tau) e^{i\tau(t - T(\mathbf{x}, \mathbf{x}_s, \mathbf{x}_r))} d\tau, \tag{4}$$

where $T(\mathbf{x}, \mathbf{x}_s, \mathbf{x}_r)$ is the two-way travel time (the time it takes for a ray to travel from the source at $\mathbf{x}_s$ to the reflector point $\mathbf{x}$ and back to the receiver at $\mathbf{x}_r$.

As we mentioned before, seismic processing involves the joint reconstruction of the background velocity and of the reflectivity. Both are a priori unknown. The reconstruction involves two distinct steps. Migration, or linearized inversion aims at the recovery of the reflectivity given the background velocity. Velocity analysis aims at estimating the background model, and is typically based on a set of reconstructions of the reflectivity.

The reconstruction of δc given the background velocity is essentially done by applying the adjoint of the linear modeling operator F to the data. This operator is precisely the migration operator and the outcome of this operation is a so-called seismic image of the reflectivity or more precisely of the high-frequency component of the velocity. The operator F has a left inverse if and only if the normal operator F^*F is invertible. In that case, a left inverse, given by $(F^*F)^{-1}F^*$ is optimal in the sense of least-squares.

It turns out that the redundancy in the data (the dimension of the data is $2n - 1$, whereas that of the reflectivity is n) can be used to simplify greatly the inversion process. The data can thus be partitioned into n-dimensional subsets, and each of this subset may be used for an independent reconstruction of the reflectivity. We will thereafter assume that the data are binned into sets of constant half-offset $h = (\mathbf{x}_s - \mathbf{x}_r)$. We write

$$T_h(\mathbf{x}, \mathbf{x}_r) = T(\mathbf{x}, \mathbf{x}_r - h, \mathbf{x}_r)$$

It has been shown by Beylkin [6] (and later Bleistein [8]) that the (asymptotic) inversion for δc is possible only if the following condition is satisfied

$$\det \begin{pmatrix} \dfrac{\partial^2 T_h}{\partial \mathbf{x} \partial \mathbf{x}_r} & \dfrac{\partial T_h}{\partial \mathbf{x}} \end{pmatrix} \neq 0.$$

This determinant is the so-called Beylkin determinant. If this condition is satisfied, then it was shown by Beylkin [6] that F_h (the modeling operator corresponding to that particular

subset of the data) is partially (and microlocally) invertible. The proof involves showing that the normal operator $F_h^* F_h$ is pseudodifferential with invertible symbol. An asymptotic inverse has the general form

$$G_h(\mathbf{x}, \mathbf{x}_r, t) \simeq (F_h^* F_h)^{-1} F_h^* \simeq F_h^* \equiv \int \phi(\mathbf{x}_r - h, \mathbf{x}_r, t) b(\mathbf{x}, h, \mathbf{x}_r, t) e^{i\tau(T_h(\mathbf{x}, \mathbf{x}_r) - t)} d\tau,$$

Here ϕ is a C^∞ cutoff function, equal to one on a subset of the data acquisition set, vanishing outside a larger compact set than the acquisition set, and smoothly varying in between. The operator G_h is called a migration operator. In particular, an asymptotic inverse to the operator in (3) yields the so-called *Kirchhoff migration* operator:

$$r(\mathbf{x}) \equiv (G_h d)(\mathbf{x}) \equiv \iiint b(h, \mathbf{x}_r, t) \delta\left(t - T_h(\mathbf{x}, \mathbf{x}_r)\right) d(h, \mathbf{x}_r, t)\ dt d\mathbf{x}_r dh \tag{5}$$

We define the operator G to be the map from data to the set of reconstructions:

$$G : d \mapsto r(h, \mathbf{x}) = (G_h d)(\mathbf{x}).$$

VELOCITY ANALYSIS VIA DIFFERENTIAL SEMBLANCE OPTIMIZATION

The *semblance principle*, which states that the images (of the singular part of the reflectivity) must agree, is the basis for the reconstruction of the background velocity. Given an estimate of the background medium, a set of images, each one corresponding to a particular "experiment" (or subset of the data) can be constructed in the manner described in the previous section. The reconstructed medium perturbation should obviously not depend on the offset h being used (there is only one Earth, hence one reflectivity), since this dependence is purely an artifact which expresses the redundancy in the data. The reconstructions will be essentially the same (as far as possible given that the inversion of the reflectivity is only partial) if the background medium is correctly chosen. However, if the background medium is incorrect, then there will some dependence on h and the images will not agree.

There are been several attempts to automate the procedure of determining the background medium. Several optimization methods have been proposed in which a functional of the set of images having an extremum when the images agree is optimized (see [15] for a comparison of various methods). In this paper, we consider functionals of the differential semblance type. Differential semblance functionals express semblance through the comparison of *neighboring bins* (in our case, offset). In particular, images are compared by taking the derivatives $\partial/\partial h$. We consider the asymptotic expression for the differential semblance functional as presented in [52]:

$$J_0[c_0] = \tfrac{1}{2} \| H \psi F[c_0] W G[c_0] d \|^2_{\mathcal{L}^2(\mathbb{R}^{2n-1})}, \tag{6}$$

This general form of differential semblance is in part due to ideas introduced in the thesis of M. Gockenbach and H. Song [25; 43]. Here $W = \partial/\partial h$ is the operator approximating

the derivative in the bin parameter(s). The application of the modeling operator $F[c_0]$ after formation of the bin difference makes the power of the output independent of the amplitude, up to an error which decays with increasing frequency. The function ϕ is a mute designed to control edge artifacts produced by differentiation in the bin direction, while the operator H is a smoothing pseudodifferential operator designed to keep the spectrum of the functional output comparable to that of the data (differentiation enhances frequency content). An appropriate choice is the inverse square root of the Helmholtz operator:

$$H = (I - \nabla^2)^{1/2}.$$

In the form presented here, the differential semblance objective is the only quadratic functional with smooth (C^∞) dependence on the background medium [47]. This is conditioned on the fact that W is a pseudodifferential operator. Moreover, for laterally homogeneous medium and with the hyperbolic moveout approximation of the travel times, it was shown in [52] that the stationary points of the differential semblance functional are actually asymptotic global minima. The hyperbolic moveout approximation assumption was relaxed by Stolk [46] who proved the same result with the correct (ray-theory) travel times, provided the medium is nonconstant above the reflector.

There is a nice connection between differential semblance optimization and pseudodifferential annihilators [23; 26]. Annihilators are a precise microlocal representation of the differential semblance concept. In fact, as mentioned in [45], the operator

$$F\frac{\partial}{\partial h}G$$

is a pseudodifferential operator which microlocally *annihilates* the data. If the background velocity is correct, then the differential semblance functional essentially vanishes. Hence, the reconstruction of the background velocity for a given data set amounts to determining the appropriate microlocal annihilator for that data set.

TIME REVERSAL IN RANDOMLY INHOMOGENEOUS MEDIA

In this section, we analyze and describe two important phenomena associated with time reversal in a randomly inhomogeneous medium:

- Super-resolution of the time-reversed, backpropagated wave field, due to multipathing.
- Self-averaging of the field which produces a statistically stable refocusing.

The analysis is based on a specific asymptotic regime, where the distance of propagation is much larger than the correlation length of the medium, fluctuations in the index of refraction are weak, and the wavelength is short compared to the correlation length of the randomly

inhomogeneous medium. This high-frequency regime is precisely the one that we described in the introduction and in which we will subsequently consider the seismic inverse problem.

The presentation that follows is based for the most part on the lectures that Professor Liliana Borcea gave during the spring semester of 2003 at Rice University, and which are collected in some notes [10]. Other references include the papers by Papanicolaou and colleagues [9; 35; 34], and the monologues by Ishimaru [27], Sobczyk [42], and by Rytov, Kravtsov and Tatarskii [40].

Introduction

In time reversal experiments, a pulse with carrier wavelength λ_0 is emitted by a point source, recorded by a transducer array of size a, and then re-emitted into the medium, time-reversed. The propagation distance L is large compared to the size a of the array. The randomly inhomogeneous medium fluctuates on a scale l which is short compared to a, but large compared to the wavelength λ_0. This regime is appropriate for the so-called parabolic approximation of the wave equation, in which the waves are assumed to propagate along a preferred direction. The z axis represents precisely that direction, whereas the $\mathbf{x}$-axis represents the plane of the coordinates (x, y) transverse to the direction of propagation. A schematic description of a time reversal experiment is shown on Figure 2.

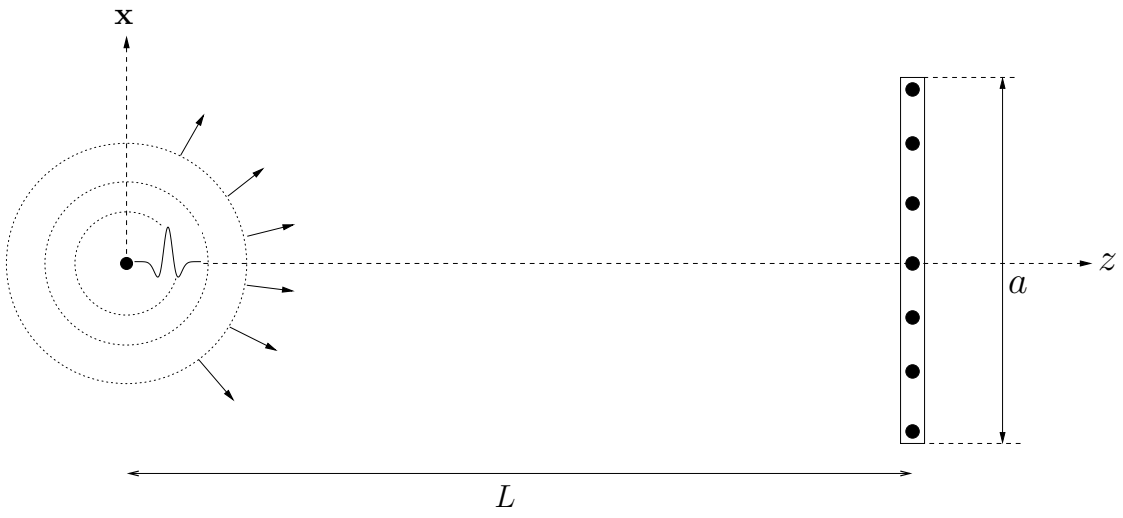

FIG. 2. Setup for the time reversal experiment

The parabolic approximation

We start with the (constant density) acoustic wave equation

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \quad \mathbf{x} = (x, y, z) \tag{7}$$

Taking its Fourier transform in time, we obtain the reduced wave equation (also known as the Helmholtz equation)

$$\Delta \widehat{u} + k^2 n^2(\mathbf{x})\widehat{u} = 0,$$

of greatest interest when the (constant) wave number $k = \omega/c_0$ is large. Here c_0 is a reference speed, $c(\mathbf{x})$ is the propagation speed, and $n(\mathbf{x}) = c_0/c(\mathbf{x})$ is the index of refraction of the medium, i.e. the ratio of the random propagation speed relative to the reference speed c_0. The transducer array is typically much smaller than the distance to the source, i.e. $a \ll L$. This puts the problem in the narrow beam regime, and the paraxial approximation [4; 3; 54] may be used as follows. Assuming that wave propagates in the z direction, the first step in obtaining the parabolic equation is to recognize that its phase progresses essentially as ikz. Therefore, writing the solution of the Helmholtz equation in the form

$$\widehat{u}(\mathbf{x}; \omega) = e^{ikz}\psi(\mathbf{x}; k),$$

i.e. a time harmonic standing wave of frequency ω, substitution into (7) yields a (parabolic) initial value problem for the wave amplitude ψ, in which the direction of propagation z plays the role of time [4]:

$$\begin{aligned} 2ik\psi_z + \Delta\psi + k^2(n^2 - 1)\psi &= 0 \\ \psi\big|_{z=0} &= \psi_0(\mathbf{x}; k). \end{aligned} \tag{8}$$

Since $\psi(\mathbf{x})$ is a slowly varying function of z and varies only over the distance of the scale l of the random fluctuations of the medium, we note that

$$k|\psi_z| \gg |\psi_{zz}| \quad \text{as long as} \quad l \gg \lambda$$

Therefore we can replace ∇^2 by the "transversal" Laplacian $\nabla_\perp^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$, and we obtain the parabolic approximation equation for $\psi(\mathbf{x})$ (also known as the Schrödinger equation)

$$2ik\psi_z + \Delta_{\mathbf{x}}\psi + k^2(n^2 - 1)\psi = 0 \tag{9}$$

where we have written $\mathbf{x} = (x, y)$ for the transverse coordinates.

The physical sense of the approximation of the wave equation (7) by the parabolic equation (9) consists in the restriction of the wave field energy to the slow and rather small scattering in the transverse direction with increasing z. This approximation ignores backscattering (i.e. neglecting the term ψ_{zz}) but includes multiple forward scattering (**multipathing**) and takes into account diffraction (in the Fresnel approximation). The parabolic approximation

is **not** valid near the point source where the beam geometry is not appropriate. Moreover, it may **not** be used when the typical size of the inhomogeneities l is comparable to the wavelength λ of the signal. When $kl = O(1)$ the variations of the index of refraction will produce oscillations in ϕ in the z direction on the scale $l \sim 1/k$. This in turn implies that $|\psi_z|$ and $|\psi_{zz}|$ are comparable. Therefore, the assumption $l \sim \lambda$ violates the validity of the parabolic approximation.

An important feature of this equation is the fact that it is *first order* in z. It is therefore sufficient to have only *one* boundary condition in the plane $z =$const.

Scaling, ordering and asymptotic analysis

In order to study the effect of the (random) heterogeneities on the backpropagated wave field, we consider the high-frequency regime, with weak fluctuations of the randomly inhomogeneous medium. More precisely, we consider the following scaling:

$$\lambda \ll l \ll L \tag{10}$$

This is the high-frequency regime: as mentioned in the introduction and in the previous section, multipathing is significant in this regime and the parabolic approximation may be used.

The index of refraction of the medium is assumed to (randomly) fluctuate about a constant background velocity:

$$\sigma\mu\left(\frac{\mathbf{x}}{l}, \frac{z}{l}\right) = n^2(\mathbf{x}, z) - 1. \tag{11}$$

Here l is the correlation length, i.e. the scale at which the medium fluctuates. The parameter σ represents the strength of the fluctuations. We assume that $\sigma \ll 1$, that is, we assume weak fluctuations of the medium so that multiple scattering occurs in the forward direction of propagation. The random field μ is assumed to be stationary, with mean zero (large scale variations do not exist), variance σ^2 and with normalized covariance (with dimensionless arguments)

$$R(\mathbf{x}, z) = \langle \mu(\mathbf{x} + \mathbf{x}', z + z')\mu(\mathbf{x}', z') \rangle$$

which decays sufficiently fast at infinity, so that the medium is *mixing*, that is, there are no long range correlations of the fluctuations. Note that large scale variations in the medium could be accommodated for in the model, at the cost of a more complex analysis.

The asymptotic analysis is done by rewriting the Schrödinger equation (9) in dimensionless form [10; 35]. Let L_z and $L_{\mathbf{x}}$ be characteristic length scales in the propagation direction, and in the transverse direction, so that $L_z \sim L$ and $L_{\mathbf{x}} \sim a$. We introduce a dimensionless wave number $k' = k/k_0$, with $k_0 = \omega_0/c_0$, and ω_0 the central frequency of the source pulse. We re-scale $\mathbf{x}$ and z by $\mathbf{x} = L_{\mathbf{x}}\mathbf{x}'$, $z = L_z z'$, and rewrite (9) in the new coordinates (thereby

substituting the expression (11) for the index of refraction) as

$$2i\frac{k_0k'}{L_z}\psi_{z'} + \frac{1}{L_{\mathbf{x}}^2}\Delta_{\mathbf{x}'}\psi + (k_0k')^2\sigma\mu\left(\frac{L_{\mathbf{x}}\mathbf{x}'}{l}, \frac{L_z z'}{l}\right)\psi = 0$$

Multiplying through by L_z/k_0, and dropping primes, we obtain

$$2ik\psi_z + \frac{L_z}{k_0L_{\mathbf{x}}^2}\Delta_{\mathbf{x}}\psi + k^2k_0L_z\sigma\mu\left(\frac{L_{\mathbf{x}}\mathbf{x}}{l}, \frac{L_z z}{l}\right)\psi = 0 \tag{12}$$

Note that the initial data should be scaled accordingly.

We now introduce three dimensionless parameters

$$\delta = \frac{l}{L_{\mathbf{x}}}, \quad \varepsilon = \frac{l}{L_z}, \quad \gamma = \frac{1}{k_0 l} = \frac{1}{2\pi}\frac{\lambda_0}{l}$$

which are the reciprocals of the transverse scale relative to correlation length, the reciprocal of the propagation distance relative to correlation length, and the central wavelength relative to correlation length.

The high-frequency regime assumption coupled to that of weak fluctuations requires that the dimensionless parameters γ, σ, ε, and δ are small, i.e.

$$\gamma \ll 1, \quad \sigma \ll 1, \quad \varepsilon \ll 1, \quad \delta \ll 1 \tag{13}$$

To make the scaling more precise, we introduce the Fresnel number

$$\theta = \frac{L_z}{k_0L_{\mathbf{x}}^2} = \gamma\frac{\delta^2}{\varepsilon}$$

Multiplying equation (12) by θ, we obtain (thereby using the above parameters)

$$2ik\theta\psi_z + \theta^2\Delta_{\mathbf{x}}\psi + k^2\frac{\delta^2}{\varepsilon^2}\sigma\mu\left(\frac{\mathbf{x}}{\delta}, \frac{z}{\varepsilon}\right)\psi = 0 \tag{14}$$

We further assume that we can relate ε and γ to σ and δ by the following relations

$$\varepsilon = \sigma^{2/3}\delta^{\zeta} \quad \gamma = \sigma^{2/3}\delta^{\beta}, \tag{15}$$

where $\zeta > 0$ and β shall be determined. With these additional assumptions, equation (14) becomes

$$\begin{aligned} &\psi_z = \frac{i}{2k}\theta\Delta_{\mathbf{x}}\psi + \frac{ik}{2}\frac{\sigma}{\gamma\varepsilon}\mu\left(\frac{\mathbf{x}}{\delta}, \frac{z}{\varepsilon}\right)\psi \\ \Rightarrow \quad &\psi_z = \frac{i}{2k}\delta^{2+\beta-\zeta}\Delta_{\mathbf{x}}\psi + \frac{ik}{2}\frac{\delta^{-\beta-\zeta}}{\sigma^{1/3}}\mu\left(\frac{\mathbf{x}}{\delta}, \frac{z}{\sigma^{2/3}\delta^{\zeta}}\right)\psi \end{aligned} \tag{16}$$

One way that the asymptotic regime (13) can be obtained is with the ordering

$$\varepsilon \ll \theta \ll \delta \ll 1. \tag{17}$$

This ordering corresponds to the so-called *white-noise* limit which is analyzed in [10]. It is different from the *transport limit* (corresponding to the case where $\lambda \sim l$) analyzed in [9; 34] and from the *high-frequency* limit treated in [35]. As the analysis in [34] shows, these various orderings all lead to a phase-space diffusion equation of the form (28). In particular, the validity of the parabolic approximation in the transport limit ordering (which was violated by the fact that it presumes the scaling $\lambda \sim l$) is restored by the so-called *narrow beam approximation*.

The ordering (17) has the following interpretation: we first take the white noise limit $\varepsilon \to 0$ (with δ fixed), then the high-frequency limit $\theta \to 0$, and then a broad beam limit $\delta \to 0$. We will analyze in detail these limits in the following sections.

The white noise limit

It is actually in the white noise limit $\varepsilon \to 0$ (with Fresnel number θ and δ fixed) that the parabolic approximation is valid, as was proved in [1]. In this limit, the wave field $\psi(z, \mathbf{x})$ satisfies an Itô-Schrödinger equation

$$d\psi^\delta = \left\{ \frac{i}{2k} \delta^{2+\beta-\zeta} \Delta_{\mathbf{x}} \psi^\delta - \frac{k^2}{8} R_0(0) \delta^{-\zeta-2\beta} \psi^\delta \right\} dz + \frac{ik}{2} \delta^{-\frac{\zeta}{2}-\beta} \psi^\delta dB\left(\frac{\mathbf{x}}{\delta}, z\right) \tag{18}$$

Here R_0 is the integrated covariance of the fluctuations μ, i.e.

$$R_0(\mathbf{x}) = \int_{-\infty}^{\infty} R(\mathbf{x}, s) ds,$$

and $B(\mathbf{x}, z)$ is an infinite-dimensional Brownian field with mean zero and covariance

$$\langle B(\mathbf{x}, z_1)\, B(\mathbf{y}, z_2) \rangle = R_0(\mathbf{x} - \mathbf{y}) z_1 \wedge z_2 = R_0(|\mathbf{x} - \mathbf{y}|) \min\{z_1, z_2\}$$

The Itô-Schrödinger equation is the result of the central limit theorem applied to equation (16) (which has the form of *Stratonovich*'s equation). In particular, defining

$$B^\varepsilon(\mathbf{x}, z) = \frac{1}{\sqrt{\varepsilon}} \int_0^z \mu\left(\mathbf{x}, \frac{s}{\varepsilon}\right) ds$$

then, as $\varepsilon \to 0$, and under certain conditions, one can show that this random process converges weakly to the Brownian field $B(\mathbf{x}, z)$ with the above covariance. The extra term in (18) is the Stratonovich correction.

The Green's function $\widehat{G}^\delta = \widehat{G}^\delta(z, z_0; \mathbf{x}, \boldsymbol{\xi}; k)$ with a point source at $(z_0, \boldsymbol{\xi})$ satisfies

$$\begin{cases} d\widehat{G}^\delta & = \left\{ \frac{i}{2k} \delta^{2+\beta-\zeta} \Delta_{\mathbf{x}} \widehat{G}^\delta - \frac{k^2}{8} R_0(0) \delta^{-\zeta-2\beta} \widehat{G}^\delta \right\} dz + \frac{ik}{2} \delta^{-\frac{\zeta}{2}-\beta} \widehat{G}^\delta dB\left(\frac{\mathbf{x}}{\delta}, z\right) \\ \widehat{G}^\delta\Big|_{z=z_0} & = \delta(\mathbf{x} - \boldsymbol{\xi}) \end{cases} \tag{19}$$

The Wigner equation

It is convenient to introduce here the tensor product of two Green's functions:

$$\begin{cases} \Gamma(z,\mathbf{x},\mathbf{y};\boldsymbol{\xi},\boldsymbol{\eta};k) & = \widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}^{\delta}(z,z_0;\mathbf{y},\boldsymbol{\eta};k)} \\ \Gamma(z_0,\mathbf{x},\mathbf{y};\boldsymbol{\xi},\boldsymbol{\eta};k) & = \delta(\mathbf{x}-\boldsymbol{\xi})\delta(\mathbf{y}-\boldsymbol{\eta}), \end{cases} \tag{20}$$

where the bar denotes complex conjugate. As we will show in a subsequent section, $\Gamma(z,\mathbf{x},\mathbf{x};\boldsymbol{\xi},\boldsymbol{\eta};k)$ describes the response, at the source plane, of a point source at $\boldsymbol{\eta}$, whose signal is recorded at the transducer array at $\mathbf{x}$, phase-conjugated, backpropagated, and observed at $\boldsymbol{\eta}$. The concepts of time reversal and backpropagation will be made clear in a subsequent section.

From equation (19) for the Green's function $\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)$, we can derive an equation for Γ as follows. The product rule writes

$$d\{\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}\} = \left[(d\widehat{G}^{\delta})\overline{\widehat{G}^{\delta}} + \widehat{G}^{\delta}(d\overline{\widehat{G}^{\delta}})\right] dz$$

where $\overline{\widehat{G}^{\delta}}$ solves (19) with k replaced by $-k$ (time reversal is equivalent to phase conjugation since G is real). Writing $\alpha = 2+\beta-\zeta$, $\Gamma(z,\mathbf{x},\mathbf{y};\boldsymbol{\xi},\boldsymbol{\eta};k)$ satisfies

$$\begin{cases} d\Gamma & = \left\{\frac{i}{2k}\delta^{\alpha}\left(\Delta_{\mathbf{x}}-\Delta_{\mathbf{y}}\right) + \frac{k^2}{4}\delta^{-\zeta-2\beta}\left[R_0\left(\frac{x-y}{\delta}\right)-R_0(0)\right]\right\}\Gamma dz \\ & \qquad + \frac{ik}{2}\delta^{-\frac{\zeta}{2}-\beta}\Gamma\left[dB\left(\frac{\mathbf{x}}{\delta},z\right)-dB\left(\frac{\mathbf{y}}{\delta},z\right)\right] \\ \Gamma\Big|_{z=z_0} & = \delta(\mathbf{x}-\boldsymbol{\xi})\delta(\mathbf{y}-\boldsymbol{\eta}) \end{cases} \tag{21}$$

We introduce the following change of transverse variables:

$$\widetilde{\mathbf{x}} = \frac{\mathbf{x}+\mathbf{y}}{2}, \qquad \widetilde{\mathbf{y}} = \frac{\mathbf{y}-\mathbf{x}}{\delta^{\alpha}}$$

so that

$$\mathbf{x} = \widetilde{\mathbf{x}} - \frac{\delta^{\alpha}\widetilde{\mathbf{y}}}{2} \quad \text{and} \quad \mathbf{y} = \widetilde{\mathbf{x}} + \frac{\delta^{\alpha}\widetilde{\mathbf{y}}}{2}.$$

Therefore:

$$\Delta_{\mathbf{x}} - \Delta_{\mathbf{y}} = (\nabla_{\mathbf{x}} - \nabla_{\mathbf{y}})\cdot(\nabla_{\mathbf{x}} + \nabla_{\mathbf{y}}) = (-\delta^{\alpha}\nabla_{\widetilde{\mathbf{y}}})\cdot(2\nabla_{\widetilde{\mathbf{x}}}) = -2\delta^{\alpha}\nabla_{\widetilde{\mathbf{x}}}\cdot\nabla_{\widetilde{\mathbf{y}}}.$$

Equation (21) thus becomes:

$$\begin{cases} d\Gamma & = \left\{ -\frac{i}{k}\nabla_{\widetilde{\mathbf{x}}}\cdot\nabla_{\widetilde{\mathbf{y}}} + \frac{k^2}{4}\delta^{-\zeta-2\beta}\left[R_0\left(-\delta^{\alpha-1}\widetilde{\mathbf{y}}\right)-R_0(0)\right]\right\}\Gamma dz \\ & \qquad + \frac{ik}{2}\delta^{-\frac{\zeta}{2}-\beta}\Gamma\left[dB\left(\frac{\widetilde{\mathbf{x}}}{\delta}-\frac{\delta^{\alpha-1}\widetilde{\mathbf{y}}}{2},z\right)-dB\left(\frac{\widetilde{\mathbf{x}}}{\delta}+\frac{\delta^{\alpha-1}\widetilde{\mathbf{y}}}{2},z\right)\right] \\ \Gamma\Big|_{z=z_0} & = \delta\left(\widetilde{\mathbf{x}}-\frac{\delta^{\alpha}\widetilde{\mathbf{y}}}{2}-\boldsymbol{\xi}\right)\delta\left(\widetilde{\mathbf{x}}+\frac{\delta^{\alpha}\widetilde{\mathbf{y}}}{2}-\boldsymbol{\eta}\right) \end{cases} \tag{22}$$

At this point, it is convenient to introduce the *Wigner distribution* which is defined as

$$W_\theta(z,\widetilde{\mathbf{x}},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k)=\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2}e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\Gamma(z,\widetilde{\mathbf{x}},\widetilde{\mathbf{y}};\boldsymbol{\xi},\boldsymbol{\eta};k)d\widetilde{\mathbf{y}} \tag{23}$$

The *Wigner distribution* is a convenient tool for the analysis of wave propagation in randomly inhomogeneous media [9; 35]. It can be interpreted as phase space wave energy. As we will discuss in a subsequent section, the time reversed, backpropagated wave field can be expressed in terms of the Wigner distribution, and in particular, the properties of the backpropagated field can be shown to be closely related to those of functionals of the Wigner distribution.

The Wigner distribution $W_\theta(z,\widetilde{\mathbf{x}},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k)$ defined by (23) satisfies the *Wigner equation*

$$\begin{aligned}dW_\theta=&\Big\{-\frac{1}{k}\mathbf{p}\cdot\nabla_{\widetilde{\mathbf{x}}}W_\theta+\frac{k^2}{4}\delta^{-\zeta-2\beta}\int\widehat{R}_0(\mathbf{q})\left[W_\theta(\mathbf{p}+\delta^{\alpha-1}\mathbf{q})-W_\theta(\mathbf{p})\right]d\mathbf{q}\Big\}dz\\&+\quad\frac{ik}{2}\delta^{-\frac{\zeta}{2}-\beta}\int d\widehat{B}(\mathbf{q},z)e^{i\mathbf{q}\cdot\frac{\widetilde{\mathbf{x}}}{\delta}}\cdot\left[W_\theta\left(\mathbf{p}-\frac{\delta^{\alpha-1}}{2}\mathbf{q}\right)-W_\theta\left(\mathbf{p}+\frac{\delta^{\alpha-1}}{2}\mathbf{q}\right)\right]d\mathbf{q},\end{aligned} \tag{24}$$

where we have omitted parameters other than wave numbers in the expressions for the Wigner distribution, for clarity of expression. The Wigner equation is a stochastic transport equation which can be obtained by taking the Fourier transform in $\widetilde{\mathbf{y}}$ of equation (22), with the Fourier transform defined as

$$\widehat{f}(\mathbf{p})=\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}e^{i\mathbf{p}\cdot\mathbf{x}}f(\mathbf{x})d\mathbf{x},\quad f(\mathbf{x})=\int_{\mathbb{R}^d}e^{-i\mathbf{p}\cdot\mathbf{x}}\widehat{f}(\mathbf{p})d\mathbf{p}.$$

Indeed, using integration by parts (note that the boundary terms vanish), it is easy to check that

$$\begin{aligned}\frac{1}{(2\pi)^2}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\left(-\frac{i}{k}\nabla_{\widetilde{\mathbf{x}}}\cdot\nabla_{\widetilde{\mathbf{y}}}\Gamma\right)d\widetilde{\mathbf{y}}&=-\frac{i}{k}\left[\frac{1}{(2\pi)^2}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\left(\nabla_{\widetilde{\mathbf{y}}}\left(\nabla_{\widetilde{\mathbf{x}}}\Gamma\right)\right)d\widetilde{\mathbf{y}}\right]\\&=\frac{i}{k}\left[\frac{1}{(2\pi)^2}\int\left(\nabla_{\widetilde{\mathbf{y}}}e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\right)\left(\nabla_{\widetilde{\mathbf{x}}}\Gamma\right)d\widetilde{\mathbf{y}}\right]\\&=-\frac{1}{k}\mathbf{p}\cdot\nabla_{\widetilde{\mathbf{x}}}\left[\frac{1}{(2\pi)^d}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\Gamma d\widetilde{\mathbf{y}}\right]\\&=-\frac{1}{k}\mathbf{p}\cdot\nabla_{\widetilde{\mathbf{x}}}W_\theta\end{aligned}$$

Similarly for the second term on the right-hand-side of (24), we have (thereby omitting the

factor in front of the integral for clarity of expression):

$$\begin{aligned}
&\frac{1}{(2\pi)^d}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\cdot\left[R_0\left(-\delta^{\alpha-1}\widetilde{\mathbf{y}}\right)-R_0(0)\right]\Gamma d\widetilde{\mathbf{y}}\\
&\qquad=\frac{1}{(2\pi)^d}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\cdot\left[\int e^{-i\mathbf{q}\cdot(-\delta^{\alpha-1}\widetilde{\mathbf{y}})}\widehat{R_0}(\mathbf{q})d\mathbf{q}-\int e^{-i\mathbf{q}\cdot 0}\widehat{R_0}(\mathbf{q})d\mathbf{q}\right]\Gamma d\widetilde{\mathbf{y}}\\
&\qquad=\frac{1}{(2\pi)^d}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\cdot\left[\int\left(e^{i\delta^{\alpha-1}\mathbf{q}\cdot\widetilde{\mathbf{y}}}-1\right)\widehat{R_0}(\mathbf{q})d\mathbf{q}\right]\Gamma d\widetilde{\mathbf{y}}\\
&\qquad=\int\widehat{R_0}(\mathbf{q})\left[\frac{1}{(2\pi)^d}\int\left[e^{i(\mathbf{p}+\delta^{\alpha-1})\cdot\widetilde{\mathbf{y}}}-e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\right]\Gamma d\widetilde{\mathbf{y}}\right]d\mathbf{q}\\
&\qquad=\int\widehat{R}_0(\mathbf{q})\left[W_\theta(\mathbf{p}+\delta^{\alpha-1}\mathbf{q})-W_\theta(\mathbf{p})\right]d\mathbf{q}
\end{aligned}$$

The last term of (24) can be obtained in a similar way.

The Fourier transform in $\widetilde{\mathbf{y}}$ of the initial condition in (22) yields

$$W_\theta(z_0,\widetilde{\mathbf{x}},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k)=\frac{1}{(2\pi)^d}\int e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\delta\left(\widetilde{\mathbf{x}}-\frac{\delta^\alpha\widetilde{\mathbf{y}}}{2}-\boldsymbol{\xi}\right)\delta\left(\widetilde{\mathbf{x}}+\frac{\delta^\alpha\widetilde{\mathbf{y}}}{2}-\boldsymbol{\eta}\right)d\widetilde{\mathbf{y}}$$

To simplify this expression, we take a smooth function ϕ of rapid decay, and consider the following functional:

$$I_\phi=\int\phi(\widetilde{\mathbf{x}})W_\theta(z_0,\widetilde{\mathbf{x}},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k)d\widetilde{\mathbf{x}}=\frac{1}{(2\pi)^2}\iint e^{i\mathbf{p}\cdot\widetilde{\mathbf{y}}}\phi(\widetilde{\mathbf{x}})\delta\left(\widetilde{\mathbf{x}}-\frac{\delta^\alpha\widetilde{\mathbf{y}}}{2}-\boldsymbol{\xi}\right)\delta\left(\widetilde{\mathbf{x}}+\frac{\delta^\alpha\widetilde{\mathbf{y}}}{2}-\boldsymbol{\eta}\right)d\widetilde{\mathbf{y}}d\widetilde{\mathbf{x}}$$

We introduce new variables:

$$\widetilde{\mathbf{z}}=\widetilde{\mathbf{x}}-\frac{\delta^\alpha\widetilde{\mathbf{y}}}{2}-\boldsymbol{\xi},\qquad\widetilde{\mathbf{w}}=\widetilde{\mathbf{x}}+\frac{\delta^\alpha\widetilde{\mathbf{y}}}{2}-\boldsymbol{\eta}.$$

Adding and subtracting these two equalities yield

$$\widetilde{\mathbf{z}}+\widetilde{\mathbf{w}}=2\widetilde{\mathbf{x}}-(\boldsymbol{\xi}+\boldsymbol{\eta})\quad\text{and}\quad\widetilde{\mathbf{z}}-\widetilde{\mathbf{w}}=\boldsymbol{\eta}-\boldsymbol{\xi}-\delta^\alpha\widetilde{\mathbf{y}},$$

so that

$$\widetilde{\mathbf{x}}=\tfrac{1}{2}\left[\widetilde{\mathbf{z}}+\widetilde{\mathbf{w}}+\boldsymbol{\xi}+\boldsymbol{\eta}\right]\quad\text{and}\quad\widetilde{\mathbf{y}}=\frac{1}{\delta^\alpha}\left[\boldsymbol{\eta}-\boldsymbol{\xi}+\widetilde{\mathbf{w}}-\widetilde{\mathbf{z}}\right].$$

Therefore, we obtain (thereby using the Jacobian of the transformation δ^α as well as the scaling rule for delta functions $\delta(\mathbf{x})=|\alpha|^2\delta(\alpha\mathbf{x})$)

$$I_\phi=\iint\frac{e^{i\mathbf{p}\cdot\left(\frac{\boldsymbol{\eta}-\boldsymbol{\xi}+\widetilde{\mathbf{y}}-\widetilde{\mathbf{x}}}{\delta^\alpha}\right)}}{(2\pi)^2}\delta(\widetilde{\mathbf{x}})\delta(\widetilde{\mathbf{y}})\phi\left(\tfrac{1}{2}\left[\widetilde{\mathbf{x}}+\widetilde{\mathbf{y}}+\boldsymbol{\xi}+\boldsymbol{\eta}\right]\right)d\widetilde{\mathbf{y}}d\widetilde{\mathbf{x}}=\frac{1}{(2\pi)^2}e^{i\mathbf{p}\cdot\left(\frac{\boldsymbol{\eta}-\boldsymbol{\xi}}{\delta^\alpha}\right)}\phi\left(\frac{\boldsymbol{\xi}+\boldsymbol{\eta}}{2}\right)$$

Therefore, the initial data for the Wigner distribution is

$$W_\theta(z_0,\widetilde{\mathbf{x}},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k)=\frac{1}{(2\pi)^2}e^{i\mathbf{p}\cdot\left(\frac{\boldsymbol{\eta}-\boldsymbol{\xi}}{\delta^\alpha}\right)}\delta\left(\widetilde{\mathbf{x}}-\frac{\boldsymbol{\xi}+\boldsymbol{\eta}}{2}\right)\equiv W_I(\widetilde{\mathbf{x}},\mathbf{p}).\tag{25}$$

Note that we have considered the Wigner distribution of Green's functions with sources at two different points. Therefore, this distribution needs not be real. It becomes real only if the two source points coincide, i.e. $\boldsymbol{\eta} = \boldsymbol{\xi}$. The exponential factor in (25) is extremely important. It carries the **phase information** that is crucial in the time-reversal refocusing phenomenon, since it takes place mostly due to phase cancellations, as we will show later.

The high-frequency and broad beam limits

In this section, we will write W_θ as a function of the wave number $\mathbf{p}$ only, to ease the notation. The first integral on the right hand side of (24) can be simplified by doing a Taylor expansion of W_θ around $\delta = 0$, thereby assuming that $\alpha - 1 > 0$:

$$W_\theta\left(\mathbf{p} + \delta^{\alpha-1}\mathbf{q}\right) = W_\theta(\mathbf{p}) + \delta^{\alpha-1}\mathbf{q}\cdot\nabla_{\mathbf{p}}W_\theta(\mathbf{p}) + \frac{\delta^{2(\alpha-1)}}{2}\mathbf{q}\cdot\Delta_{\mathbf{p}}W_\theta(\mathbf{p})\cdot\mathbf{q} + \dots$$

Using the fact that $\widehat{R}_0(\mathbf{q})$ is an even function (isotropy property), we have that

$$\int \mathbf{q}_i\widehat{R}_0(\mathbf{q})d\mathbf{q} = 0,$$

and

$$\int \mathbf{q}_i\mathbf{q}_j\widehat{R}_0(\mathbf{q})d\mathbf{q} = 0 \quad \text{for } i \neq j,$$

$$\text{and} \quad \int \mathbf{q}_i^2\widehat{R}_0(\mathbf{q})d\mathbf{q} = \left[\frac{\partial^2}{\partial {x_i}^2}\int e^{i\mathbf{q}\cdot\mathbf{x}}\widehat{R}_0(\mathbf{q})d\mathbf{q}\right]_{\mathbf{x}=0} = -\widehat{R}_0''(0) \equiv 4D.$$

Here D is the phase space diffusion coefficient. Therefore:

$$\int \widehat{R}_0(\mathbf{q})\left[W_\theta(\mathbf{p}+\delta^{\alpha-1}\mathbf{q}) - W_\theta(\mathbf{p})\right]d\mathbf{q} \simeq 2\delta^{2(\alpha-1)}D\Delta_{\mathbf{p}}W_\theta$$

Similarly, the second integral on the right hand side of (24) can be simplified by expanding W_θ in the following way:

$$W_\theta\left(\mathbf{p} - \frac{\delta^{\alpha-1}}{2}\mathbf{q}\right) = W_\theta\left(\mathbf{p} + \frac{\delta^{\alpha-1}}{2}\mathbf{q}\right) + (-\delta^{\alpha-1}\mathbf{q})\cdot\nabla_{\mathbf{p}}W_\theta(\mathbf{p}) + \dots$$

We also use the fact

$$d\nabla_{\widetilde{\mathbf{x}}}B\left(\frac{\widetilde{\mathbf{x}}}{\delta}, z\right) = i\int \mathbf{q}e^{i\mathbf{q}\cdot\frac{\widetilde{\mathbf{x}}}{\delta}}d\widehat{B}(\mathbf{q}, z)d\mathbf{q},$$

where we have used the Fourier transform of B to shift the derivative in the transverse coordinates to a multiplication in phase space. Therefore:

$$\int d\widehat{B}(\mathbf{q}, z)e^{i\mathbf{q}\cdot\frac{\widetilde{\mathbf{x}}}{\delta}}\cdot\left[W_\theta\left(\mathbf{p} - \frac{\delta^{\alpha-1}}{2}\mathbf{q}\right) - W_\theta\left(\mathbf{p} + \frac{\delta^{\alpha-1}}{2}\mathbf{q}\right)\right]d\mathbf{q} = i\delta^{\alpha-1}\nabla_{\mathbf{p}}W_\theta\cdot d\nabla_{\widetilde{\mathbf{x}}}B\left(\frac{\widetilde{\mathbf{x}}}{\delta}, z\right)$$

The Wigner equation (24) can therefore be rewritten as

$$dW_\theta = \left[-\frac{1}{k}\mathbf{p}\cdot\nabla_{\widetilde{\mathbf{x}}}W_\theta + \frac{k^2}{2}\delta^{2(\alpha-1)-\zeta-2\beta}D\Delta_{\mathbf{p}}W_\theta\right]dz - \frac{k}{2}\delta^{\alpha-1-\frac{\zeta}{2}-\beta}\nabla_{\mathbf{p}}W_\theta\cdot d\nabla_{\widetilde{\mathbf{x}}}B\left(\frac{\widetilde{\mathbf{x}}}{\delta},z\right) \quad (26)$$

In order for this stochastic partial differential equation to have a limit, we need to have $\alpha > 1$ on the one hand (this condition is necessary to be able to perform the above expansions), but also

$$\alpha - 1 = \frac{\zeta}{2} + \beta$$

Recalling that $\alpha = 2 + \beta - \zeta$, we obtain

$$\zeta = \frac{2}{3}, \quad \text{and} \quad \beta = \alpha - \frac{4}{3},$$

so that α may be chosen as a free parameter (greater than 1). With respect to the relations (15), we obtain:

$$\varepsilon = \sigma^{2/3}\delta^{2/3}, \quad \gamma = \sigma^{2/3}\delta^{\alpha-4/3} \quad (\alpha > 1), \quad \text{and} \quad \theta = \delta^\alpha \ll \delta$$

Nevertheless, we have the desired ordering

$$\varepsilon \ll \theta \ll \delta \ll 1$$

As $\delta \to 0$ (i.e. as $\theta \to 0$), the limit of (26) is the Itô-Liouville equation

$$dW = \left[-\frac{1}{k}\mathbf{p}\cdot\nabla_{\widetilde{\mathbf{x}}}W + \frac{k^2 D}{2}\Delta_{\mathbf{p}}W\right]dz - \frac{k}{2}\nabla_{\mathbf{p}}W\cdot d\nabla_{\widetilde{\mathbf{x}}}B\left(\frac{\widetilde{\mathbf{x}}}{\delta},z\right). \quad (27)$$

Going back to the initial data (25) for this equation, we can see that in the limit $\delta \to 0$, only points $\boldsymbol{\eta}$ close to $\boldsymbol{\xi}$ will matter. It is also important to note that the initial condition for W in (27) still depends on δ even though the asymptotic limit in the equation has been taken.

Moments of the Wigner distribution

From the Itô-Liouville equation (27), we can get closed equations for all the moments of the Wigner distribution W, not only for its mean, but also for moments with different wave numbers k. The wave number enters equation (27) as a parameter.

The average $W^{(1)} \equiv \langle W \rangle$ of W can be computed as the solution of

$$\begin{cases} \dfrac{dW^{(1)}}{dz} & = -\dfrac{1}{k}\mathbf{p}\cdot\nabla_{\widetilde{\mathbf{x}}}W^{(1)} + \dfrac{k^2 D}{2}\Delta_{\mathbf{p}}W^{(1)} \\ W^{(1)}\Big|_{z=z_0} & = W_I(\widetilde{\mathbf{x}},\mathbf{p}). \end{cases} \quad (28)$$

which has the form of an advection-diffusion equation in phase space.This equation is the basis of a simple explanation of the super-resolution phenomenon. Taking the Fourier transform of (28) with respect to both $\widetilde{\mathbf{x}}$ and $\mathbf{p}$, and using the method of characteristics, the solution can be shown to be

$$W^{(1)}(z,\widetilde{\mathbf{x}},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k) = \frac{1}{(2\pi)^{2d}}\iint\left[\iint e^{i\left[\frac{\mathbf{r}\cdot(\mathbf{p}-\mathbf{p}_0)}{k}+\mathbf{w}\cdot(\widetilde{\mathbf{x}}-\widetilde{\mathbf{x}}_0)-z\frac{\mathbf{w}\cdot\mathbf{p}}{k}\right]-\frac{Dz}{2}\left[r^2-(\mathbf{r}\cdot\mathbf{w})z+w^2\frac{z^2}{3}\right]}d\mathbf{r}d\mathbf{w}\right]\frac{W_I(\widetilde{\mathbf{x}}_0,\mathbf{p}_0)}{k}d\mathbf{p}_0d\widetilde{\mathbf{x}}_0$$

Now, using the definition (23) of the Wigner distribution, and the definition (20) of the Γ function, we obtain after some further computations, thereby assuming $\mathbf{x}=\mathbf{y}$:

$$\begin{aligned}\left\langle\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\eta};k)}\right\rangle &= \int W^{(1)}(z,\mathbf{x},\mathbf{p};\boldsymbol{\xi},\boldsymbol{\eta};k)d\mathbf{p}\\ &= \widehat{G}_0^{\delta}(z,z_0,\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}_0^{\delta}(z,z_0,\mathbf{x},\boldsymbol{\eta};k)}e^{-\left(k^2|\boldsymbol{\xi}-\boldsymbol{\eta}|^2a_e^2\right)/(2z^2)},\end{aligned} \tag{29}$$

where

$$a_e=\sqrt{\frac{Dz^3}{3}} \tag{30}$$

is the so-called **effective aperture** of the transducer array. Note that the exponential term in the inverse Fourier transform reduced to one because $\mathbf{x}=\mathbf{y}$ in this setting, and therefore $\widetilde{\mathbf{y}}=0$. A remarkable feature of (29) is that all of the statistics of the randomly inhomogeneous medium is contained in a_e. Furthermore, the effective aperture is independent of frequency, and thus the same expression is valid in the time domain. The importance of this moment formula will become clear in the next section, where we derive an explicit expression for the time-reversed signal.

The method of characteristics can also be used to study the second moment of the Wigner distribution at different points and in particular, to show that the processes $W(z,\mathbf{x}_1,\mathbf{p}_1;k_1)$ and $W(z,\mathbf{x}_2,\mathbf{p}_2;k_2)$ are *decorrelated* for $k_1\neq k_2$, i.e.

$$\langle W(z,\mathbf{x}_1,\mathbf{p}_1;k_1)\,W(z,\mathbf{x}_2,\mathbf{p}_2;k_2)\rangle \simeq \langle W(z,\mathbf{x}_1,\mathbf{p}_1;k_1)\rangle\,\langle W(z,\mathbf{x}_2,\mathbf{p}_2;k_2)\rangle\,,$$

In particular, this yields the other key property:

$$\begin{aligned}\left\langle\left(\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\eta};k)}\right)^2\right\rangle &\simeq \int\langle W(z,\mathbf{x},\mathbf{p}_1;k)\rangle\,d\mathbf{p}_1\int\langle W(z,\mathbf{x},\mathbf{p}_2;k)\rangle\,d\mathbf{p}_2\\ &= \left\langle\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\eta};k)}\right\rangle^2.\end{aligned} \tag{31}$$

That is, the variance of $\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}$ is essentially zero (note that this result holds for fixed frequency in this regime). This means that for any $c>0$, we have

$$P\left(\left|\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}-\left\langle\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}\right\rangle\right|\geq c\right)\leq\frac{1}{c^2}\left\langle\left(\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}-\left\langle\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}\right\rangle\right)^2\right\rangle\sim 0,$$

where $P(\cdot)$ denotes the probability, and where we have used Chebyshev inequality and (31). Therefore, we have

$$\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\eta};k)} \approx \left\langle \widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\xi};k)\overline{\widehat{G}^{\delta}(z,z_0;\mathbf{x},\boldsymbol{\eta};k)} \right\rangle, \tag{32}$$

that is, $\widehat{G}^{\delta}\overline{\widehat{G}^{\delta}}$ is **self-averaging** in this asymptotic regime, even at fixed k in this scaling (so that the stabilization is due to the decorrelation of W at different wave vectors $\mathbf{p}$).

Application to time reversal

We will now apply these results to the time reversal problem described in the introduction. A signal $f(t)$ emitted from a point source located at $\mathbf{y} = (0, 0, L)$ is recorded by an array of (point) transducers located at $\mathbf{x}_r = (rh/2, 0, 0)$, for $r = -N, \ldots, N$. We denote the manifold on which the transducers lie by Σ_r. The range of the source with respect to the array, as seen from the central element $\mathbf{x}_0$ is $|\mathbf{x}_0 - \mathbf{y}| = L$. It is then reversed in time and re-emitted into the medium. A diagram of the experiment is shown on Figure 3. We assume that the

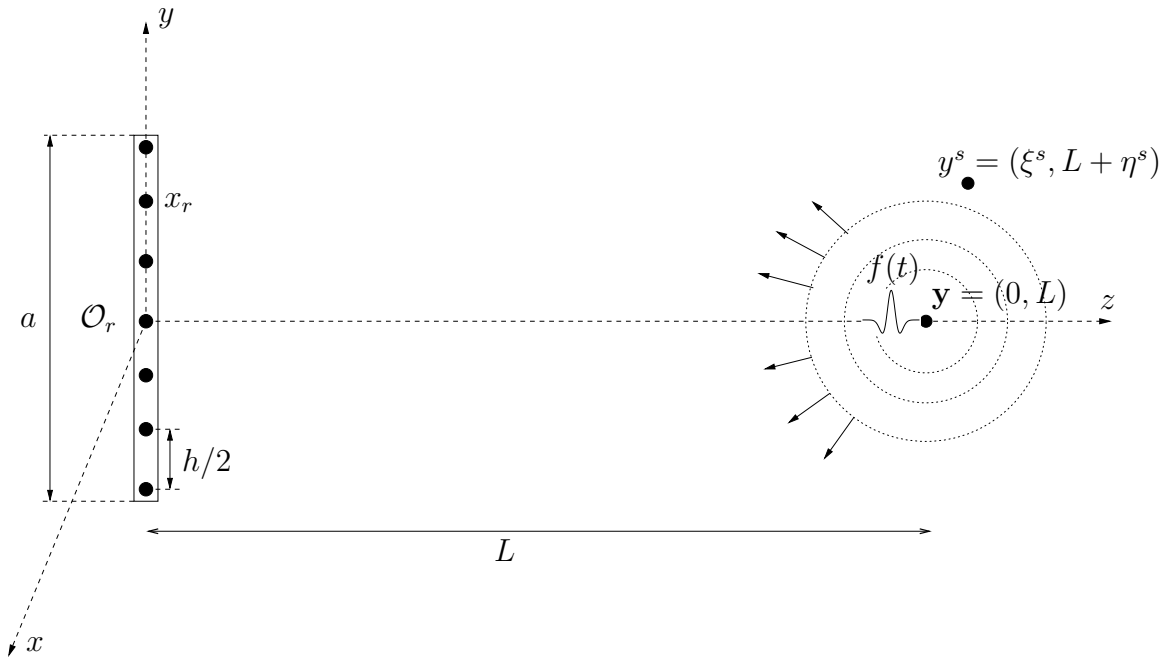

FIG. 3. Setup for the localized source experiment

pulse f emitted by the active target has the form

$$f(t) = -\frac{d}{dt}\left(\frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-i\omega_0 t - \frac{t^2}{2\sigma_t^2}}\right) = \frac{i\omega_0 + \frac{t}{\sigma_t^2}}{\sqrt{2\pi\sigma_t^2}} e^{-i\omega_0 t - \frac{t^2}{2\sigma_t^2}},$$

where ω_0 is the carrier (central) frequency and σ_t represents the width of the pulse. Writing $F(t)$ for the time derivative of $f(t)$ (the quantity in parenthesis above), the Fourier transform of f is given by:

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt = -\int_{-\infty}^{\infty} \frac{d}{dt}F(t)e^{i\omega t}dt = \frac{i\omega}{\sqrt{2\pi\sigma_t^2}} \int_{-\infty}^{\infty} e^{i(\omega-\omega_0)t-\frac{t^2}{2\sigma_t^2}}dt = i\omega\ e^{-\frac{1}{2}\sigma_t^2(\omega-\omega_0)^2}$$

where we have used integration by parts (the boundary terms vanish because of the presence of t^2 in the exponential term).

At fixed frequency ω, the two point Green's function satisfies the Helmholtz equation

$$\begin{aligned} \Delta\widehat{G}(\mathbf{x},\mathbf{y},\omega) + k^2n^2(\mathbf{x})\widehat{G}(\mathbf{x},\mathbf{y},\omega) &= -\delta(\mathbf{x}-\mathbf{y}), \\ \lim_{r\to\infty} r\left(\partial\widehat{\psi}/\partial r - ikn\widehat{\psi}\right) &= 0, \quad r = |\mathbf{x}-\mathbf{y}| \end{aligned} \tag{33}$$

The signal measured at a transducer at $\mathbf{x}_r$ can be computed as a superposition of Green's functions (Duhamel's principle):

$$d(\mathbf{x}_r,t) = \left(f(\cdot) *_t G(\mathbf{x}_r,\mathbf{y},\cdot)\right)(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f}(\omega)\widehat{G}(\mathbf{x}_r,\mathbf{y},\omega)e^{-i\omega t}d\omega,$$

Let $g(\mathbf{x},t) = f(t)\delta(\mathbf{x}-\mathbf{y})$ be the right-hand-side of the scalar wave equation corresponding to (33). We now view the correspondence between the source $g(\mathbf{x},t)$ and the sampled signals d as a linear operator

$$L : g \mapsto d. \tag{34}$$

The determination of the source g of acoustic wave motion from remote measurements is a notoriously ill-posed problem. This inverse problem can be separated into two parts: an optimal control for the wave equation and an inverse mixed initial-boundary value problem [53]. In the latter, the initial data of the solution of the wave equation are to be determined from its trace on a time-like hyperplane. It is a result proved in [53] that the initial data provide precisely the information needed to recover the source from its trace on the hyperplane. Suppose now that the forward model (34) is reformulated as this so-called inverse mixed problem. To recover the initial data, we may pose it as a output least-squares problem and attempt to compute the least-squares inverse. Because the source g is typically *oscillatory*, and because the normal operator L^*L is pseudodifferential (hence pseudolocal, i.e. it preserves the locii of high-frequency components), the singularities of the initial data *localized about the source location* should be reproduced by applying the adjoint operator L^* to the data. This is precisely the idea behind the *point spread function* which we will formally introduce later in this section. We now derive an expression for the adjoint operator L^*, thereby idealizing reality by assuming that the transducers form a continuous array of

size a. Following the definition of adjointness, we write

$$\langle L^* d, g\rangle = \langle d, Lg\rangle = \langle d, u\big|_{\Sigma_r}\rangle = \iint \left[\iint G(\mathbf{x}_r, \mathbf{y}, t-\tau) g(\mathbf{y}, \tau) d\tau d\mathbf{y}\right] d(\mathbf{x}_r, t) dt d\mathbf{x}_r$$
$$= \iint \left[\iint G(\mathbf{x}_r, \mathbf{y}, t-\tau) d(\mathbf{x}_r, -\tau) d\tau d\mathbf{x}_r\right] g(\mathbf{x}, t) dt d\mathbf{y}$$

Therefore, the adjoint operator applied to the data yields:

$$\left(L^* d\right)(\mathbf{y}; t) = \iint G(\mathbf{x}_r, \mathbf{y}, t-\tau) d(\mathbf{x}_r, -\tau) d\tau d\mathbf{x}_r \simeq \sum_{r=-N}^{N} d(\mathbf{x}_r, -t) *_t G(\mathbf{x}_r, \mathbf{y}, t) \tag{35}$$

Note that because the pulse is real, i.e. $f(t) = \overline{f(t)} \leftrightarrow \widehat{f}(\omega) = \overline{\widehat{f}(-\omega)}$, time reversal is **equivalent** to complex conjugation in frequency domain:

$$d(\mathbf{x}_r, -t) = \left(f(\cdot) *_t G(\mathbf{x}_r, \mathbf{y}, \cdot)\right)(-t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\widehat{f}(\omega)}\, \overline{\widehat{G}(\mathbf{x}_r, \mathbf{y}, \omega)} e^{-i\omega t} d\omega. \tag{36}$$

For the sake of comparison, we first show that in a homogeneous medium, the size of the refocused spot is approximately $\lambda L/a$. The underlying assumption here is that the array is operating in the remote-sensing regime so that $a \ll L$. There are two striking features of this refocusing in randomly inhomogeneous media. The first one is the super-resolution phenomenon, for which we derived a theoretical explanation in the previous section (see (29)). An intuitive explanation for this is the following: because of multipathing, the transducer array can capture waves that were initially moving away from it but get scattered onto it by the heterogeneities. As the result, the array captures a larger part of the total wave field emanating from the point source and thus appears to be larger than its physical size. Therefore, the heterogeneities of the medium *enhance* the resolution. We will derive an explicit formula which shows that the refocused spot is in this case $\lambda L/a_e$, where $a_e > a$ is the effective size of the array. The other key feature is that the time-reversed and backpropagated field is self-averaging.

Time reversal in homogeneous media

For a homogeneous medium in 3-D, the free space Green's function is given by

$$\widehat{G}_0(\mathbf{x}, \mathbf{y}) = \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|},$$

In time reversal, the back-propagated field is estimated at a *search point* $\mathbf{y}^S$, taken to be in the plane determined by $\mathbf{y}$ and the array, at range $L+\eta$ and cross-range ξ, i.e., $\mathbf{y}^S = (\xi, 0, L+\eta)$. The adjoint operator (35) is commonly referred to as the *point spread function* for time reversal:

$$\Gamma_0^{\mathtt{TR}}(\mathbf{y}^S; t) = \sum_{r=-N}^{N} d(\mathbf{x}_r, -t) *_t G_0(\mathbf{x}_r, \mathbf{y}^S, t) \tag{37}$$

Using (36) we obtain:

$$\begin{aligned}\widehat{\Gamma}_0^{\mathtt{TR}}(\mathbf{y}^S;\omega) &= \overline{\widehat{f}(\omega)} \sum_{r=-N}^{N} \overline{\widehat{G}_0(\mathbf{x}_r,\mathbf{y};\omega)}\widehat{G}_0(\mathbf{x}_r,\mathbf{y}^S;\omega) \\ &= \overline{\widehat{f}(\omega)} \sum_{r=-N}^{N} \frac{e^{ik\left(|\mathbf{x}_r-\mathbf{y}^S|-|\mathbf{x}_r-\mathbf{y}|\right)}}{\overline{16\pi^2|\mathbf{x}_r-\mathbf{y}||\mathbf{x}_r-\mathbf{y}^S|}}\end{aligned} \tag{38}$$

In a remote sensing regime ($a \ll L$), we can use the so-called *parabolic approximation* of the phase:

$$\begin{aligned}|\mathbf{x}_r-\mathbf{y}| &= \sqrt{L^2+x_r^2} \approx L + \frac{x_r^2}{2L} \\ |\mathbf{x}_r-\mathbf{y}^S| &= \sqrt{(L+\eta)^2+(x_r-\xi)^2} \approx L+\eta+\frac{(x_r-\xi)^2}{2(L+\eta)}\end{aligned}$$

(x_r refers to the x-coordinate of $\mathbf{x}_r$). This yields:

$$\begin{aligned}|\mathbf{x}_r-\mathbf{y}^S|-|\mathbf{x}_r-\mathbf{y}| &\approx \eta + \frac{(x_r-\xi)^2}{2(L+\eta)} - \frac{x_r^2}{2L} \\ &= \eta + \frac{\xi^2}{2(L+\eta)} - \left(\frac{\eta x_r^2}{2L(L+\eta)} + \frac{\xi x_r}{L+\eta}\right)\end{aligned}$$

and

$$|\mathbf{x}_r-\mathbf{y}^S||\mathbf{x}_r-\mathbf{y}| \approx L^2,$$

where we have used the fact that $x_r = \frac{rh}{2}$, and the relations $h \ll L$, $\eta \ll L$. Substitution of these estimates into (38), thereby using the fact that $k = \frac{\omega}{c_0}$, yields:

$$\widehat{\Gamma}_0^{\mathtt{TR}}(\mathbf{y}^S;\omega) \approx -\frac{i\omega\ e^{-\frac{1}{2}\sigma_t^2(\omega-\omega_0)^2} e^{i\frac{\omega}{c_0}\left(\eta+\frac{\xi^2}{2(L+\eta)}\right)}}{16\pi^2L^2} \sum_{r=-N}^{N} e^{-i\frac{\omega}{c_0}\left(\frac{\eta x_r^2}{2L(L+\eta)}+\frac{\xi x_r}{L+\eta}\right)}$$

Substituting this expression into (37), we obtain:

$$\Gamma_0^{\mathtt{TR}}(\mathbf{y}^S;t) \approx -\frac{1}{32\pi^3L^2}\int_{-\infty}^{\infty} i\omega\ e^{-\frac{1}{2}\sigma_t^2(\omega-\omega_0)^2} e^{i\omega\left[t-\frac{1}{c_0}\left(\eta+\frac{\xi^2}{2(L+\eta)}\right)\right]} \sum_{r=-N}^{N} e^{-i\frac{\omega}{c_0}\left(\frac{\eta x_r^2}{2L(L+\eta)}+\frac{\xi x_r}{L+\eta}\right)} d\omega,$$

Because the separation $h/2$ between the array elements is chosen so that the elements behave like an array of aperture $a = Nh \ll L$ and not like separate entities, the sum in the above expression may be replaced by an integral over the (continuous) interval $[-a/2, a/2]$, i.e.,

$$\int_{-a/2}^{a/2} e^{-i\frac{\omega}{c_0}\left(\frac{\eta x^2}{2L(L+\eta)}+\frac{\xi x}{L+\eta}\right)} dx\ = \frac{h}{2}\sum_{r=-N}^{N} e^{-i\frac{\omega}{c_0}\left(\frac{\eta x_r^2}{2L(L+\eta)}+\frac{\xi x_r}{L+\eta}\right)}$$

We obtain:

$$\Gamma_0^{\mathtt{TR}}(\mathbf{y}^S;t) \approx -\frac{1}{16\pi^3 L^2 h}\int_{-\infty}^{\infty} i\omega\, e^{-\frac{1}{2}\sigma_t^2(\omega-\omega_0)^2} e^{i\omega\left[t-\frac{1}{c_0}\left(\eta+\frac{\xi^2}{2(L+\eta)}\right)\right]} \int_{-a/2}^{a/2} e^{-i\frac{\omega}{c_0}\left(\frac{\eta x^2}{2L(L+\eta)}+\frac{\xi x}{L+\eta}\right)} dx\, d\omega,$$

Evaluation of this expression at the exact range $\eta = 0$, and at the arrival time $t = \frac{\xi^2}{2c_0 L}$ (corresponding to the parabolic shift in time) yields:

$$\Gamma_0^{\mathtt{TR}}\left(\xi,\eta=0;t=\frac{\xi^2}{2c_0L}\right) \approx \frac{-ic_0}{4\pi^2 Lh\xi} e^{-\frac{\xi^2 a^2}{8c_0L^2\sigma_t^2}} \sin\left(\frac{\omega_0\xi a}{2c_0L}\right) \sim \frac{1}{\xi}\sin\left(\frac{\pi\xi a}{\lambda_0 L}\right) e^{-\xi^2/2s^2}$$

where

$$s = \frac{2c_0L\sigma_t}{a} = \frac{2}{B}\frac{\lambda_0 L}{2}, \quad B = \frac{2\pi}{\omega_0,\sigma_t}.$$

Here B is the bandwidth of the pulse. This expression shows that the refocusing resolution is roughly given by the product of a sinc-like function with a Gaussian $e^{-\xi^2/2s^2}$. The focal spot size $\lambda_0 L/a$ is easily obtained from the sinc function. However, spurious oscillations (the so-called Fresnel zones) come along as a byproduct of the sinc. This is particularly true for narrow-band pulses, for which s is large and therefore the Gaussian is wide. However, for broad-band pulses, s is on the order of the spot size $\lambda_0 L/a$, so the refocusing resolution is still roughly $\lambda_0 L/a$, but the Fresnel zones have been eliminated by the narrow Gaussian. In either case, we have that the following (intuitive) relation between resolution and physical aperture of the array: the larger a, the better the refocusing.

Time reversal in randomly inhomogeneous media

In a randomly inhomogeneous medium, the point spread function in frequency domain is given by

$$\widehat{\Gamma}^{\mathtt{TR}}(\mathbf{y}^S;\omega) = \overline{\widehat{f}(t)} \sum_{r=-N}^{N} \overline{\widehat{G}(\mathbf{x}_r,\mathbf{y};\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S;\omega),$$

where $\widehat{G}$ is the random, time harmonic Green's function. Because the time-reversed and backpropagated field $\Gamma^{\mathtt{TR}}(\mathbf{y}^S,t)$ is self-averaging and because in the high-frequency regime $\widehat{\Gamma}^{\mathtt{TR}}(\mathbf{y}^S;\omega)$ is itself self-averaging, we may replace the product of the random Green's functions by its expectation. Then we use the moment formula (29) derived previously (see also the appendix) to obtain:

$$\begin{aligned}
\widehat{\Gamma}^{TR}(\xi,\eta=0,\omega) &\approx \overline{\widehat{f}(\omega)} \sum_{r=-N}^{N} \left\langle \overline{\widehat{G}(\mathbf{x}_r,\mathbf{y};\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S;\omega)\right\rangle \\
&\approx \overline{\widehat{f}(\omega)} e^{-\frac{k^2\xi^2 a_e^2}{2L^2}} \sum_{r=-N}^{N} \overline{\widehat{G}_0(\mathbf{x}_r,\mathbf{y};\omega)}\widehat{G}_0(\mathbf{x}_r,\mathbf{y}^S;\omega) \\
&= \widehat{\Gamma}_0^{TR}(\xi,\eta=0,\omega)\, e^{-\frac{k^2\xi^2 a_e^2}{2L^2}}
\end{aligned} \tag{39}$$

In the time domain, the point spread function evaluated at the arrival time $t = \xi^2/2c_0L$ yields:

$$\begin{aligned}\Gamma\left(\xi, \eta = 0, t = \frac{\xi^2}{2c_0L}\right) &\approx \frac{c_0}{4\pi^2 Lh\xi}\frac{-i}{\sqrt{2\pi\sigma_t^2}}\sin\left(\frac{\omega_0\xi a}{2c_0L}\right)e^{-\frac{A_e^2\xi^2}{8c_0^2L^2\sigma_t^2}} \\ &= \Gamma_0^{TR}\left(\xi, \eta = 0, t = \frac{\xi^2}{2c_0L}\right)e^{-\frac{2\pi^2 a_e^2\xi^2}{\lambda_0^2L^2}}\end{aligned} \tag{40}$$

Here A_e is the broad-band effective aperture given by

$$A_e^2 = a_e^2 + \left(\frac{4\pi a_e}{B}\right)^2$$

The expressions (39) and (40) show the extra Gaussian arising in the random medium case. This Gaussian factor is precisely the reason behind super-resolution. The spot size in this case becomes $\lambda_0 L/a_e$. The expression for the broad-band effective aperture shows that for narrow-band signals the physical aperture a is negligible while for broad-band signals it may contribute to super-resolution. In any case, it is often the case that $a \ll A_e$ when there is substantial multipathing.

Concluding remark

A key factor in the self-averaging of the backpropagated field (i.e. the "imaging" functional or point spread function Γ^{TR}) is the approximate cancellation of phases of the random Green's functions between the array and the point $\mathbf{y}$ and between the array and the point $\mathbf{y}^S$, respectively. Heuristically, the random Green's function has the form

$$\widehat{G}^\delta(\mathbf{x}, \mathbf{y}; k) \sim \frac{e^{ikr+i\phi}}{4\pi r}$$

so the random phase ϕ in the product

$$\widehat{G}^\delta(\mathbf{x}, \mathbf{y}; k)\overline{\widehat{G}^\delta(\mathbf{x}, \mathbf{y}^S; k)}$$

almost vanishes. This observation is the basis for all of the ideas that will be investigated as part of the proposed work. We present and explain these ideas in the following section.

APPLICATIONS TO SEISMIC IMAGING

The connection between the problem formulation from the previous section and that of section where we introduced the seismic inverse problem can be made as follows. The index of refraction contains now three distinct scales of velocity. The medium is described precisely

as in section (see (11)), except the component of velocity on the wavelength scale is now taken into account. We have:

$$n^2(\mathbf{x}) = 1 + \sigma\mu(\mathbf{x}) + \varrho(\mathbf{x}), \tag{41}$$

where $\varrho(\mathbf{x})$ corresponds to the high-frequency component of velocity - the reflectivity which can be resolved via migration. The data are modeled with a Born approximation, which yields the theoretical expression for the scattered field at a receiver at $\mathbf{x}_r$:

$$d(\mathbf{x}_s, \mathbf{x}_r, t) = \int_{-\infty}^{\infty} \frac{k^2 \widehat{f}(\omega)}{2\pi} \left[\int_{\mathcal{D}} \varrho(\mathbf{y}) \widehat{G}(\mathbf{x}_s, \mathbf{y}, \omega) \widehat{G}(\mathbf{x}_r, \mathbf{y}, \omega) d\mathbf{y} \right] e^{-i\omega t} d\omega \tag{42}$$

Note that this expression is the precise analog to (2), written in the frequency domain. However, the Green's functions in the above formula are random, contrary to those appearing in (2). Nevertheless, the relation between the reflectivity $\varrho(\mathbf{x})$ and the predicted data $d(\mathbf{x}_s, \mathbf{x}_r, t)$ is linear.

We omit momentarily the fact that seismic imaging techniques deal exclusively with reduced data set (i.e. common-offset, common shot, etc...) and we denote by t_{sr} the deterministic travel time

$$t_{sr}(\mathbf{y}) \equiv T(\mathbf{y}, \mathbf{x}_s) + T(\mathbf{y}, \mathbf{x}_r) = \frac{|\mathbf{x}_s - \mathbf{y}|}{c_0} + \frac{|\mathbf{x}_r - \mathbf{y}|}{c_0}, \tag{43}$$

i.e. the time it takes for a ray to travel from a source at $\mathbf{x}_s$ to a point $\mathbf{y}$ in the medium and back to a receiver at $\mathbf{x}_r$. Using the Kirchhoff migration method (recall equation (5) in section), the reconstruction of the reflectivity ϱ at a point $\mathbf{y}$ in the medium is performed as follows:

$$\varrho(\mathbf{y}) \equiv \iint b\left(\mathbf{x}_s, \mathbf{x}_r, T(\mathbf{y}, \mathbf{x}_s) + T(\mathbf{y}, \mathbf{x}_r)\right) d\left(\mathbf{x}_s, \mathbf{x}_r, T(\mathbf{y}, \mathbf{x}_s) + T(\mathbf{y}, \mathbf{x}_r)\right) \, d\mathbf{x}_r d\mathbf{x}_s \tag{44}$$

That is, backpropagation using Kirchhoff migration consists of evaluating expression (42) at the arrival time t_{sr}, and then summing (the discrete analog of the integral in (44)) the backpropagated fields over (a subset of) all sources and receivers with proper weighting.

Because the fluctuations in the medium are **not** known a priori, the backpropagation is done *fictitiously*, in a reference medium with no fluctuations. There remain random phases in such functionals corresponding to long random paths from the source to the reflector and back to the receiver (there is no conjugated Green's function, hence no random phase cancellation). The backpropagated field $d(\mathbf{x}_s, \mathbf{x}_r, t)$ is a randomly fluctuating function, spread out over a large time window, due to varying arrival times corresponding to different random paths in the medium. Consequently, spurious artifacts are created in the reconstructions. These techniques **lack** statistical stability.

To achieve self-averaging functionals, we must somehow cancel the random phases in $d(\mathbf{x}_s, \mathbf{x}_r, t)$. One way to achieve this goal, as suggested in [11], is to divide the data set into smaller subsets

and construct **local data covariances** by taking the **cross-correlation** of nearby traces $d(\mathbf{x}_s, \mathbf{x}_r, t)$ and $d(\mathbf{x}_{s'}, \mathbf{x}_{r'}, t)$, i.e.

$$d(\mathbf{x}_s, \mathbf{x}_r, t) \star d(\mathbf{x}_{s'}, \mathbf{x}_{r'}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\widehat{d}(\mathbf{x}_s, \mathbf{x}_r, \omega)} \widehat{d}(\mathbf{x}_{s'}, \mathbf{x}_{r'}, \omega) e^{-i\omega t} d\omega,$$

where we have used $\star$ to denote cross-correlation. Substitution of (44) yields

$$d(\mathbf{x}_s, \mathbf{x}_r, t) \star d(\mathbf{x}_{s'}, \mathbf{x}_{r'}, t) =$$
$$\frac{1}{2\pi} \int_{-\infty}^{\infty} k^4 |\widehat{f}(\omega)|^2 \left[\iint_{\mathcal{D}} \overline{\widehat{G}(\mathbf{x}_s, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{s'}, \mathbf{y}, \omega) \overline{\widehat{G}(\mathbf{x}_r, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{r'}, \mathbf{y}, \omega) d\mathbf{y} d\mathbf{y}' \right] e^{-i\omega t} d\omega$$

Clearly we obtain near cancellation of the random phases corresponding to nearby random paths from the source at $\mathbf{x}_s$ to a reflector at $\mathbf{y}$ to a receiver at $\mathbf{x}_r$, and from a source at $\mathbf{x}_{s'}$ to a reflector at $\mathbf{y}'$ to a receiver at $\mathbf{x}_{r'}$ (see Figure 4). We assume here that $|\mathbf{x}_s - \mathbf{x}_{s'}| \sim |\mathbf{x}_r - \mathbf{x}_{r'}| = O(l)$.

In essence, this approach can be viewed as a pre-processing step in which, starting with the randomly fluctuating data $d(\mathbf{x}_s, \mathbf{x}_r, t)$, we obtain a reduced, self-averaging data set.

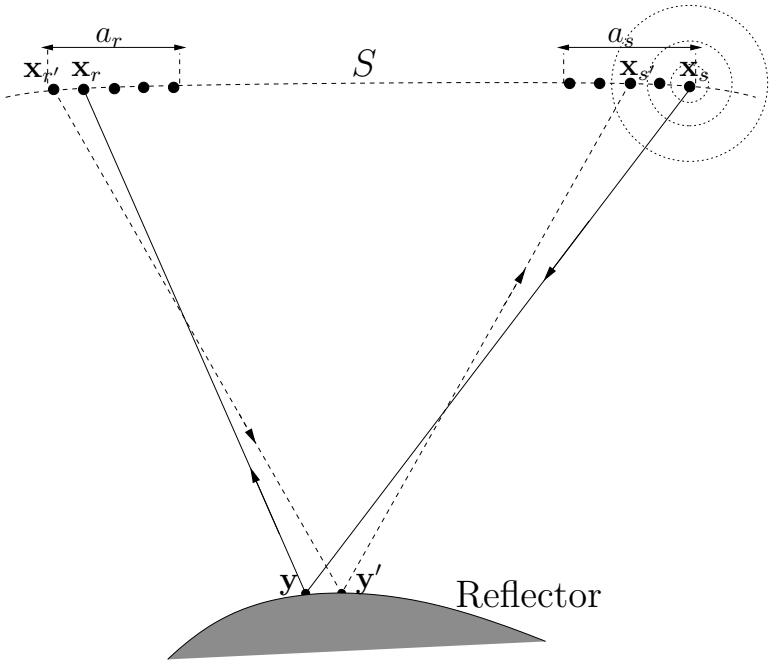

FIG. 4. A data pre-processing step

Cross-correlation tomography

We now explain how these local data covariances may be used for the purpose of estimating the background velocity. It would seem at first that all the travel time information (hence the velocity information) might be lost in the cross-correlation process. However, as we now show, there is still some *differential* time information present in the cross-correlated data. The problem then becomes one of somehow retrieving this information from the local data covariances. We then construct, in the case of a layered medium, a pseudodifferential annihilator which when applied to the (pre-processed) data with the *correct* background medium yields a vanishing outcome.

We will at first restrict ourselves to media that are laterally homogeneous, that is, the medium coefficient (the velocity) depends solely on depth. The layered medium assumption leads to simple explicit expressions for all quantities appearing in the differential semblance approach to velocity estimation [52]. Starting with the Kirchhoff modeling operator (3), the layered assumption assumption coupled to a stationary phase argument leads to the so-called *convolutional model* of primaries-only reflections. The derivation of this model is explained very carefully in [59]. Although being one of the simplest model in which to pose the velocity analysis problem, the convolutional model is widely used in the industry and therefore the following analysis is relevant to current seismic processing practices.

A natural binning scheme for this model is the *common midpoint gather*. Because all the gathers are in principle the same for a layered model, the data consists of a single common midpoint gather. The bins contain single traces, parametrized by offset $h = (\mathbf{x}_s - \mathbf{x}_r)$. The velocity parameter is the interval velocity $c_0(z)$, whereas the reflectivity is regarded as bin-dependent, i.e. $r = r(z, h)$. We parametrize velocity and reflectivity by the *two-way vertical travel time*

$$t_0 = 2 \int_0^z \frac{1}{v(\zeta)} d\zeta$$

rather than depth. Thus $c_0 = c_0(t_0)$ and $r = r(t_0, h)$. We denote by $T(t_0, h)$ the *two-way travel time* corresponding to a depth at t_0 and offset h, and by $T_0(t, h)$ the inverse function, i.e.

$$T\left(T_0(t,h),h\right) = t, \quad T_0\left(T(t_0,h),h\right) = t_0$$

With these conventions, the convolutional model reads

$$d(t,h) = f(t) *_t \left[a(t,h) r\left(T_0(t,h),h\right)\right] \equiv \left(F[c_0]r\right)(t,h),$$

where

$$r(z) = \tfrac{1}{2} \frac{d}{dz} \frac{\delta c(z)}{c_0(z)}$$

is yet another form of reflectivity. For simplicity, we will further assume that source signature deconvolution has been applied so that $f(t)$ is essentially impulsive ($f \sim \delta$), thus removing the convolution from the above expression:

$$d(t,h) = a(t,h) r\left(T_0(t,h),h\right) \equiv \left(F[c_0]r\right)(t,h),$$

Motivation

To motivate our approach, suppose for a moment that there are **no** fluctuations, and that we have a single reflector at depth z. We sketch this oversimplified situation on Figure 5. Because there is only one reflector, the trace has only one event at t_0 (no need to use a model

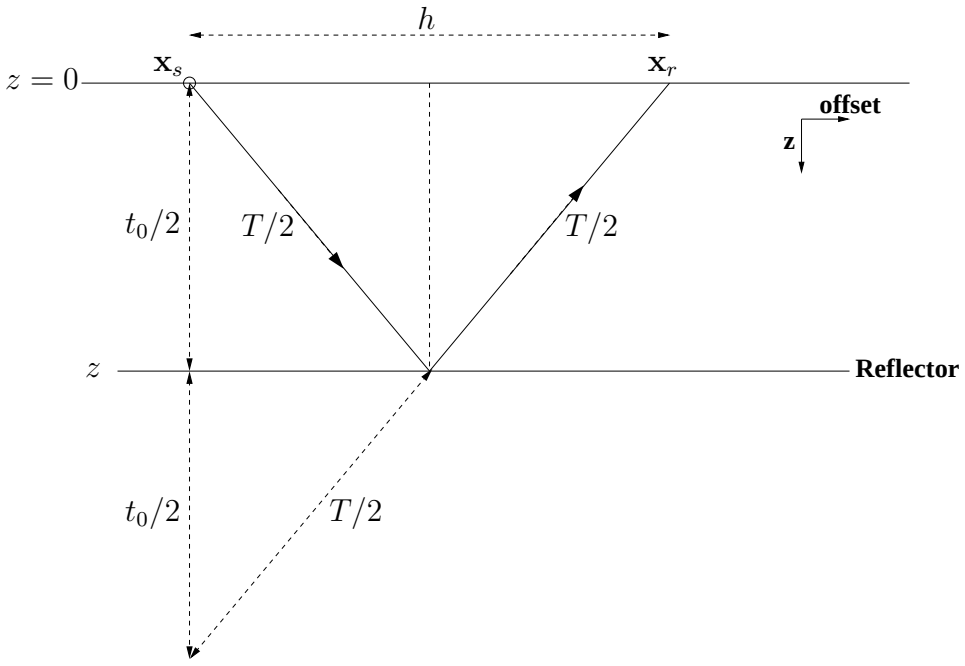

FIG. 5. A single reflector at depth z

in this case) so

$$d(t,h) = f\left(t - T(t_0,h)\right),$$

where f is the source pulse, and therefore

$$\begin{aligned}\left(d(h,\cdot) \star d(h',\cdot)\right)(t) &= f \star f\left(t + T(h,t_0) - T(h',t_0)\right)\\ &\simeq \left(f \star f\right)'(t)\frac{\partial T}{\partial h}(h,t_0)(h'-h),\end{aligned}$$

That is, the cross-correlation $d \star d$ contains **arrival time slowness** (derivative of the travel time with respect to surface location) information. This simple example shows that cross-correlations of traces do in fact contain velocity information. We now show how to retrieve that information.

The pseudodifferential annihilator

We denote by c_0^* the *correct* background velocity. The corresponding two-way travel time and inverse travel time are denoted $T^*(h,t_0)$ and $T_0^*(h,t)$, respectively. We assume that the

data are *model-consistent* (noise-free), i.e.

$$d(t,h) = r^*\left(T_0^*(t,h)\right),$$

where we have set the amplitude factor to 1 for simplicity (it can be re-introduced with almost no changes in the results that follow). Note that

$$\frac{\partial d}{\partial t}(t,h) = \frac{\partial T_0^*}{\partial t}(t,h)\frac{\partial r^*}{\partial t_0}\left(T_0^*(t,h)\right) \quad \Rightarrow \quad \frac{\partial r^*}{\partial t_0}\left(T_0^*(t,h)\right) = \left(\frac{\partial T_0^*}{\partial t}(t,h)\right)^{-1}\frac{\partial d}{\partial t}(t,h) \tag{45}$$

and

$$\frac{\partial d}{\partial h}(t,h) = \frac{\partial T_0^*}{\partial h}(t,h)\frac{\partial r^*}{\partial t_0}\left(T_0^*(t,h)\right). \tag{46}$$

The arrival (horizontal) slowness of the ray passing offset h at time t is defined as follows:

$$p(t,h) \equiv \frac{\partial T}{\partial h}\left(T_0(t,h),h\right) \tag{47}$$

The *stretch factor* is defined as:

$$s(t,h) \equiv \frac{\partial T_0}{\partial t}(t,h) = \left(\frac{\partial T}{\partial t_0}\left(T_0(t,h),h\right)\right)^{-1} \tag{48}$$

With these defined, we also have:

$$0 = \frac{\partial}{\partial h}T\left(T_0(t,h),h\right) = \frac{\partial T}{\partial t_0}\left(T_0(t,h),h\right)\frac{\partial T_0}{\partial h}(t,h) + \frac{\partial T}{\partial h}\left(T_0(t,h),h\right)$$

Therefore:

$$\frac{\partial T_0}{\partial h}(t,h) = -s(t,h)p(t,h) \tag{49}$$

Now, given a trial velocity c_0, we consider the following **weighted** cross-correlations:

$$\begin{aligned} C_t(t,h,h') &= \int_{-\infty}^{\infty}\left[d(t+\tau,h)\frac{\partial T_0}{\partial \tau}(\tau,h)\int_{-\infty}^{\tau} d(\cdot,h')\right]d\tau \\ C_h(t,h,h') &= \int_{-\infty}^{\infty}\left[d(t+\tau,h)\frac{\partial T_0}{\partial h}(\tau,h)\int_{-\infty}^{\tau} d(\cdot,h')\right]d\tau \end{aligned} \tag{50}$$

Both (deterministic) weights contain the stretch factor (48) and can therefore be expected to be slowly varying. In particular, we think that the self-averaging property of these cross-correlations is preserved. Of course, that this is really the case remains to be proved.

We now define the functional:

$$I(t,h) = \left(\frac{\partial C_t}{\partial h'} + \frac{\partial C_h}{\partial t}\right)(t,h,h'=h) \tag{51}$$

Using (45) and (46), we can compute the derivatives of (50). We have:

$$\begin{aligned}\frac{\partial C_t}{\partial h'}(t,h,h'=h) &= \int_{-\infty}^{\infty} d(t+\tau,h)\frac{\partial T_0}{\partial \tau}(\tau,h)\left[\int_{-\infty}^{\tau}\frac{\partial T_0^*}{\partial h}(\tau',h)\frac{\partial r^*}{\partial t_0}\left(T_0^*(\tau',h)\right)d\tau'\right]d\tau \\ &= \int_{-\infty}^{\infty} d(t+\tau,h)\frac{\partial T_0}{\partial \tau}(\tau,h)\left[\int_{-\infty}^{\tau}\frac{\partial T_0^*}{\partial h}(\tau',h)\left(\frac{\partial T_0^*}{\partial \tau'}(\tau',h)\right)^{-1}\frac{\partial d}{\partial \tau'}(\tau',h)d\tau'\right]d\tau\end{aligned}$$

Now, thereby using integration by parts,

$$\int_{-\infty}^{\tau}\frac{\partial T_0^*}{\partial h}(\tau',h)\left(\frac{\partial T_0^*}{\partial \tau'}(\tau',h)\right)^{-1}\frac{\partial d}{\partial \tau'}(\tau',h)d\tau' \simeq \frac{\partial T_0^*}{\partial h}(\tau,h)\left(\frac{\partial T_0^*}{\partial \tau}(\tau,h)\right)^{-1}d(\tau,h) - \ldots,$$

where the elided terms are of lower frequency content than the term explicitly displayed. They are of the same relative order of frequency as the terms neglected in the derivation of the convolutional model from the acoustic wave equation. We finally obtain:

$$\frac{\partial C_t}{\partial h'}(t,h,h'=h) \simeq \int_{-\infty}^{\infty} d(t+\tau,h)\left(\frac{\partial T_0}{\partial \tau}\frac{\partial T_0^*}{\partial h}\right)(\tau,h)\left(\frac{\partial T_0^*}{\partial \tau}(\tau,h)\right)^{-1}d(\tau,h)d\tau$$

Similarly,

$$\frac{\partial C_h}{\partial t}(t,h,h'=h) = \int_{-\infty}^{\infty}\frac{\partial d}{\partial t}(t+\tau,h)\frac{\partial T_0}{\partial h}(\tau,h)\left[\int_{-\infty}^{\tau} d(\tau',h)d\tau'\right]d\tau$$

We use a first integration by parts to shift the time integral to the other terms of the integrand (thereby assuming that the weight $\partial T_0/\partial h$ is slowly varying):

$$\begin{aligned}\frac{\partial C_h}{\partial t}(t,h,h'=h) &= -\int_{-\infty}^{\infty} d(t+\tau,h)\frac{\partial T_0}{\partial h}(\tau,h)\left[\int_{-\infty}^{\tau}\frac{\partial d}{\partial \tau'}(\tau',h)d\tau'\right]d\tau \\ &= -\int_{-\infty}^{\infty} d(t+\tau,h)\frac{\partial T_0}{\partial h}(\tau,h)\left[\int_{-\infty}^{\tau}\frac{\partial T_0^*}{\partial \tau'}(\tau',h)\left(\frac{\partial T_0^*}{\partial \tau'}(\tau',h)\right)^{-1}\frac{\partial d}{\partial \tau'}(\tau',h)d\tau'\right]d\tau\end{aligned}$$

Note that we can neglect the boundary terms because they are of lower frequency content. We proceed then as above to obtain

$$\frac{\partial C_h}{\partial t}(t,h,h'=h) = -\int_{-\infty}^{\infty} d(t+\tau,h)\left(\frac{\partial T_0}{\partial h}\frac{\partial T_0^*}{\partial \tau}\right)(\tau,h)\left(\frac{\partial T_0^*}{\partial \tau'}(\tau,h)\right)^{-1}d(\tau,h)d\tau.$$

Therefore, (51) becomes:

$$\begin{aligned}I(t,h) &= \int_{-\infty}^{\infty} d(t+\tau,h)\left[\left(\frac{\partial T_0^*}{\partial h}\frac{\partial T_0}{\partial \tau} - \frac{\partial T_0^*}{\partial \tau}\frac{\partial T_0}{\partial h}\right)(\tau,h)\left(\frac{\partial T_0^*}{\partial \tau}(\tau,h)\right)^{-1}\right]d(\tau,h)d\tau \\ &= \int_{-\infty}^{\infty} d(t+\tau,h)s(\tau,h)\left[p(\tau,h)-p^*(\tau,h)\right]d(\tau,h)d\tau\end{aligned} \tag{52}$$

This functional **vanishes** when $p^* = p$, i.e. when $c_0 = c_0^*$. It measures the mismatch of event slowness, weighted by data autocorrelation. This functional is precisely an annihilator in the sense of Section refsec:two. A procedure to determine the optimal velocity c_0 can be formulated as the following optimization problem:

$$\min_{c_0} J = \tfrac{1}{2}\|I(t,h)\|^2, \tag{53}$$

where $\|\cdot\|$ denotes the $\mathcal{L}^2$ norm. This is a variant of differential semblance optimization which we introduced in Section . It also appears to be a waveform variant of *stereotomography* [7] (an innovative approach to slope tomography).

Statement of the hypothesis

In this section, we formulate a set of conjectures which form the central points that the proposed work will address with respect to the velocity analysis problem. These conjectures are based on similar results that have been proved for other variants of differential semblance optimization (see for example [52]. It may be the case that we need to adopt the so-called *hyperbolic moveout approximation* of the two-way traveltime function:

$$T(t_0,h) \approx \sqrt{t_0^2 + \frac{h^2}{v_{\text{RMS}}^2(t_0)}}, \tag{54}$$

where

$$v_{\text{RMS}}(t_0) = \sqrt{\frac{1}{t_0}\int_0^{t_0} v^2} \tag{55}$$

is the root-mean-square velocity. This approximation, also known as the *Dix formula*, is derived and explained carefully in [52]. It allows, in particular, to obtain explicit formulations of terms such as the stretch factor (48) and the arrival time slowness (47).

The proposed thesis will seek to validate the following set of conjectures:

1. All stationary points of the objective function (53) are (asymptotic) **global** minima.
2. In the case where intermediate scale random fluctuations are allowed in the model, the cross-correlations (50) with **slowly** varying weights are statistically stable, as is the case without weights. Therefore, the objective function (53) is self-averaging, i.e. deterministic.
3. The gradient of the objective function J is also self-averaging. Thus, the optimization problem is itself statistically stable.
4. The stationary points of J with cross-correlation weights computed from the long-scale velocity component are optimal estimators of the background velocity.

Note that the demonstration of the validity of the above conjectures would entail that, under the assumptions made, **velocity analysis is essentially stable against random fluctuations on the medium scale** l.

Assessing the quality of migration

We consider in this section the second part of the seismic inverse problem. Assuming that we can reconstruct the background velocity in the way described in the last section, what can we say about the reconstruction of the reflectivity? As we mentioned in the introduction, it is quite clear that the uncertainty at the medium scales of velocity can be expected to pollute the resolution of migration techniques. Is it possible to quantify that (loss of) resolution explicitly?

Using the theory of time reversal and wave propagation in randomly inhomogeneous media, the resolution problem has been addressed in [11], but in a slightly different context. The problem considered in that paper is that of imaging the reflectivity ϱ of an extended target compactly supported in a domain $\mathcal{D} \subset \mathbb{R}^3$. By using a variant of the so-called *matched field* imaging functional [5; 13], the authors of [11] manage to transform the imaging problem into a *deterministic deblurring problem*. How relevant these results are to the quantification of the (loss of) resolution in practical migration techniques is unclear, but they are certainly worth understanding, with the hope that we can find a way to apply them for practical purposes.

The setup is as shown on Figure 6. The diameter of the target is assumed small compared to both L_s and L_r. We assume a remote-sensing regime, i.e. $a_s \ll L_s$ and $a_r \ll L_r$. The system of coordinates has its origin at $\mathcal{O}_s$, the center of the sub-array of sources. The z-axis passes through $\mathcal{O}_s$ and $\mathbf{y}_0$, a reference point in the target.

The matched field approach is extended in this case by taking the convolution of $d(\mathbf{x}_s, \mathbf{x}_r, t + t_{sr}(\mathbf{y}^S))$ with $d(\mathbf{x}_{s'}, \mathbf{x}_{r'}, -t - t_{s'r'}(\mathbf{y}^S))$ and evaluating at zero lag $t = 0$ (t_{sr} and $t_{s'r'}$ are defined according to (43)):

$$\Gamma^{IM}(\mathbf{y}^S) = \int_{-\infty}^{\infty} \frac{k^4 |\widehat{f}(\omega)|^2}{2\pi} \Big\{ \iint_{\mathcal{D}} \overline{\rho(\mathbf{y})} \rho(\mathbf{y}') \sum_{s,s'} \overline{\widehat{G}(\mathbf{x}_s, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{s'}, \mathbf{y}', \omega) e^{ik\left(|\mathbf{x}_s - \mathbf{y}^S| - |\mathbf{x}_{s'} - \mathbf{y}^S|\right)} \sum_{r,r'} \overline{\widehat{G}(\mathbf{x}_r, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{r'}, \mathbf{y}', \omega) e^{ik\left(|\mathbf{x}_r - \mathbf{y}^S| - |\mathbf{x}_{r'} - \mathbf{y}^S|\right)} d\mathbf{y} d\mathbf{y}' \Big\} d\omega \tag{56}$$

We assume that the sources lie on the surface orthogonal to the axis $\mathcal{O}_s - \mathbf{y}_0$, and that the receivers lie on a surface orthogonal to $\mathcal{O}_r - \mathbf{y}_0$. Note that this approach in effect supposes that the target $\mathcal{D}$ shrinks to the point $\mathbf{y}_0$. We write:

$$\mathbf{x}_s = (\mathbf{u}_s, 0), \quad \mathbf{x}_r = (\mathbf{u}_r, 0)$$

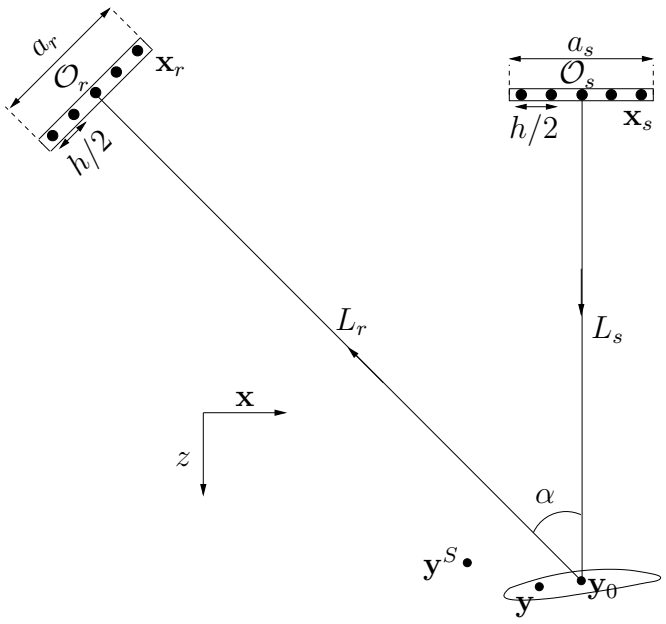

FIG. 6. The setup for distributed reflectivity imaging

Because we are in the remote sensing regime, we have near cancellation of the random phases in the pairs of Green's functions in (56), and by summing over the frequencies, we obtain a self-averaging operator. In particular, we can replace these pairs of random Green's functions by their respective expectations. We again idealize reality by assuming that the sources and receivers form a continuum indexed by $\mathbf{u}_s$ and $\mathbf{u}_r$, respectively. We obtain:

$$\Gamma^{IM}(\mathbf{y}^S) = \iint\limits_{\mathcal{D}} \frac{8\varrho(\mathbf{y})\varrho(\mathbf{y}')}{\pi h^4} \left[\int_{-\infty}^{\infty} k^4 |\widehat{f}(\omega)|^2 \left\langle \widehat{\mathcal{F}}_s(\mathcal{O}_s, \mathbf{y}, \mathbf{y}', \mathbf{y}^S, \omega) \right\rangle \left\langle \widehat{\mathcal{F}}_r(\mathcal{O}_r, \mathbf{y}, \mathbf{y}', \mathbf{y}^S, \omega) \right\rangle d\omega \right] d\mathbf{y} d\mathbf{y}',$$

where

$$\left\langle \widehat{\mathcal{F}}_s(\mathcal{O}_s, \mathbf{y}, \mathbf{y}', \mathbf{y}^S, \omega) \right\rangle = \iint \left\langle \overline{\widehat{G}(\mathbf{x}_s, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{s'}, \mathbf{y}', \omega) \right\rangle e^{ik\left(|\mathbf{x}_s - \mathbf{y}^S| - |\mathbf{x}_{s'} - \mathbf{y}^S|\right)} d\mathbf{u}_s d\mathbf{u}_{s'}$$

$$\left\langle \widehat{\mathcal{F}}_r(\mathcal{O}_r, \mathbf{y}, \mathbf{y}', \mathbf{y}^S, \omega) \right\rangle = \iint \left\langle \overline{\widehat{G}(\mathbf{x}_r, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{r'}, \mathbf{y}', \omega) \right\rangle e^{ik\left(|\mathbf{x}_r - \mathbf{y}^S| - |\mathbf{x}_{r'} - \mathbf{y}^S|\right)} d\mathbf{u}_r d\mathbf{u}_{r'}$$

The second moments in the above expression involve Green's functions at four distinct points. Such general moment formula is derived in [27; 40] and given explicitly in [22]. In the appendix, we show that the formula as given in [22] can be reduced in our context to the following expression:

$$\left\langle \overline{\widehat{G}(\mathbf{x}_s, \mathbf{y}, \omega)} \widehat{G}(\mathbf{x}_{s'}, \mathbf{y}', \omega) \right\rangle \approx \overline{\widehat{G}_0(\mathbf{x}_s, \mathbf{y}, \omega)} \widehat{G}_0(\mathbf{x}_{s'}, \mathbf{y}', \omega) e^{-\frac{k^2 D L_s}{6}\left[|\mathbf{u}_s - \mathbf{y}_{s'}|^2 + (\mathbf{u}_s - \mathbf{u}_{s'}) \cdot (\mathbf{y} - \mathbf{y}')^\perp + |(\mathbf{y} - \mathbf{y}')^\perp|^2\right]}$$

In the system of coordinates described previously, and with the notation

$$\mathbf{y} = (\boldsymbol{\xi}, L_s + \eta), \quad \mathbf{y}' = (\boldsymbol{\xi}', L_s + \eta'), \quad \mathbf{y}^S = (\boldsymbol{\xi}^S, L_s + \eta^S),$$

we obtain:

$$\left\langle \widehat{\mathcal{F}}_s(\mathcal{O}_s, \mathbf{y}, \mathbf{y}', \mathbf{y}^S, \omega) \right\rangle =$$
$$\iint \frac{e^{ik\left(|\mathbf{x}_{s'} - \mathbf{y}'| - |\mathbf{x}_{s'} - \mathbf{y}^S| - |\mathbf{x}_s - \mathbf{y}| + |\mathbf{x}_s - \mathbf{y}^S|\right)}}{16\pi^2 |\mathbf{x}_s - \mathbf{y}||\mathbf{x}_{s'} - \mathbf{y}'|} e^{-\frac{k^2 D L_s}{6}\left[|\mathbf{u}_s - \mathbf{u}_{s'}|^2 + (\mathbf{u}_s - \mathbf{u}_{s'})\cdot(\boldsymbol{\xi} - \boldsymbol{\xi}')^\perp + |(\boldsymbol{\xi} - \boldsymbol{\xi}')^\perp|^2\right]} d\mathbf{u}_s d\mathbf{u}_{s'}$$

The phases and amplitudes can now be approximated with the parabolic approximation, and with some additional approximations (see [11] for details), we finally obtain:

$$\left\langle \widehat{\mathcal{F}}_s(\mathcal{O}_s, \mathbf{y}, \mathbf{y}', \mathbf{y}^S, \omega) \right\rangle \approx \frac{a_s^2}{8\pi k^2 D L_s |\mathcal{O}_s - \mathbf{y}^S|^2} e^{ik(\eta - \eta') - \frac{3k^2 D L_s |\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}{8} - \frac{|\boldsymbol{\xi} + \boldsymbol{\xi}' - 2\boldsymbol{\xi}^S|^2}{8 D L_s^3}}$$

The calculation of the term corresponding to the sources is very similar. With these results, and after further simplifications that we omit here, expression (56) becomes, thereby writing $\eta \equiv y_1$ and $\boldsymbol{\xi} \equiv (y_2, y_3)$) :

$$\Gamma(\mathbf{y}^S) \approx \widetilde{C}(\mathbf{y}^S, D, f) \int_{\mathcal{D}} \varrho^2(\mathbf{y}) e^{-\frac{(y_2 - y_2^S)^2}{2}\left(\frac{1}{D L_s^3} + \frac{1}{D L_r^3}\right) - \frac{(y_1 - y_1^S)^2}{2 D L_s^3} - \frac{\left[(y_1 - y_1^S)\cos\alpha - (y_3 - y_3^S)\sin\alpha\right]^2}{2 D L_r^3}} d\mathbf{y} \quad (57)$$

This functional does not reconstruct the reflectivity ϱ pointwise, but rather the squared reflectivity ϱ^2 averaged over an area defined by the intersection of three Gaussians (one for each component) centered at $\mathbf{y}^S$ and with variance proportional to the effective apertures (recall that $a_e = \sqrt{DL^3/3}$). As can be seen from the expression (57), some range information remains in the matched field functional. However, as the angle α decreases, the variance of the Gaussian in the range direction increases (the Gaussian is quite elongated in depth), so that a single array completely loses depth information. The key to resolution enhancement appears therefore to obtain multiple views of the reflectivity from the source-receiver sub-arrays. Then, the reflectivity could be obtained by estimating ϱ^2 via a least-squares deblurring method using data for several source-receiver offset angle α. This remains to be investigated, though.

Practical aspects

A large part of the proposed work relies naturally on a numerical investigation of the properties laid out in the previous section. The implementation of all the operators (modeling, migration) and functionals that we have defined thus far will be done in the framework of the following software packages. The Standard Vector Library (SVL) is a collection of C++ base classes which implement the basic mathematical components of calculus in Hilbert space such as vectors and operators, among others key concepts. It also implements numerous Newton-based optimization algorithms. The Timestepping Simulation for Optimization (TSOpt) provides a framework to construct efficient simulation-driven optimization applications, and serves as the interface to SVL.

DISCUSSION AND CONCLUSIONS

As discussed in this proposal, we wish to address the seismic inverse problem from a very innovative point of view. In essence, the proposed work represents an attempt to connect two very well established yet inherently different theories. On the one hand, seismic imaging techniques, typically deterministic, are based on the commonly assumed dichotomy of the velocity field between long and short scales. On the other hand, the theory of wave propagation in stochastic media, based on the conviction that classical (i.e. deterministic) models and theories are too idealized and do not adequately reflect the complexity and heterogeneities of numerous real media, usually views the short and medium scales of velocity as random fields. The proposed work will make the distinction more subtle. In particular, only the intermediate scale of velocity is regarded as randomly fluctuating and it is precisely the uncertainty at this scale that we wish to understand, and especially its influence on the resolution of the long (background velocity) and short (image) scales. Using ideas on time-reversed acoustics developed by G. Papanicolaou and his group, we propose to pre-process the seismic data in such a way that we obtain self-averaging data subsets, that is, data that are essentially stable with respect to the random fluctuations at the intermediate scale of velocity. Because the pre-processing step involves the cross-correlation of nearby traces, it would seem that all of the travel time information (hence information on the background medium) is lost. We demonstrated in this proposal, using the convolutional model (thereby assuming a laterally homogeneous medium), that this is not the case, and that the cross-correlated data do contain *slowness* information. We then formulated an explicit velocity analysis algorithm to retrieve that information, hence the background medium, as well as a certain number of hypothesis which constitute the main points that this thesis plans to address in a first stage. With respect to the quality of seismic images, we have described how the matched field functional can be used to transform the imaging problem into a deterministic deblurring problem. We propose to investigate how these ideas can be used in practice for the purpose of estimating the resolution of migration techniques.

APPENDIX A: THE MOMENT FORMULA

The *mutual coherence function* or second moment of $\widehat{G}$ as given in [22] is:

$$\left\langle \widehat{G}(\mathbf{r}_1,\mathbf{r}_2,\omega)\overline{\widehat{G}(\mathbf{r}_3,\mathbf{r}_4,\omega)}\right\rangle = \widehat{G}_0(\mathbf{r}_1,\mathbf{r}_2,\omega)\overline{\widehat{G}_0(\mathbf{r}_3,\mathbf{r}_4,\omega)}e^{-H(\mathbf{r}_c,\boldsymbol{\rho}_{13},\boldsymbol{\rho}_{24})},$$

where the (free space) two-point Green's function is

$$\widehat{G}_0(\mathbf{x},\mathbf{y}) = \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}, \quad k=\frac{\omega}{c_0}$$

and where

$$H\left(\mathbf{r}_c,\boldsymbol{\rho}_{13},\boldsymbol{\rho}_{24}\right) = \frac{k^2|\mathbf{r}_c|}{4}\int_0^1 \left\{R(\mathbf{0}) - R\left[(1-\zeta)\boldsymbol{\rho}_{24}+\zeta\boldsymbol{\rho}_{13}\right]\right\}d\zeta. \tag{58}$$

These equations allow for four distinct spatial points $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, and $\mathbf{r}_4$. The center coordinate and the projected difference coordinates are given by

$$\mathbf{r}_c = \frac{(\mathbf{r}_1 - \mathbf{r}_2) + (\mathbf{r}_3 - \mathbf{r}_4)}{2}$$
$$\boldsymbol{\rho}_{ij} = (\mathbf{r}_i - \mathbf{r}_j) - \left(\frac{\mathbf{r}_c \cdot (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_c|}\right)\frac{\mathbf{r}_c}{|\mathbf{r}_c|}$$

In the remote-sensing regime, the mean-square phase deviation of the signal

$$\frac{k^2|\mathbf{r}_c|R(\mathbf{0})}{4}$$

will be a reasonably large number [22]. This allows the use of a second-order Taylor expansion of $R(\boldsymbol{\rho})$ about $\boldsymbol{\rho} = \mathbf{0}$ in the integrand of (58) since the exponentiated error produced by such an expansion will be very small. The Taylor expansion of $R(\boldsymbol{\rho})$ about $\boldsymbol{\rho} = \mathbf{0}$ yields:

$$R(\boldsymbol{\rho}) = R(\mathbf{0}) + \boldsymbol{\rho}R'(\mathbf{0}) + \frac{\rho^2}{2}R''(\mathbf{0}) + \dots$$

We use the assumption that R is isometric, i.e. $R(\boldsymbol{\rho}) = R(|\boldsymbol{\rho}|)$ to cancel the first derivative:

$$R'(\mathbf{0}) = \frac{R(\mathbf{0}+\mathbf{h}) - R(\mathbf{0}-\mathbf{h})}{2\mathbf{h}} = 0$$

Therefore, thereby using the fact that $R''(\mathbf{0}) = -4D$, we obtain:

$$R(\mathbf{0}) - R(\boldsymbol{\rho}) = -\frac{\rho^2}{2}R''(\mathbf{0}) = 2D\rho^2,$$

and

$$\begin{aligned}\int_0^1 \left\{R(\mathbf{0}) - R\left[(1-\zeta)\boldsymbol{\rho}_{24} + \zeta\boldsymbol{\rho}_{13}\right]\right\} d\zeta &= 2D\int_0^1 \left[(1-\zeta)\boldsymbol{\rho}_{24} + \zeta\boldsymbol{\rho}_{13}\right]^2 d\zeta \\ &= \frac{2D}{3}\left(\rho_{24}^2 + \rho_{13}^2 + \boldsymbol{\rho}_{24}\cdot\boldsymbol{\rho}_{13}\right)\end{aligned}$$

Therefore, we obtain:

$$H\left(\mathbf{r}_c, \boldsymbol{\rho}_{13}, \boldsymbol{\rho}_{24}\right) \approx \frac{k^2 D|\mathbf{r}_c|}{6}\left[\rho_{24}^2 + \rho_{13}^2 + \boldsymbol{\rho}_{24}\cdot\boldsymbol{\rho}_{13}\right]$$

and the second moment formula becomes:

$$\left\langle \widehat{G}(\mathbf{r}_1, \mathbf{r}_2, \omega)\overline{\widehat{G}(\mathbf{r}_3, \mathbf{r}_4, \omega)}\right\rangle = \widehat{G}_0(\mathbf{r}_1, \mathbf{r}_2, \omega)\overline{\widehat{G}_0(\mathbf{r}_3, \mathbf{r}_4, \omega)}e^{-\frac{k^2 D|\mathbf{r}_c|}{6}\left[\rho_{24}^2 + \rho_{13}^2 + \boldsymbol{\rho}_{24}\cdot\boldsymbol{\rho}_{13}\right]}.$$

REFERENCES

F. Bailly, J.-F. Clouet, and J.-P. Fouque. Parabolic and white noise approximation for waves in random media. *SIAM J. Appl. Math.*, 56:1445–1470, 1996.

G. Bal, G. C. Papanicolaou, and L. Ryzhik. Self-averaging in time reversal for the parabolic wave equation. *Stochastics and Dynamics*, 2:507–531, 2002.

A. Bamberger, B. Engquist, L. Halpern, and P. Joly. Higher order paraxial wave equation approximations in heterogeneous media. *SIAM Journal on Applied Mathematics*, 48:129–154, 1988.

A. Bamberger, B. Engquist, L. Halpern, and P. Joly. Parabolic wave equation approximations in heterogeneous media. *SIAM Journal on Applied Mathematics*, 48:99–128, 1988.

J. Berryman, L. Borcea, G. C. Papanicolaou, and C. Tsogka. Statistically stable ultrasonic imaging in random media. *Journal of Acoustical Society of America*, 112:1509–1522, 2002.

G. Beylkin. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform. *J. Math. Phys.*, 26:99–108, 1985.

F. Billette and G. Lambaré. Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophys. J. Int.*, 135:671–680, 1998.

N. Bleistein. On the imaging of reflectors in the earth. *Geophysics*, 52:931–942, 1987.

P. Blomgren, G. C. Papanicolaou, and H. Zhao. Super-resolution in time-reversal acoustics. *Journal of the Acoustical Society of America*, 111:238–248, 2002.

L. Borcea. Imaging and time reversal in random media. Lecture notes, 2003.

L. Borcea, G. C. Papanicolaou, and C. Tsogka. Imaging the reflectivity of extended scatterers in random media. *preprint*, 2003.

L. Borcea, G. C. Papanicolaou, and C. Tsogka. A resolution study for imaging and time reversal in random media. *To appear in Contemporary Math.*, 2003.

L. Borcea, G. C. Papanicolaou, C. Tsogka, and J. Berryman. Imaging and time reversal in random media. *Inverse Problems*, 18(5):1247–1279, 2002.

J. Carazzone and W. W. Symes. Velocity inversion by differential semblance optimization. *Geophysics*, 56(5):654–663, 1991.

H. Chauris. *Analyse de vitesse par migration pur l'imagerie des structures complexes en sismique réflexion.* PhD thesis, École des Mines de Paris, 2000.

H. Chauris and M. Noble. Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications synthetic and real data sets. *Geophys. J. Int.*, 144:14–26, 2001.

H. Chauris, M. Noble, and P. Podvin. Testing the behaviour of differential semblance for velocity optimization. In *60th Conference EAGE*, Leipzig, 1998. European Association for Geoscientists and Engineers. Expanded Abstract.

L. A. Chernov. *Wave propagation in a random medium*. McGraw-Hill, New-York, 1960. Silverman, R. A. translator.

J. F. Claerbout. *Imaging the Earth's Interior*. Blackwell Scientific Publishers, Oxford, 1985.

A. Derode, P. Roux, and M. Fink. Robust acoustic time reversal with high-order multiple scattering. *Physical Review Letters*, 75:4206–4209, 1995.

D. R. Dowling and D. R. Jackson. Phase conjugation in underwater acoustics. *J. Acoust. Soc. Am.*, 89:171–181, 1990.

D. R. Dowling and D. R. Jackson. Narrow band performance of phase conjugate arrays in dynamic random media. *J. Acoust. Soc. Am.*, 91:3257–3277, 1992.

J. J. Duistermaat. *Fourier Integral Operators*. Birkhauser, New York, 1996.

M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, and M. Tanter. Time-reversed acoustics. *Rep. Prog. Phys.*, 63:1933–1994, 2000.

M. Gockenbach. *An abstract analysis of differential semblance optimization*. PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A., 1994.

L. Hörmander. *The Analysis of Linear Partial Differential Operators*, volume I-IV. Springer Verlag, Berlin, 1983.

A. Ishimaru. *Wave Propagation and Scattering in Random Media Vols. 1, 2*. IEEE/OUP Series on Electromagnetic Wave Theory. IEEE Press - Oxford University Press Classic Reissue, Oxford, 1997.

M. Jannane, W. Beydoun, E. Crase, D. Cao, Z. Koren, E. Landa, M. Mendes, A. Picas, M. Noble, G. Roeth, S. Singh, R. Snieder, A. Tarantola, D. Trezeguet, and M. Xie. Wavelengths of earth structures that can be resolved from seismic reflection data. *Geophysics*, 54(7):906–910, July 1989.

W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson. Phase conjugation in the ocean : Experimental demonstration of an acoustic time reversal mirror. *J. Acoust. Soc. Am.*, 103:25–40, 1998.

P. Lailly. The seismic inverse problem as a sequence of before-stack migrations. In J.B. Bednar et al., editors, *Conference on Inverse Scattering: Theory and Applications*, pages 206–220. SIAM, Philadelphia, 1983.

W. Mulder and A. ten Kroode. Automatic velocity analysis by differential semblance optimization. In *Expanded Abstracts, Society of Exploration Geophysicists, 71st Annual International Meeting*, Tulsa, 2001. SEG.

B. Nair and B. S. White. High-frequency wave propagation in random media - a unified approach. *SIAM J. Appl. Math.*, 51(2):374–411, 1991.

R. O'Dougherty and N. Anstey. Reflections on amplitudes. *Geophysical Prospecting*, 19:430–458, 1971.

G. C. Papanicolaou, L. Ryzhik, and K. Solna. The parabolic wave approximation and time reversal. *Matematica Contemporanea, in press*, 2002.

G. C. Papanicolaou, L. Ryzhik, and K. Solna. Statistical stability in time reversal. *To appear in the SIAM J. on Appl. Math.*, 2003.

C. Prada, S. Manneville, D. Spolianski, and M. Fink. Decomposition of the time reversal operator: Detection and selective focusing on two scatterers. *J. Acoust. Soc. Am.*, 99:2067–2076, 1996.

C. Prada, J. L. Thomas, and M. Fink. The iterative time reversal mirror: Analysis of convergance. *J. Acoust. Soc. Am.*, 97:62–71, 1995.

C. Prada, F. Wu, and M Fink. The iterative time reversal mirror: A solution to self-focusing in the pulse echo mode. *J. Acoust. Soc. Am.*, 90:1119–1129, 1991.

Rakesh. A linearized inverse problem for the wave equation. *Comm. on P.D.E.*, 13(5):573–601, 1988.

S. M. Rytov, Yu. A. Kravtsov, and V. I. Tatarskii. *Introduction to Statistical Radiophysics Vol. 4 Wave Propagation Through Random Media*. Springer-Verlag, Berlin, 1988.

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

K. Sobczyk. *Stochastic wave propagation*. Elsevier, Warsow, 1985.

H. C. Song. *On a transmission inverse problem*. PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A., 1994.

H. C. Song, W. A. Kuperman, and W. S. Hodgkiss. Iterative time reversal in the ocean. *J. Acoust. Soc. Am.*, 105:3176–3184, 1999.

C. Stolk. *On the modeling and inversion of seismic data*. PhD thesis, Universiteit Utrecht, 2000.

C. C. Stolk. On the stationary points of the seismic reflection tomography and differential semblance functionals in laterally homogeneous media. *Inverse problems*, 18:253–264, 2002.

C. C. Stolk and W. W. Symes. Smooth objective functionals for seismic velocity inversion. *Inverse Problems*, 19:73–89, 2003.

W. W. Symes. Mathematical foundations of reflected wave imaging. Technical Report 2, Rice University, February 1990.

W. W. Symes. A differential semblance algorithm for the inverse problem of reflection seismology. *Comput. Math. Appl.*, 22:147–178, 1991.

W. W. Symes. A differential semblance criterion for inversion of multioffset seismic reflection data. *J. Geoph. Res.*, 98:2061–2073, 1993.

W. W. Symes. High frequency asymptotics, differential semblance, and velocity analysis. In *Expanded Abstracts, Society of Exploration Geophysicists, 68th Annual International Meeting*, pages xx–xx, Tulsa, 1998. SEG.

W. W. Symes. All stationary points of differential semblance are asymptotic global minimizers: layered acoustics. In *The Rice Inversion Project: Annual Report 1999*, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005-1892, 1999. also appeared in SEP-95.

W. W. Symes and E. Lansing. A trace theorem for solutions of the wave equation, and the remote determination of acoustic sources. *Math. Meth. in the Appl. Sci.*, 5:131–152, 1983.

F. D. Tappert. The parabolic equation method. In Keller and Papadakis, editors, *Wave Propagation and Underwater Acoustics*. Springer Verlag, New York, 1977.

A. Tarantola. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49:1259–1266, 1984.

A. P. E. tenKroode, D. J. Smit, and A. R. Verdel. A microlocal analysis of migration. *Wave Motion*, 28:149–172, 1998.

B. J. Uscinski. *The elements of wave propagation in random media*. McGraw-Hill, New York, 1979.

B. S. White. The stochastic caustic. *SIAM J. Appl. Math.*, 44(1):127–149, 1984.

N. Winslow. Joint inversion using the convolutional model. Technical report, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA, 2000. (MA thesis).