Spectral Analysis of Randomly Scattered Signals using the Wavelet Transform *

J.F. Clouet and J.P. Fouque
Centre de Mathématiques Appliquées
Ecole Polytechnique
91128 Palaiseau Cedex

M. Postel
Laboratoire d'Analyse Numérique
Université Pierre et Marie Curie
75252 Paris Cedex 05

Abstract

Estimation of power spectra is a central question in signal analysis and its applications such as Geophysics. This is classical for stationary signals for which the power spectral density, i.e. the Fourier transform of the covariance, is defined as a function of the frequency only. In the general case, signals under study are not stationary. This work has been motivated by signals obtained in the context of wave propagating in random media. A one-dimensional model of random media is given by a layered medium whose properties are random variables in each layer independent from one layer to another. It has been shown ,[1], that if a pulse of typical wavelengths ϵ is sent in such a medium with small layers of order ϵ^2, the reflected signal is a non-stationary stochastic process which converges in distribution, as ϵ goes to 0, to a locally stationary gaussian process. This convergence will be the key property in the estimation of the power spectrum for the reflected signal. We introduce an estimation of the local power spectrum using wavelet transform. In our main theoretical result we identify the coefficients to be computed and we present the corresponding algorithm with numerical experiments. In the last section we show how this wavelet method enables us to get a fairly good estimation of the small parameter ϵ when this one is not known a priori.

*Spectral Analysis of Randomly Scattered Signals

1 Introduction

We will summarize briefly the mathematical results detailed in [1]. Their scope is a class of reflected signals propagated in a medium satisfying two main hypotheses:

- The medium is layered.
- The incident pulse is broad compared to the typical width of the layers but narrow compared to the large scale variations in the properties of the scattering half space.

If, in dimensionless variables, macroscopic variations of the medium properties occur on an order one scale then the pulse width is of order ϵ, which is a small positive parameter, and the size of the layers is of order ϵ^2 which is much smaller than ϵ. A typical situation is given by an acoustic signal entering a random slab $[0, L]$; the acoustic equations are:

$$\left\{ \begin{array}{l} \rho(x)\frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} = 0 \\ \frac{1}{K(x)}\frac{\partial p}{\partial t} + \frac{\partial u}{\partial x} = 0 \end{array} \right. \tag{1}$$

where $u = u(x,t)$ is the velocity, $p = p(x,t)$ is the pressure, $\rho(x)$ is the density and $K(x)$ the bulk modulus; we assume that the coefficients $\rho(x)$ and $K(x)$ are fluctuating around macroscopic smooth functions $\rho_0(x)$ and $K_0(x)$ in such a way that:

$$\left\{ \begin{array}{l} \rho(x) = \rho_0(x)(1 + \eta(\frac{x}{\epsilon^2})) \\ \frac{1}{K(x)} = \frac{1}{K_0(x)})(1 + \nu(\frac{x}{\epsilon^2})) \end{array} \right. \tag{2}$$

where $\eta(\frac{x}{\epsilon^2})$ and $\nu(\frac{x}{\epsilon^2})$ are centered ergodic processes varying on the small scale ϵ^2 corresponding to the size of the inhomogeneities of the medium. We suppose that $|\eta(\frac{x}{\epsilon^2})|$ and $|\nu(\frac{x}{\epsilon^2})|$ are bounded by a constant less than 1. Notice that we do not assume that the size of the fluctuations is small.

We consider the solution (u, p) of (1), we assume that $\rho(x)$ and $K(x)$ are constant outside of $[0, L]$ and we impose a continuity condition at $x = 0$ and $x = L$. We consider the following boundary conditions:

$$\left\{ \begin{array}{ll} A_\epsilon(x,t) \equiv u(x,t) + \frac{p(x,t)}{\sqrt{\rho_0(x)K_0(x)}} & A_\epsilon(0,t) = \frac{1}{\sqrt{\epsilon}} f(\frac{t}{\epsilon}) \\ B_\epsilon(x,t) \equiv u(x,t) - \frac{p(x,t)}{\sqrt{\rho_0(x)K_0(x)}} & B_\epsilon(L,t) = 0 \end{array} \right. \tag{3}$$

where A_ϵ is the wave traveling to the right, B_ϵ is the wave traveling to the left and f is a smooth function with compact support representing the entering signal at $x = 0$ on the scale ϵ, the typical wavelength, with a normalizing factor $\frac{1}{\sqrt{\epsilon}}$ to keep a constant energy entering the medium. The quantity $B_\epsilon(0,t)$

is then the random reflected signal. By the hyperbolicity of (1), for L large enough, $B_\epsilon(0,t)$ for $0 \leq t \leq T$ is independent of L.

It has been shown in [1] that $R^\epsilon_{f,t}(\sigma) \equiv B_\epsilon(0, t+\epsilon\sigma)$ converges in distribution, as a process in σ, to a centered stationary gaussian process $R_{f,t}(\sigma)$. Defining the windowed correlation function by:

$$\begin{aligned} C^\epsilon_f(t,\sigma) &= \mathbb{E}\{B_\epsilon(0,t-\epsilon\tfrac{\sigma}{2})B_\epsilon(0,t+\epsilon\tfrac{\sigma}{2})\} \\ &= \mathbb{E}\{R^\epsilon_{f,t}(-\tfrac{\sigma}{2})R^\epsilon_{f,t}(+\tfrac{\sigma}{2})\} \end{aligned} \tag{4}$$

the previous result implies that $C^\epsilon_f(t,\sigma)$ converges to the correlation function $C_f(t,\sigma)$ of the limiting process $R_{f,t}(\sigma)$. Then:

$$\lim_{\epsilon\to 0} C^\epsilon_f(t,\sigma) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\imath\omega\sigma}|\hat{f}(\omega)|^2\Lambda(\omega,t)d\omega \tag{5}$$

where $\hat{f}(\omega)$ is the Fourier transform of the incident pulse and $\Lambda(\omega,t)$ is the local power spectral density; the convergence in [2] is pointwise in σ but only weakly in t which means, [3], that (5) holds when both sides are multiplied by a smooth function of compact support and integrated over t. The local power spectrum density $\Lambda(\omega,t)$ depends on the macroscopic speed of propagation $c_0(x) = \sqrt{\frac{K_0(x)}{\rho_0(x)}}$ and on the statistical properties of the medium through the non-negative coefficient α defined by:

$$\alpha = \frac{1}{4}\int_0^\infty \mathbb{E}\{(\eta(0)-\nu(0))(\eta(x)-\nu(x))\}dx \tag{6}$$

In general $\Lambda(\omega,t)$ is difficult to compute explicitly ; for its determination a system of partial differential equations has to be solved [1]. This quantity is particularly interesting in the inverse problem which consists in reconstructing $c_0(x)$ from the observation of the reflected signal (see again [1] for details). Its statistical estimation is therefore essential. In what follows we propose an estimation of $\Lambda(\omega,t)$ based on a wavelet transform method. Fortunately, in the homogeneous case which corresponds to $\rho_0(x)$ and $K_0(x)$ constant, we have an explicit form:

$$\Lambda(\omega,t) = \frac{\alpha\omega^2}{(1+\alpha\omega^2 t)^2} \tag{7}$$

This situation will be used in our numerical experiments to compare our estimation with the true value of $\Lambda(\omega,t)$.
From (4) and (5) we are led to consider the expression

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega\sigma} C_f^\epsilon(t,\sigma)d\sigma \tag{8}$$

as an approximation for $|\hat{f}(\omega)|^2\Lambda(\omega,t)$ for ϵ small. We note first that although (8) is well defined for $\epsilon > 0$ it is not a nonnegative function. This is well known in signal analysis [4]. Recall that in the stationary case the positiveness of the power spectrum can be shown through the equivalence between its definition

$$\Lambda_{stat}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega\sigma} C_f^\epsilon(\sigma)d\sigma \tag{9}$$

which depends only on ω and the periodogram approximation

This is not true anymore in the non-stationary case. Therefore the estimation of (8) is an open problem. Second, if we try to use (8) as an estimate for $|\hat{f}(\omega)|^2\Lambda(\omega,t)$ it will depend sensitively on how we truncate the infinite Fourier integral. With enough experimenting with the data we can make a suitable version of (8) which provides useful estimates for $\Lambda(\omega,t)$. This was done in [5] as well as in [6] and [7] with data generated by direct numerical simulation of the scattering problem. The question is then: what is a good way to estimate $\Lambda(\omega,t)$ from $R_f^\epsilon(t)$ when we only know the parameter ϵ? By good we mean primarily robust, that is, by an algorithm that does not need fine tuning each time we look at a different set of data. Of course the estimation must also be unbiased. In [1](appendix E) a method to estimate $\Lambda(\omega,t)$, using the local or windowed Fourier transform, has been described and tested.
Using a wavelet transform we are able to identify the coefficients to be computed to obtain an estimation of $\Lambda(\omega,t)$. This is done in the next section.
In the following sections we describe the algorithm for the computation of these coefficients, we explain the numerical implementation and finally we give numerical experiments in the homogeneous case (7). In the last section we show that our method is also efficient when one wants to estimate parameters ϵ and α when these are not known a priori.

2 Wavelet Transform and Convergence Result

A wavelet basis is a tight frame $(g_{mn}(x))_{m,n\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ defined by the dilations and translations of a mother wavelet g:

$$g_{mn}(x) = \sqrt{\frac{1}{a_0^m}}\, g\left(\frac{x}{a_0^m} - nb_0\right) \tag{10}$$

where $a_0 > 1$ and $b_0 > 0$. We will use a wavelet described (see [8]) by its Fourier transform.

Let $\nu(x)$ be a C^∞ function satisfying:

$$\nu(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 & \text{for } x \geq 1 \end{cases}$$
$$\nu(x) + \nu(1-x) = 1$$

Let $l = \frac{2\pi}{b_0(a_0^2-1)}$, we define:

$$\hat{h}(y) = \begin{cases} 0, & y \leq l \\ \sin\left[\frac{\pi}{2}\nu\left(\frac{y-l}{l(a_0-1)}\right)\right], & l \leq y \leq a_0 l \\ \cos\left[\frac{\pi}{2}\nu\left(\frac{y-a_0 l}{l a_0(a_0-1)}\right)\right], & a_0 l \leq y \leq a_0^2 l \\ 0, & y \geq a_0^2 l \end{cases}$$

and:

$$\hat{g}(y) = \sqrt{\frac{b_0}{2\pi}} e^{\imath \frac{b_0 y}{2}} \left[\hat{h}(y) + \hat{h}(-y)\right]. \tag{11}$$

Then $g(x) = \sqrt{\frac{1}{2\pi}} \int e^{-\imath xy} \hat{g}(y) dy$ is the mother wavelet of a basis of $L^2(\mathbb{R})$ and $\hat{g}$ satisfies the following normalization properties:

$$\begin{array}{ll} \hat{g} \text{ has compact support: } & \begin{array}{l} Supp(\hat{g}) = [-a_0^2 l, -l] \cup [l, a_0^2 l] \\ length(Supp(\hat{g})) = \frac{4\pi}{b_0} \end{array} \\ \int \frac{|\hat{g}(y)|^2}{|y|} dy = \frac{b_0 \log a_0}{\pi} & \\ \sum_{m\in\mathbb{Z}} |\hat{g}(a_0^m y)|^2 = \frac{b_0}{2\pi} \mathbb{I}_{y\neq 0} & \end{array} \tag{12}$$

The main property of the wavelet transform is that the reflected signal R_f^ϵ can be recovered from its local wavelet coefficients:

$$R_f^\epsilon(x) = \sum_{n,m\in\mathbb{Z}} R_{nm}^\epsilon g_{nm}(x) \tag{13}$$

with

$$R_{nm}^\epsilon = \int R_f^\epsilon(t) g_{nm}(t)^* dt \tag{14}$$

and we can express the local power spectral density in term of the coefficients. This spectral density is defined as:

$$\Lambda_f(t,\omega) = \lim_{\epsilon\to 0} \frac{1}{\sqrt{2\pi}} \int e^{-\imath\omega\sigma} C_f^\epsilon(t,\sigma) d\sigma \tag{15}$$

where $C_f^\epsilon(t,\sigma)$ is the local correlation function:

$$C_f^\epsilon(t,\sigma) = \mathbb{E}\{R_f^\epsilon(t - \frac{\epsilon\sigma}{2}) R_f^\epsilon(t + \frac{\epsilon\sigma}{2})^*\} \tag{16}$$

let:

$$\Lambda_f^\epsilon(t,\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-\imath\omega\sigma} C_f^\epsilon(t,\sigma) d\sigma \tag{17}$$

Therefore a naive way to estimate the density would be to compute the limit:

$$\Lambda_f^\epsilon(t,\omega) = \lim_{\epsilon\to 0} \sum_{n,m,n',m'} \mathbb{E}\{R_{nm}^\epsilon R_{m'n'}^\epsilon\} \frac{1}{\sqrt{2\pi}} \int e^{-\imath\omega\sigma} g_{nm}(t-\frac{\epsilon\sigma}{2}) g_{n'm'}(t+\frac{\epsilon\sigma}{2})^* d\sigma$$

But we will prove the following result which enables us to compute in a simple way an estimate for the density:

Theorem

Assume that $\Lambda_f^\epsilon(t,\omega) \to_{\epsilon\to 0} \Lambda_f(t,\omega)$ uniformly in ω and t on compact sets where $\Lambda_f(t,\omega)$ is continuous in both its arguments, and that the pulse f is smooth enough. Then we have:

$$\Lambda_f(t,\omega) = \lim_{a_0\to 1} \lim_{\epsilon\to 0} \sum_n \sum_{m=M+1,M+2} \mathbb{E}\{|R_{nm}^\epsilon|^2\} G_{nm}(t,\omega) \tag{18}$$

where

$$G_{nm}(t,\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-\imath\omega\sigma} g_{nm}(t-\frac{\epsilon\sigma}{2}) g_{nm}(t+\frac{\epsilon\sigma}{2})^* d\sigma \tag{19}$$

and (for $\omega \neq 0$):

$$M = [\frac{\log \frac{l\epsilon}{|\omega|}}{\log a_0}]$$

Moreover, as will be seen later, $G_{nm}(t,\omega)$ is very well localized in n; so with ω held fixed, we will estimate the density by computing an average for $|R_{M+1,n}^\epsilon|^2$ and $|R_{M+2,n}^\epsilon|^2$ for a finite number of values of n and multiplying them with the fixed functions $G_{nm}(t,\omega)$, with a_0 close to 1 and ϵ close to zero. It is a relatively nonstandard argument to make the dilation coefficient a_0 converge to 1 so we will give a full proof of this theorem.

The proof is based on the Poisson summation formula and on the properties of $\hat{g}$ (12). It is detailed in the appendix and we now present the numerical implementation of the theorem, with simulations that actually illustrate the convergence of the algorithm.

3 Numerical Algorithm

It has just been shown that in the limit only two terms in m contribute in the sum (18). Hence the estimation of Λ requires the computation of only 2 wavelet coefficients R_{mn}. We first show how considerations of support compactness simplify the computation of the relevant functions of t and ω, G_{mn}. In particular we will present an efficient tabulation of those functions on a non rectangular grid. Then we will describe how the corresponding wavelet coefficients R_{mn} are computed.

3.1 Computation of $G_{mn}(t,\omega)$

Since what is known is the Fourier transform of $g(t)$ we prefer to the definition of the function G_{mn} (19) its equivalent form in the dual space (36) which we rewrite as

$$G_{mn}(\omega,t)=\frac{2}{\epsilon\sqrt{2\pi}}e^{-i\omega n b_0 2a_0^m/\epsilon}e^{2i\omega t/\epsilon}\int e^{-2inb_0\tilde{\omega}}e^{2ita_0^{-m}\tilde{\omega}}\hat{g}(-\tilde{\omega})\hat{g}^*(\omega 2a_0^m/\epsilon+\tilde{\omega})d\tilde{\omega}$$

which decomposes into four separate integrals due to the form of $\hat{g}$ (11)

$$\begin{aligned}G_{mn}(\omega,t) &= \frac{2b_0}{\sqrt{(2\pi)^3}}e^{-i\omega b_0 a_0^m(2n+1)/\epsilon}e^{2i\omega t/\epsilon}\\ & \quad \Big\{\int e^{-ib_0\tilde{\omega}(2n+1)}e^{2ita_0^{-m}\tilde{\omega}}\hat{h}(-\tilde{\omega})\hat{h}(2\omega a_0^m/\epsilon+\tilde{\omega})d\tilde{\omega}\\ &+ \int e^{-ib_0\tilde{\omega}(2n+1)}e^{2ita_0^{-m}\tilde{\omega}}\hat{h}(\tilde{\omega})\hat{h}(2\omega a_0^m/\epsilon+\tilde{\omega})d\tilde{\omega}\\ &+ \int e^{-ib_0\tilde{\omega}(2n+1)}e^{2ita_0^{-m}\tilde{\omega}}\hat{h}(-\tilde{\omega})\hat{h}(-2\omega a_0^m/\epsilon-\tilde{\omega})d\tilde{\omega}\\ &+ \int e^{-ib_0\tilde{\omega}(2n+1)}e^{2ita_0^{-m}\tilde{\omega}}\hat{h}(\tilde{\omega})\hat{h}(-2\omega a_0^m/\epsilon-\tilde{\omega})d\tilde{\omega}\Big\} \end{aligned} \tag{20}$$

Let us write the frequency ω where we are computing the spectrum as intermediate between two consecutive frequencies M and $M+1$ on the irregular "grid" $\omega=a_0^{-m}l\epsilon$.

$$a_0^{-M-1}l\epsilon \;<\; \omega \;<\; a_0^{-M}l\epsilon \tag{21}$$

This can be done by introducing a coefficient $1<a<a_0$ such that

$$\omega=a_0^{-M}l\epsilon/a \tag{22}$$

What are the values of m such that G_{mn} will enter summation (18) ? This happens when functions $\hat{g}(-\tilde{\omega})$ and $\hat{g}^*(2la_0^{m-M}/a+\tilde{\omega})$ have overlapping supports. We know from (38) that 2 values of m only contributes in the limit: $M+1$ and $M+2$. Furthermore due to support properties of $\hat{h}$ only one cross-product in the equation above is non-zero for $m=M+1$ and $m=M+2$. This leads to the following final expression of G_{mn} after centering of the integral at zero in order to use standard FFT algorithm

$$\begin{aligned} G_{M+1,n}(l\epsilon a_0^{-M}/a,t) &= \frac{b_0}{\epsilon\sqrt{(2\pi)^3}} e^{il(a_0^2+1-2a_0/a)(b_0(2n+1)/2-ta_0^{-M-1})} \int_{(1-a_0^2)l/2}^{(a_0^2-1)l/2} e^{-ib_0\tilde{\omega}(2n+1)} \\ & e^{2it\tilde{\omega}a_0^{-M-1}} \hat{h}(-\tilde{\omega}+(a_0^2+1)l/2)\ \hat{h}(\tilde{\omega}+2a_0 l/a-(a_0^2+1)l/2)\ d\tilde{\omega} \end{aligned}$$

and

$$\begin{aligned} G_{M+2,n}(l\epsilon a_0^{-M}/a,t) &= \frac{b_0}{\epsilon\sqrt{(2\pi)^3}} e^{il(a_0^2(1-2/a)+1)(b_0(2n+1)/2-ta_0^{-M-2})} \int_{(1-a_0^2)l/2}^{(a_0^2-1)l/2} e^{-ib_0\tilde{\omega}(2n+1)} \\ & e^{2it\tilde{\omega}a_0^{-M-2}} \hat{h}(-\tilde{\omega}+(a_0^2+1)l/2)\ \hat{h}(\tilde{\omega}+2a_0^2 l/a-(a_0^2+1)l/2)\ d\tilde{\omega} \end{aligned}$$

Recall the rules to use the FFT algorithm: let the discrete samples of a frequency series $\hat{s}$ be noted by $\hat{s}_j = \hat{s}(j\Delta\omega)$ for $-N/2+1 \leq j \leq N/2$. Under hypothesis of regularity for $\hat{s}$ and if $](-N/2+1)\Delta\omega, N\Delta\omega/2[\ \in \ Supp\{\hat{s}\}$ its inverse Fourier Transform f can be estimated at time $t_k = k\Delta t$ with $N\Delta t\Delta\omega = 2\pi$ by

$$s_k = s(t_k) = \frac{\Delta\omega}{2\pi} \sum_{j=-N/2+1}^{N/2} \hat{s}_j e^{-ikj\frac{2\pi}{N}} \tag{23}$$

The numerical error thus introduced can be neglected if we assume sufficient regularity of $\hat{s}(\omega)$. Support compactness and regularity of $\hat{g}$ ensure (see [9]) the exponential convergence of the Finite Fourier Transform towards the continuous Fourier Transform in the domain of analyticity of the integrand.

At this stage, to get N samples in time domain (index n) we sample $\hat{h}$ by $2N$ samples of sampling rate $\Delta\omega$ such that $2N\Delta\omega = (a_0^2-1)l$ - which amounts to $\Delta t = b_0$ - and we use only the odd samples of the FFT output. In fact there is only one FFT to be performed for $t=0$ and $G_{M+1,n}$ for all t can be deduced from it by translation: let

$$\begin{aligned} G_n^1 = \ & \tfrac{b_0}{\epsilon\sqrt{(2\pi)^3}} e^{ib_0 l(a_0^2+1-2a_0/a)(2(n-N)+1)/2} \cdot \\ & \cdot \textstyle\sum_{k=-N+1}^{N} e^{-i\frac{\pi}{N}kn} \hat{h}(-k\Delta\omega+(a_0^2+1)l/2)\ \hat{h}(k\Delta\omega+2a_0 l/a-(a_0^2+1)l/2) \end{aligned} \tag{24}$$

We get for any $t = Nb_0 a_0^{M+1}$

$$G_{M+1,n}(\epsilon l a_0^{-M}/a, Nb_0 a_0^{M+1}) = G^1_{2(n-N)+1} \tag{25}$$

If we use this method to compute $G_{M+1,n}$ it cannot be used for the other function $G_{M+2,n}$ since the time sampling would have to be different. Let the series G^2 be computed by the FFT algorithm:

$$\begin{aligned} G_n^2 &= \frac{b_0}{\epsilon\sqrt{(2\pi)^3}} e^{ib_0 l(a_0^2(1-2/a)+1)n/2} \sum_{k=-N+1}^{N} e^{-i\frac{\pi}{N}kn} \\ &\quad \hat{h}(-k\Delta\omega + (a_0^2+1)l/2)\ \hat{h}(k\Delta\omega + 2a_0^2 l/a - (a_0^2+1)l/2) \end{aligned} \tag{26}$$

The value of the function $G_{M+2,n}$ at any time $t = Nb_0a_0^{M+1}$ is computed by linear interpolation of the function at the 2 nearest samples

$$G_{M+2,n}(\epsilon l a_0^{-M}/a, Nb_0a_0^{M+1}) = G^2_{2(n-N/a_0)+1} \tag{27}$$

Functions G_{mn} are represented on figure (1) at time $t = 0$ as a function of n for two values of the frequency ω. The first one $\omega = 10$ is exactly on the irregular grid (28) for $a_0 = 1.1168$, $\epsilon = 0.075$, $b_0 = 1$ and $M = -17$. The second one $\omega = 10.58$ corresponds to a value of $a = 1.0556$ intermediate between two grid points. For $\omega = 10$ only the coefficient $m = M+1 = -16$ contributes and $G_{-15,n}(10,0) = 0$ while for $\omega = 10.58$ both terms contribute in about the same proportion and $G_{-16,n}(10.58,0)$ and $G_{-15,n}(10.58,0)$ have similar shape and height. For both frequencies G_{mn} is concentrated around the origin. This illustrates the fast decrease with respect of the index n and will justify the numerical evaluation of (18) with a finite summation over n. Note that if (ω, t) happens to be exactly on the irregular grid: figure 1

$$\left(\omega = \frac{2\pi\epsilon}{b_0(a_0^2-1)a_0^M}, t = Nb_0a_0^{M+1}\right) \tag{28}$$

the spectrum can be estimated using only one wavelet coefficient because $G_{M+2,n}$ is identically zero and function $G_{M+1,n}$ simplifies for $a = 1$

$$G_{M+1,n}(\epsilon l a_0^{-M}, Nb_0a_0^{M+1}) = G^1_{2(n-N)+1} \tag{29}$$

with

$$
\begin{aligned}
G_n^1 = \; & \frac{b_0}{\epsilon\sqrt{(2\pi)^3}} e^{i2\pi\frac{a_0-1}{a_0+1}(2(n-N)+1)/2} \cdot \\
& \sum_{k=-N+1}^{N} e^{-i\frac{\pi}{N}kn} \hat{h}(-k\Delta\omega + (a_0^2+1)l/2) \; \hat{h}(k\Delta\omega + 2a_0 l - (a_0^2+1)l/2)
\end{aligned} \tag{30}
$$

We would like to use this simplification whenever possible. However we will illustrate with numerical simulations that the averaging introduced by summing over 2 wavelet coefficients instead of using only $R^\epsilon_{M+1,n}$ lowers significantly the fluctuations, especially when there are only a few realizations over which the statistical expectation of $|R^\epsilon_{mn}|^2$ can be performed. This seems logical since the distance $\Delta\omega$ between two consecutive frequencies on the grid $l\epsilon a_0^m$ is of order 1. Actually the theory tells us that random reflection coefficients at frequencies spaced apart at least one unit on the order one scale are uncorrelated; hence we can expect corresponding wavelet coefficients to be uncorrelated.

3.2 Computation of the wavelet coefficient R^ϵ_{mn}

Here again we start from the definition of the wavelet coefficient (14) and transform it to use the form $\hat{g}(\omega)$ that is known explicitly rather than $g(t)$. Now we use the definition of R^ϵ_f which is specific to the set-up described in the introduction and is introduced in section 2 of [1] for instance.

$$
R^\epsilon_f(t) = \frac{1}{2\pi\epsilon}\int e^{-i\omega t/\epsilon}\hat{R}^\epsilon(\omega)\hat{f}(\omega)d\omega \tag{31}
$$

where $\hat{R}^\epsilon(\omega)$ is the reflection coefficient at $z=0$ and at frequency ω/ϵ for a randomly layered medium where the layer thickness is ϵ^2. For clarity we let $\hat{R}(\omega) = \hat{R}^\epsilon(\epsilon\omega)$ though it still depends on ϵ through the layer thickness. We use this expression in (14) and use the fact that $R^\epsilon_f(t)$, $f(t)$ and $\hat{h}$ are real . Since the support of $\hat{h}$ is $[l, a_0^2 l]$ and we want to use standard FFT algorithms we need to translate and center it at 0. We get eventually

$$
\begin{aligned}
R^\epsilon_{mn} \; = \; & \sqrt{\frac{2b_0}{2\pi a_0^m}}\mathrm{Real}[e^{-ib_0\frac{a_0^2+1}{2}l(n+\frac{1}{2})}\sum_{k=-N/2+1}^{N/2} e^{-i\frac{2\pi}{N}k(n+1/2)} \\
& \hat{R}((k\Delta\omega + \tfrac{a_0^2+1}{2}l)a_0^{-m})\hat{f}((k\Delta\omega + \tfrac{a_0^2+1}{2}l)a_0^{-m}\epsilon)\hat{h}(k\Delta\omega + \tfrac{a_0^2+1}{2}l)]
\end{aligned} \tag{32}
$$

In (32) we chose to dilate the frequency ω by ωa_0^m so that we will always use the same window $\hat{h}$ and will do the dilation in the sampling of the signal $\hat{R}$.

We see that to compute R^{ϵ}_{mn} by a standard FFT algorithm we need to sample $\hat{F}(\omega)$ and $\hat{h}(\omega)$ on the interval $]\frac{-a_0^2+1}{2}l, \frac{a_0^2-1}{2}l[$ with N samples so that $2\pi = N\Delta\omega b_0$ and $N\Delta\omega = (a_0^2-1)l$, which is the same condition since $2\pi = (a_0^2-1)lb_0$. Note that after numerical FFT we will have $N/2$ samples for negative values of n

The sampling of $\hat{F}$ at rate $\Delta\omega$ requires the resampling of the signal $\hat{R}(\omega)$ from the recording sampling rate h_0 to the desired sampling for the current wavelet coefficient - that is $\Delta\omega a_0^{-m}$.

This is done by linear interpolation which introduces an additional source of error of order h_0^2. An alternative method would be to perform the FFT with the actual sampling of the data. The number of samples has to be a power of two and must also be chosen so that the support of the window is larger than the support of $\hat{h}$ which will be prolonged with zeros. As long as the ratio of h_0 over $\Delta\omega a_0^{-m}$ is smaller than one the first method allows us to use fewer points in the numerical FFT with the same quality of results. For ratio of the sampling rates greater than one the same method can still be used but with a higher order interpolation method. Numerical simulations indicate that interpolation of $\hat{R}$ over 6 points is necessary if $h_0 = 2\Delta\omega a_0^{-m}$.

3.3 Numerical implementation

Before turning to the numerical simulations we summarize the algorithm: in all the following simulations we assume that we know the scale factor ϵ.

- Randomly reflected signals are available in the frequency domain at a given sampling rate h_0 - This will be detailed in the next subsection.
- Estimation of the power spectrum is sought at a given frequency ω. Let the value of b_0 be fixed to one for instance, and a_0 to a value slightly above $1+\epsilon$. Inequality (21) determines the two dilation indexes $M+1$ and $M+2$ that will enter in the estimation. The position of the frequency in the grid is obtained through the parameter a (22)
- Compute the series G^0_n and G^1_n using (24) and (26)
- For each independent realization of the random medium compute the two wavelet coefficients $R^{\epsilon}_{M+1,n}$ and $R^{\epsilon}_{M+2,n}$ and average the square of their absolute value
- Series $\mathbb{E}\{|R^{\epsilon}_{mn}|^2\}$ are now available for a finite number N_{max} of indexes n. For each time $t = Nb_0a_0^{M+1}$ with $N \leq N_{max}$ perform the numerical analogue of summation (18)

$$\Lambda(\omega, Nb_0a_0^{M+1}) = \sum_{n=0}^{N_{max}} \left[\mathbb{E}\{|R^{\epsilon}_{M+1,n}|^2\}G^0_{2(n-N)+1} + \mathbb{E}\{|R^{\epsilon}_{M+2,n}|^2\}G^0_{2(n-N/a_0)+1}\right] \quad (33)$$

4 Numerical simulations

4.1 Description of the data used to compute estimations

In this section we describe actual numerical simulations that have been performed to illustrate the method described above. First of all a set of randomly scattered acoustic signals was generated using the layer code described in [1]. Because it is much faster to generate and hence allows us to get more redundancy for the same cost we treat only plane wave signals. Furthermore, since we are mainly interested in the convergence of the algorithm we use only homogeneous backgrounds where we have an analytical expression for the power spectrum (7) where α measures the intensity of the randomness of the medium. In the simulations below, the random medium is a succession of layers of constant thickness ϵ^2. ρ, the density, equals 1 everywhere. K, the bulk modulus, is constant within each layer and takes random values from one layer to the other such that $1/K$ is uniformly distributed between 0.7 and 1.3. This set-up leads to a value of $\alpha = 0.3^2/24 = 0.00375$.

The source spectrum $\hat{f}(\omega)$ is a Gaussian pulse. The fact that the signals are generated directly in the frequency domain is certainly an artificial numerical advantage since it saves an FFT - both in terms of CPU and precision. But it should not be considered essential. If time signals are to be treated, only step (31) would be interposed at this stage. In the geophysical application that we have in mind this will require an inverse Fourier Transform of the time series recorded by geophones.

In the simulation we propagate an acoustic wave through a pile of thin layers whose bulk modulus take random values that are uniformly distributed and independent from one layer to the other. We can repeat this with as many independent realizations of the random properties as necessary. In practice, the assumption that the medium properties are random variables is only a tool, or rather - there will be only one realization of such a medium. This difficulty cannot be overcome in the estimation of the spectrum. However in the final aim, which is to recover the average properties of the medium by comparing the spectral estimates with their theoretical mean, it will be possible to replace the statistical averaging by the redundancy of information provided by the lateral space dependence available if a point source is used instead of a plane wave.

4.2 Numerical Behaviour of the algorithm

From the description above we see that several causes will account for the discrepancies between the estimates and their theoretical mean beside the fact that the averaging is of course done on a finite number of random realizations of the medium. We will first illustrate the numerical convergence of the algorithm, *i.e.* is it justified to use (33) for $a_0 \neq 1$ and $\epsilon \neq 0$ and do the estimations thus obtained get closer to their theoretical value when a_0 and ϵ get closer to their limit? Then we would like to study the influence of the position of the frequency ω where the spectrum is estimated within the irregular grid (28). When ω is exactly on the grid only one wavelet coefficient is used. When ω is intermediate, two consecutive wavelet coefficients enter into the summation and we shall illustrate numerically that it suppresses a large part of the fluctuations.

Eventually we present numerical experiments showing that it is possible to estimate the scale factor ϵ from the data. These results rely on the robustness of the algorithm with respect to the other parameters - a_0 in particular- and in general require the same framework that the plane wave inversion method used to recover the profile $\alpha/c_0(\tau)$ [10],[1].

4.2.1 Convergence as $a_0 \to 1_+$ and $\epsilon \to 0$

To study independently the factors described above we do the following numerical experiments: the spectrum $\Lambda(\omega, t)$ is estimated for values of (ω, t) exactly on the nonrectangular grid (28) (only $M+1, n$ terms enter the summation (18)). The random signals are generated with frequency sampling rate h_0 exactly equal to $d\omega a_0^m$ so that no resampling is needed to perform (32). In the first set of experiments the influence of the wavelet parameter a_0 is studied: $\hat{\Lambda}$ is estimated at frequency $\omega = 10$ using 300 realizations of a randomly reflected signal on a stack of 3000 layers of thickness $\epsilon^2 = 0.03^2 = 0.0009$. Six different values of a_0 are used, chosen so that the equation

$$\omega = l\epsilon a_0^{-M} = \frac{2\pi\epsilon a_0^{-M}}{b_0(a_0^2-1)} \tag{34}$$

has a solution in M in the set of integers. Figure (2) shows the continuous variation of the wavelet index of dilation M as a function of a_0 on the left and the corresponding sampling rate of $\hat{R}(\omega)$ used in the FFT that appears in the computation of $R^\epsilon_{M+1,n}$ on the right. The dots on the curves correspond to the 6 sets of parameters (a_0, M, h_0) that have actually been tried. The sampling rate is normalized by ϵ. On the same figure the position of the center of the window $\hat{g}_{M+1,n}(\omega)$ used to compute $R^\epsilon_{M+1,n}$ is also represented. This center frequency is $\omega(a_0^2+1)/(2\epsilon a_0)$ which goes to ω/ϵ when $a_0 \to 1$. One can see that for all values tried we are already close to the limit value. Figure (3) compares the estimation of $\hat{\Lambda}$ done with the six different sets of parameters with the theoretical value given by (7). On each plot the spectrum at fixed frequency $\omega = 10$ is represented as a function of time. Remember that the sampling rate in time depends both on the frequency and the parameter a_0 through (28). As a_0 goes to 1, this sampling rate increases and there are less and less points on the curves. As a_0 goes away from 1, the sampling rate decreases and time points are closer to each other but the error also increases (figure (3)). One can see on the bottom right plot both large fluctuations and a deviation from the theoretical mean for small times. What is less clear is the behaviour as a_0 gets really close to one, past $1+\epsilon$ say, as is the case on the top left plot. Since there are fewer points on the curve it is hard to say if they do tend to depart from the theoretical curve, as they should since the error also has an $O(\epsilon)/(a_0-1)$ term. The time after which theory and estimations don't agree anymore on all plots correspond to the two way travel time to the bottom of the stack of random layers on which the reflected signals have been recorded. In our case it is around 2x3000x0.03^2 = 5.4. After this time the wave travels in homogeneous medium without being reflected thus the power spectrum should be zero. On the other hand the theoretical power spectrum formula (7) holds for infinite stack of random layers.

figure 2

figure 3

In the second experiment we have kept the ratio $\epsilon/(a_0-1)$ constant and equal to 0.75, and have made a_0 go to one - thus making ϵ go to zero. Figure (4) shows the continuous variation of the wavelet index of dilation M, on the left, along with the corresponding sampling rate h_0 required to compute coefficient $R^{\epsilon}_{M+1,n}$ normalized by $\epsilon = 0.75(a_0-1)$. The four dots correspond to the four sets of parameter actually tried. Two of them do not fall exactly on the curve because the values of ϵ actually used 0.03,0.075,0.1 and 0.15 could only roughly obey the rule $\epsilon = 0.75(a_0-1)$ with integer values of M.

figure 4

Figure (5) compares the estimations of power spectrum done with the four sets of parameters described above and their theoretical mean given by (7). They are represented at the fixed frequency $\omega = 10$ as functions of time. For each plot 300 different simulations of wave propagation were done, on stacks of random layers of different thicknesses ϵ^2, to perform the statistical averaging $< |R^{\epsilon}_{M+1,n}|^2 >$ in (18), (in the following mathematical expectation will also be denoted by brackets). In order to be able to compare the different plots down to t=10., 9000 layers were used for $\epsilon = 0.03$, but only 300 were necessary for $\epsilon = .15$. This choice of parameters illustrate the convergence as a_0 goes to 1 since both the fluctuations and the bias diminishes.

figure 5

These two sets of experiments show that we can reproduce numerically the behaviour predicted by the convergence result (18).

4.2.2 Estimation on and beside the grid

Estimations of the spectrum at (ω, t) exactly on the grid (28) can be obtained using only one wavelet coefficient. However the simplicity of the computation is counterbalanced by the fluctuations of the estimations thus obtained which are much higher than estimations for intermediate frequencies - where the contributions of two wavelet coefficients are averaged. It is this phenomenon that is illustrated by the following experiments: the data is the same as used in the bottom left plot of figure (5). The spectrum is estimated at three different frequencies 10., 10.58 and 11.168 which correspond to $\omega = l\epsilon a_0^{-}17/a$ with $a = 1$, $a = a_0 = 1.117$ and an intermediate value $a = 1.0556$. Two wavelet coefficients enter the computation (33) $R^{\epsilon}_{-16,n}$ and $R^{\epsilon}_{-15,n}$. On figure (6) the signal in the frequency domain multiplied by the wavelet $\hat{g}_{mn}$ is represented as a function of frequency for one of the 300 independent realizations. The data on the right hand side plot is used to compute coefficient $R^{\epsilon}_{-15,n}$. It is centered around the first target frequency 10, normalized by $\epsilon = 0.075$. On the left hand side plot the data used to compute the second coefficient $R^{\epsilon}_{-16,n}$ is represented and it is centered around the third target frequency $11.168/0.075 = 149$. A numerical FFT is performed on each independent realization of these two plots, in the manner described in (32). figure 6

In figure (7) the average of the squared modulus of R^{ϵ}_{mn} is represented as a function of index n. The main feature is that $|R^{\epsilon}_{-15,n}|$ and $|R^{\epsilon}_{-16,n}|$ are of comparable amplitude. It is those two series, $< |R^{\epsilon}_{-15,n}|^2 >$ and $< |R^{\epsilon}_{-16,n}|^2 >$, that will be combined with the functions $G_{-15,n}(\omega, t)$ and $G_{-16,n}(\omega, t)$ described in the previous section and represented on figure (1) to obtain the spectrum estimation. Due to the form of the function G, for the intermediate frequency 10.58 both coefficients enter into (33) while for $\omega = 10.$ only $R^{\epsilon}_{-15,n}$ will contribute and for $\omega = 11.168$ only $R^{\epsilon}_{-16,n}$. figure 7

In figure (8) the resulting estimations are represented with R^{ϵ}_{mn} averaged over 300 realizations on the left hand side plots and only 10 on the right hand side plots. When the statistical redundancy is small - as it is the case when 10 realizations are used, the averaging effect obtained by combining independent frequency coefficients is clearly illustrated. figure 8

5 Determination of the scaling factor ϵ

Here we want to illustrate one major advantage of our wavelet-based estimator over the estimator used before. This advantage is the clear definition of the window width as a function of the scaling factor ϵ through the parameter a_0. We have seen in the previous sections that in terms of bias, the estimator is quite stable with respect with the parameter a_0 even if the fluctuations do increase as a_0 departs from 1. We want to use this observation to determine the scaling factor ϵ as a parameter that minimizes a certain functional that we describe below. The final goal is to make the minimal assumptions on the model - namely we would like to recover the speed profile as a piece-wise linear function of depth, and the unknowns for the problem would be the discrete values of the speed plus the scaling factor ϵ. With the speed profile thus described, Λ may be determined by solving numerically the transport equation as mentioned earlier. In fact in this paper we are going to reduce the complexity of the method by assuming that the background is homogeneous and that besides ϵ, we are just looking for the noise intensity α that appears in formula (7). We emphasize this dependence of Λ on α by writing $\Lambda(t, \omega; \alpha)$.

To construct the maximum likelihood estimator of ϵ, α we need to know the probability density of $\hat{\Lambda}(t_j, \omega_l)$ when (ϵ, α) is known. We have chosen here not to use the fact that the reflection coefficient tends to a Gaussian process when ϵ goes to zero. In previous works on inversion [10] this was used to derive the exponential distribution for the estimations of the power spectrum. But in the present case we do not know that the exponential distribution holds when the estimations are computed using an arbitrary value for ϵ. Therefore we rather choose for the likelihood function the standard least squares type functional

$$F(\epsilon, \alpha) = C \;+\; \sum_{j=1}^{N_t} \sum_{l=1}^{N_f} (\hat{\Lambda}(t_j, \omega_l; \epsilon) - \Lambda(t_j, \omega_l; \alpha))^2 \;\; p_{j,l} \tag{35}$$

The weights $p_{j,l}$ in the expression above are the variances of the estimations $\hat{\Lambda}$ and are not necessarily the same for all times t_j and frequencies ω_l. In fact we know that if the true value of ϵ is used to compute $\hat{\Lambda}$ the variance will exactly equal the mean therefore strongly depend on t and ω. We choose to keep the same dependence here and arbitrarily set the $p_{j,l}$ equal to $1/\Lambda(t_j, \omega_l; \alpha)^2$.

The maximum likelihood estimator for α may therefore be constructed by minimizing the expression in (35) over ϵ, α. This may be accomplished by using a numerical optimization code with an objective function given by (35) to be minimized in two dimensions.

Here we would like to go into more details that will partly justify our choice of probability law for the power spectrum estimations: the frequency-time grid on which these estimations are computed depends both on the unknown ϵ and the computational parameter a_0. This last one being fixed, the grid is different for different trial values of ϵ. For a given data set characterized by a frequency sampling rate h_0 and a number N of samples $[i_1, i_N]$ the possible dilation indices will have to satisfy the following criteria:

- the sampling rate to compute the wavelet coefficient with (32) should be at least greater than the available sampling rate
  $M \leq \log(\frac{\Delta\omega}{h_0 a_0}) / \log(a_0)$

- the wavelet coefficient support should be within the data frequency range $[l, a_0^2 l] \in [a_0^{-M-1} i_1 h_0, a_0^{-M-1} i_N a_0]$ which leads to $\log(\frac{2\pi a_0}{b_0 h_0 i_N * (a_0^2-1)})/\log(a_0) \leq M \leq \log(\frac{2\pi}{b_0 h_0 a_0 i_1 * (a_0^2-1)})/\log(a_0)$

To each of the dilation indices contained within the bounds (M_{min}, M_{max}) defined above, (28) associates a frequency value which depends on ϵ ($\omega_M = l\epsilon a_0^{-M}$) and a time sampling rate ($\delta_t = b_0 a_0^{M+1}$).

To define precisely the objective function (35) constant - that is independent from the unknown ϵ - macroscopic frequency and time sampling rates must be chosen . They should be large enough so that consecutive frequencies or time estimates can be considered uncorrelated. The range of frequency $[\omega_{min}, \omega_{max}]$ and time $[0, t_{max}]$ over which the summations in (35) are to be performed should be also constant. This means that the range of the unknown ϵ has to be fixed *a priori* to $[\epsilon_{min}, \epsilon_{max}]$ and should be such that $[\omega_{min}, \omega_{max}] = [l\epsilon_{max}/a_0^{M_{min}}, l\epsilon_{min}/a_0^{M_{max}}] \neq \emptyset$

To summarize, the functional which has to be minimized with ϵ, and eventually with α, depends on several parameters: a_0, the possible scope of the unknown $[\epsilon_{min}, \epsilon_{max}]$, the macroscopic frequency and time steps. We now present a few numerical experiments that represent the behaviour of the functional for different values of these parameters. The available data consists in 314 independent realizations of the reflection coefficient from a stack of 3000 layers of thickness $\epsilon^2 = 0.03^2$ which means that times included in the functional should be less than 5.4. In the frequency domain the sampling rate is 0.05 which corresponds to an available macroscopic frequency range of [3.,12.3] for the exact ϵ value , 0.03 .

Figure (9) shows the variation of the functional as a function of ϵ for a fixed a_0. Each of the two curves correspond to a different value of a_0 in $\{1.03, 1.06\}$. The fact that they all present a minimum, more or less flat and more or less centered around the exact value $\epsilon = 0.03$, indicates that the functional is quite robust with respect to a_0. figure 9

The next figure (10) shows the variation of the functional for ϵ varying in two different intervals $[0.02, 0.04]$ and $[0.01, 0.08]$. This time $a_0 = 1.06$. Here again, the two curves present a clear minimum, even for this small number of realisations upon which the averaging is done (10). In this last experiment the range of frequencies included in the functional varies with ϵ, so as to include as much data as possible. Even though this is hard to justify it does not seem to worsen the behaviour of the functional. In both experiments the true value $\alpha = 0.00375$ is used. In view of this first set of experiments we feel entitled to improve our model by establishing a link between the algorithm numerical parameter a_0 and the unknown ϵ. This is based on the error estimate established by the theorem. We do not know exactly how ϵ and a_0 should be linked and any algebraic relation that enables $\epsilon/(a_0 - 1) \to 0$ and $a_0 - 1 \to 0$ simultaneously is a good candidate. Relations of the form $\epsilon = c(a_0 - 1)^b$

with $b > 1$ and c two arbitrary constants have been extensively tested. They all reproduce the same behaviour of the functional with respect to ϵ and we eventually settle on $\epsilon = (a_0 - 1)^{1.1}$ figure 10

We now minimize the functional simultaneously versus ϵ and versus α - the parameter of equation (7). For this purpose, we use the optimisation routine POWELL of Numerical Recipes [11]. It is an improved relaxation scheme which is not very competitive in terms of iterations number but has the big advantage of dealing with functionals whose derivatives are impossible or hard to compute. At this stage we have not been concerned with the efficiency of the code since we wanted first to verify that the functional was sensitive enough with respect to the two parameters and that we could minimize it. The minimisation algorithm is the following

0. initialization of search directions
$u_\epsilon = (1, 0)$
$u_\alpha = (0, 1)$

1. initialization of solution $P_0 = (\epsilon_0, \alpha_0)$

2. minimization of $f(x) = F(P_0 + x.u_\epsilon)$ versus x and set P_1 at the minimum

3. minimization of $f(x) = F(P_1 + x.u_\alpha)$ versus x and set P_2 at the minimum

4. change the search directions
$u_\alpha \rightarrow u_\epsilon$
$P_2 - P_0 \rightarrow u_\alpha$

5. minimization of $f(x) = F(P_2 + x.u_\alpha)$ versus x and set P_0 at the minimum

6. go back to step 2

The scalar minimizations at steps 2,3 and 5 are performed using procedure BRENT , also found in [11]. For each trial value of the parameter x the functional has to be evaluated for a new set of parameters ϵ and α. α appears only in the theoretical value of the spectrum which is simple to compute in the homogeneous case. On the other hand, trying a new ϵ means computing new estimates of the power spectrum, with a different pulse width $\hat{f}$, and a different wavelet parameter $a_0 = 1 + \epsilon^{\frac{1}{1.1}}$

Figures (11) to (13) present the solution of such an inversion for the dataset described earlier. Figure (11) compares the solution obtained with two different datasets each made of one single realisation. The two top graphs show the variation of the unknown ϵ, the two graphs in the middle the variation of α and the two graphs on the bottom show the value of the functional whose minimum is sought. The abcissa is the iteration number at the lowest level in the algorithm described above. Every evaluation of the functional F in the scalar minimization of $f(x)$ corresponds to one point on the graphs. Obviously from this experiment the redundancy of information coming from using different frequencies and time windows is not sufficient to find the unknowns ϵ and α accurately. figure 11

We present on figure (12) the solutions obtained when using data averaged on different numbers of independent realizations The left hand side graphs show the value of ϵ as a function of the iteration number, the graph in the middle show the value of α and the right hand side graph represents the value of the objective function. There are some interesting features to notice in this experiment: we have mentioned earlier that the iterative minimization stops when the functional has stabilized around a minimum according to an arbitrary criterion set here to 0.001 in relative variation. The figure shows that except for the 10 realizations case, the number of iterations needed to converge diminishes as the data becomes more coherent. Also, it can be noticed from the right hand side column plots that the absolute value of the functional diminishes when the number of realizations involved increases. These two remarks are consistent with the theory: we expect the functional to be smoother and have a minimum nearer zero when more realizations are used. The solution (ϵ, α) becomes acceptable as soon as five independent realizations are used, so we can expect to make this method practical when using point source reflected signals recorded at different receiver locations on the surface. figure 12

figure 13

6 Appendix: proof of the theorem

In order to use the normalization properties of the wavelet, we express the sum in the Fourier domain. Throughout the paper C will denote a constant independent of a_0, ϵ and of any summation index but whose value might change from line to line; we assume also ω to be non-negative because $\Lambda_f(t, w)$ is even. A straightforward computation leads to:

$$G_{nm} = \frac{1}{\sqrt{2\pi}} \frac{2}{\epsilon} \int e^{2\imath(\frac{a_0^m}{\epsilon}\omega - h)(\frac{t}{a_0^m} - nb_0)} \hat{g}(h) \hat{g}(2\frac{a_0^m}{\epsilon}\omega - h)^* dh \tag{36}$$

Let $S_{nm} = \mathbb{E}\{|R^\epsilon_{nm}|^2\} G_{nm}$; then:

$$\begin{aligned} S_{nm} = \; & \left(\frac{1}{\sqrt{2\pi}}\right)^3 \frac{2}{a_0^m} \int\int\int\int\int C_f^\epsilon(\tau, \sigma) e^{2\imath(\frac{a_0^m}{\epsilon}\omega - h)(\frac{t}{a_0^m} - nb_0)} \hat{g}(h) \hat{g}(2\frac{a_0^m}{\epsilon}\omega - h)^*. \\ & .e^{\imath \frac{\omega_1 - \omega_2}{a_0^m}\tau} e^{-\imath \frac{\omega_1 + \omega_2}{a_0^m} \frac{\epsilon\sigma}{2}} e^{\imath(\omega_2 - \omega_1) nb_0} \hat{g}(\omega_1)^* \hat{g}(\omega_2) d\omega_1 d\omega_2 d\tau d\sigma dh \end{aligned}$$

We write the Poisson summation formula in the form:

$$\sum_n e^{\imath\omega n x} = \frac{2\pi}{\omega} \sum_p \delta(x + \frac{2p\pi}{\omega})$$

Using this formula, we can now integrate each term with respect to the σ and h variables and, after the change of variables:

$$v = \frac{\omega_1+\omega_2}{2} \quad u = \frac{\omega_1-\omega_2}{2}$$

we obtain $\sum_n S_{nm} = \sum_p \tilde{S}_{mp}$ where:

$$\begin{aligned} \tilde{S}_{mp} &= \frac{2}{b_0}\frac{a_0^m}{\epsilon}\int\int\int \Lambda_f^\epsilon(\tau,v)\left(\frac{e^{2\imath\frac{u(\tau-t)}{\epsilon}}}{\epsilon}\right)e^{\frac{2\imath\pi pt}{b_0 a_0^m}}. \\ &.\hat{g}\left(\frac{a_0^m}{\epsilon}(v+u)\right)^*\hat{g}\left(\frac{a_0^m}{\epsilon}(v-u)\right)\hat{g}\left(\frac{a_0^m}{\epsilon}(\omega-u)+\frac{\pi p}{b_0}\right)^*\hat{g}\left(\frac{a_0^m}{\epsilon}(\omega+u)-\frac{\pi p}{b_0}\right)dudvd\tau \end{aligned} \tag{37}$$

It is easy to see for which m and p, due to the compact support of $\hat{g}$, $\tilde{S}_{mp}$ is equal to zero; we find that $\frac{a_0^m}{\epsilon}\omega \in [-a_0^2 l, -l] \cup [-\frac{\pi}{b_0}, \frac{\pi}{b_0}] \cup [l, a_0^2 l]$. So, for ω held fixed, there are only two kinds of values for m for which $\tilde{S}_{mp}$ is non-zero:

$$\begin{array}{ll} \frac{a_0^m}{\epsilon}\omega \in [l, a_0^2 l] & \Rightarrow m = M+1, M+2 \text{ with } M = [\frac{\log\frac{l\epsilon}{\omega}}{\log a_0}] \\ 0 < \frac{a_0^m}{\epsilon}\omega \leq \frac{\pi}{b_0} & \Rightarrow m \in]-\infty, \frac{\log\frac{\pi\epsilon}{\omega b_0}}{a_0}] \end{array} \tag{38}$$

We assume that a_0^2 is less than three so that these two sets do not intersect each other. Moreover p takes only 11 values given by: $p = -1, 0, 1$ or $|p \mp \frac{2(a_0^2+1)}{a_0^2-1}| \leq 2$.

We now study the limit of the sum as ϵ goes to zero and then as a_0 goes to 1. We have already changed $\sum_n S_{nm}$ into $\sum_p \tilde{S}_{mp}$; the idea is then to replace $\tilde{S}_{mp}$ by an approximation $\tilde{\tilde{S}}_{mp}$ whose limit is the same as ϵ goes to zero but which is much easier to study for a_0 close to 1. So we first prove that $\lim_{\epsilon\to 0}\sum_{nm} S_{nm} = \lim_{\epsilon\to 0}\sum_{mp}\tilde{\tilde{S}}_{mp}$ where:

$$\tilde{\tilde{S}}_{mp} = \frac{2\pi}{b_0}\frac{a_0^m}{\epsilon}e^{\frac{2\imath\pi pt}{b_0 a_0^m}}\int \Lambda_f(t,v)|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2\hat{g}(\frac{a_0^m}{\epsilon}\omega - \frac{\pi p}{b_0})\hat{g}(\frac{a_0^m}{\epsilon}\omega + \frac{\pi p}{b_0})^* dv$$

let:

$$\begin{aligned} A(u,v) &= \hat{g}\left(\frac{a_0^m}{\epsilon}(u+v)\right)^*\hat{g}\left(\frac{a_0^m}{\epsilon}(v-u)\right) \\ B(u) &= \hat{g}\left(\frac{a_0^m}{\epsilon}(\omega+u)-\frac{\pi p}{b_0}\right)\hat{g}\left(\frac{a_0^m}{\epsilon}(\omega-u)+\frac{\pi p}{b_0}\right)^* \end{aligned}$$

then:

$$\begin{aligned} \tilde{S}_{mp} &= \frac{2}{b_0}\frac{a_0^m}{\epsilon}e^{\frac{2\imath\pi pt}{b_0 a_0^m}}\int\int\int \Lambda_f^\epsilon(\tau,v)\left(\frac{e^{2\imath\frac{u(\tau-t)}{\epsilon}}}{\epsilon}\right)A(u,v)B(u)dudvd\tau \\ &= \frac{1}{b_0}\frac{a_0^m}{\epsilon}e^{\frac{2\imath\pi pt}{b_0 a_0^m}}\int\int \Lambda_f(t,v)\left(\int du e^{\imath\tau u}A(u,v)B(u)\right)dvd\tau + \frac{a_0^m}{\epsilon}R \end{aligned}$$

We have

$$\begin{aligned} \int\left(\int du e^{\imath\tau u}A(u,v)B(u)\right)d\tau &= 2\pi A(0,v)B(0) \\ &= |\hat{g}(\frac{a_0^m}{\epsilon}v)|^2\hat{g}(\frac{a_0^m}{\epsilon}\omega - \frac{\pi p}{b_0})\hat{g}(\frac{a_0^m}{\epsilon}\omega + \frac{\pi p}{b_0})^* \end{aligned}$$

Hence $\tilde{S}_{mp} = \tilde{\tilde{S}}_{mp} + \frac{a_0^m}{\epsilon} R$ and the remaining part R can be bounded by:

$$|R| \le C \int\int \left|\Lambda_f^\epsilon(t+\epsilon\frac{\tau}{2}, v) - \Lambda_f(t,v)\right| . \left|\int e^{\imath\tau u} A(u,v)B(u)du\right| dv d\tau$$

It follows from the assumptions made on Λ and f that:

$$\begin{cases} \Lambda_f^\epsilon(t+\epsilon\frac{\tau}{2}, v) \to \Lambda_f(t,v) & \begin{array}{l}\text{in } L_1(dv) \text{ as } \epsilon \to 0 \\ \text{uniformly in } \tau \text{ on compact sets}\end{array} \\ \exists T, \int |\sup_{t>T} \Lambda_f(t,v)|\, dv < \infty & \end{cases}$$

because the assumption on the smoothness of the pulse shows that, as $V \to \infty$:

$$\begin{array}{l} \int_{|v|>V} \sup_{t,\epsilon} |\Lambda_f^\epsilon(t,v)| dv \to 0 \\ \int_{|v|>V} \sup_t |\Lambda_f(t,v)| dv \to 0 \end{array}$$

$|\int e^{\imath\tau u} A(u,v)B(u)du|$ is bounded uniformly in v, m and p. Since A and B are C^∞ with compact supports, their Fourier transforms are rapidly decreasing and we can change multiplication into convolution, leading to:

$$\begin{aligned} \int e^{\imath\tau u} A(u,v)B(u)du &= (2\pi)^{-5/2} \left(\frac{\epsilon}{a_0^m}\right)^4 \int\int\int e^{-\imath x v} e^{\imath y v} e^{\imath z(-\omega + \frac{\epsilon}{a_0^m}\frac{\pi p}{b_0})} e^{\imath(\tau - x - y - z)(\omega + \frac{\epsilon}{a_0^m}\frac{\pi p}{b_0})}. \\ &\quad .g(-\frac{\epsilon}{a_0^m} x)^* g(-\frac{\epsilon}{a_0^m} y) g(\frac{\epsilon}{a_0^m} z) g\left(\frac{\epsilon}{a_0^m}(\tau - x - y - z)\right)^* dx dy dz \end{aligned}$$

Hence:

$$\int \left|\int e^{\imath\tau u} a(u,v) b(u,v) c(u) d(u) du\right| d\tau \le (2\pi)^{-5/2} \|g\|_{L_1}^4$$

and $\|g\|_{L_1}$ is bounded for a_0 close to 1. Let $T > 0$. We write:

$$\begin{aligned} |R| &\le C\|g\|_{L_1}^4 \int \sup_{\tau \le T} \left|\Lambda_f^\epsilon(t+\epsilon\frac{\tau}{2}, v) - \Lambda_f(t,v)\right| dv \\ &\quad + C\left(\int \sup_{t>T} |\Lambda_f(t,v)|\, dv\right)\left(\int_{\tau>T} \left|\int e^{\imath\tau u} A(u,v)B(u)du\right| d\tau\right) \end{aligned}$$

For T big enough, the second term in the right hand-side is close to zero and we find that, as ϵ goes to zero, $|R| \le O(\epsilon)$. Therefore:

$$\begin{aligned} \sum_m \left|\sum_n S_{nm} - \sum_p \tilde{\tilde{S}}_{mp}\right| &= \sum_m \left|\sum_p \tilde{S}_{mp} - \sum_p \tilde{\tilde{S}}_{mp}\right| \\ &\le C\left(\sum_{0 < \frac{a_0^m}{\epsilon}\omega \le \frac{\pi}{b_0}} \frac{a_0^m}{\epsilon} O(\epsilon) + \sum_{\frac{a_0^m}{\epsilon}\omega \in [l, a_0^2 l]} \frac{a_0^m}{\epsilon} O(\epsilon)\right) \\ &\le \frac{1}{a_0 - 1} O(\epsilon) \end{aligned}$$

We prove now that only the $m = M+1, M+2$ and $p = 0$ terms are not negligible in the limit $\epsilon \to 0$ and $a_0 \to 1$.

- If $p = 0$ then $\tilde{\tilde{S}}_{mp} = \frac{2\pi}{b_0}\frac{a_0^m}{\epsilon} e^{\frac{2\imath\pi p t}{b_0 a_0^m}} \int \Lambda_f(t,v) |\hat{g}(\frac{a_0^m}{\epsilon} v)|^2 \hat{g}(\frac{a_0^m}{\epsilon}\omega - \frac{\pi p}{b_0}) \hat{g}(\frac{a_0^m}{\epsilon}\omega + \frac{\pi p}{b_0})^* dv$ is nonzero only for m such that $\frac{a_0^m}{\epsilon}\omega \in [l, a_0^2 l]$ so $m = M+1, M+2$.

- If $p \neq 0$ then $\hat{g}(\frac{a_0^m}{\epsilon}\omega - \frac{\pi p}{b_0})\hat{g}(\frac{a_0^m}{\epsilon}\omega + \frac{\pi p}{b_0})^*$ is nonzero for m such that $0 \leq \frac{a_0^m}{\epsilon}\omega \leq \frac{\pi}{b_0}$ and, since p takes only a finite number of values, we can bound the sum:

$$\begin{array}{ll} |\sum_{m\in\mathbb{Z},p\in\mathbb{Z}^*} \tilde{\tilde{S}}_{mp}| & = |\sum_{p\neq 0, 0\leq \frac{a_0^m}{\epsilon}\omega \leq \frac{\pi}{b_0}} \tilde{\tilde{S}}_{mp}| \\ & \leq C \sum_{0\leq \frac{a_0^m}{\epsilon}\omega\leq\frac{\pi}{b_0}} \frac{2\pi}{b_0}\frac{\pi}{b_0\omega} \int_{\frac{a_0^m}{\epsilon}v\in[-a_0^2 l,-l]\cup[l,a_0^2 l]} |\Lambda_f(t,v)|.|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2 dv \\ & \leq C \int_{\frac{2\omega}{a_0^2-1}}^{\infty} |\Lambda_f(t,v)| dv \end{array}$$

Thus, we find the following estimate:

$$\sum_m \left| \sum_n S_{nm} - \tilde{\tilde{S}}_{m,0} \right| \leq \left(\frac{O(\epsilon)}{a_0 - 1} + O(a_0 - 1) \right)$$

with $\sum_m \tilde{\tilde{S}}_{m,0} = \tilde{\tilde{S}}_{M+1,0} + \tilde{\tilde{S}}_{M+2,0}$ which explains why only two indices $m = M+1$ and $m = M+2$ are significant in the limit. The only thing left to do is to compute the limit as a_0 goes to 1 of:

$$\sum_m \tilde{\tilde{S}}_{m,0} = \sum_m \frac{2\pi}{b_0}\frac{a_0^m}{\epsilon} \int \Lambda_f(t,v)|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2|\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv$$

We suppose $v > 0$ ($\Lambda_f(t,v)$ is even) and we decompose the sum as

$$\frac{1}{2}\sum_m \tilde{\tilde{S}}_{m,0} = \sum_m (S_m^1 + S_m^2 + S_m^3)$$

with

$$\begin{array}{ll} S_m^1 & = \frac{2\pi}{b_0}\frac{a_0^m}{\epsilon}\int_0^\infty (\Lambda_f(t,v) - \Lambda_f(t,\omega))\,|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2|\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv \\ S_m^2 & = \frac{2\pi}{b_0}\int_0^\infty \Lambda_f(t,\omega)\left(\frac{a_0^m}{\epsilon} - \frac{2\pi}{b_0}\frac{1}{2v\log a_0}\right)|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2|\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv \\ S_m^3 & = \int_0^\infty \Lambda_f(t,\omega)\left(\frac{2\pi}{b_0}\right)^2 \frac{1}{2v\log a_0}|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2|\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv \end{array}$$

Because of the normalization properties (12), it is easy to see that:

$$\sum_m \int_0^\infty \frac{|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2}{v}|\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv = \left(\sum_m |\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2\right)\int_0^\infty \frac{|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2}{v} dv = \left(\frac{b_0}{2\pi}\right)^2 \log a_0$$

which implies $\sum_m S_m^3 = \frac{1}{2}\Lambda_f(t,\omega)$. The two other sums go to zero; the support property shows that:

$$\begin{array}{c} l \leq \frac{a_0^m}{\epsilon}\omega \leq a_0^2 l \\ l \leq \frac{a_0^m}{\epsilon}v \leq a_0^2 l \end{array}$$

Thus $|\omega - v| \leq (a_0^2-1)\omega$ and $\left|\frac{a_0^m}{\epsilon} - \frac{2\pi}{b_0}\frac{1}{2v\log a_0}\right| \sim \frac{4\pi}{b_0 v}$ for a_0 close to 1. Hence:

$$\begin{array}{ll} \sum_m S_m^1 & \leq \sup_{|\omega-v|\leq(a_0^2-1)\omega} |\Lambda_f(t,v) - \Lambda_f(t,\omega)| \sum_m \frac{2\pi}{b_0}\frac{a_0^m}{\epsilon}\int_0^\infty |\hat{g}(\frac{a_0^m}{\epsilon}v)|^2|\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv \\ & \leq C \sup_{|\omega-v|\leq(a_0^2-1)\omega} |\Lambda_f(t,v) - \Lambda_f(t,\omega)| \to 0 \end{array}$$

and:

$$\begin{array}{ll}\sum_m S_m^2 & \leq C|\Lambda_f(t,\omega)| \sum_m \int_0^\infty \frac{|\hat{g}(\frac{a_0^m}{\epsilon}v)|^2}{v} |\hat{g}(\frac{a_0^m}{\epsilon}\omega)|^2 dv \\ & \leq C \log a_0 \to 0\end{array}$$

Finally we obtain:

$$\left|\sum_{nm} S_{nm} - \Lambda_f(t,\omega)\right| \leq \frac{O(\epsilon)}{a_0 - 1} + O(a_0 - 1)$$

which proves our claim.

7 Conclusion

This paper makes a precise description of the power spectral density of a nonstationary signal. In the case of signals reflected by a random layered medium it describes a numerical method to estimate the power spectrum on a discrete time-frequency grid. Wavelet transform techniques allow us to use rigorously the local stationarity of the signal. The theorem proved in the paper validates the intuition that only the wavelet coefficients corresponding to dilation factors of the order of the local stationarity scale contribute to the estimate. Furthermore the proof also establishes a link between the scale factor ϵ and the wavelet dilation coefficient a_0. This link can be and is used extensively in the numerical simulations. The convergence of the algorithm and its robustness with respect to a_0 is illustrated using the true value of ϵ. Then we present preliminary experiments showing the feasibility of the inverse problem which consists in the determination of epsilon. In this paper we only deal with random medium of constant background and we show that the wavelet estimation algorithm allows us to recover both the scale factor and the random noise intensity. In the near future new numerical simulations should verify whether the estimates of the power spectrum are exponentially distributed or not when obtained using an arbitrary scale factor. In the inverse problem the two next steps are the simultaneous inversion of the scale factor and a varying background for multiple realizations with plane wave data and the determination of ϵ for a single realization point source data.

8 Figures

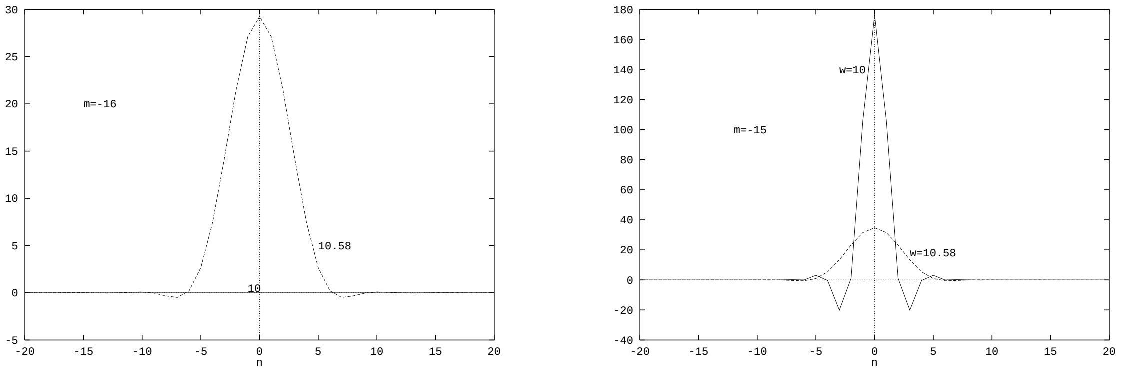

Figure 1 :

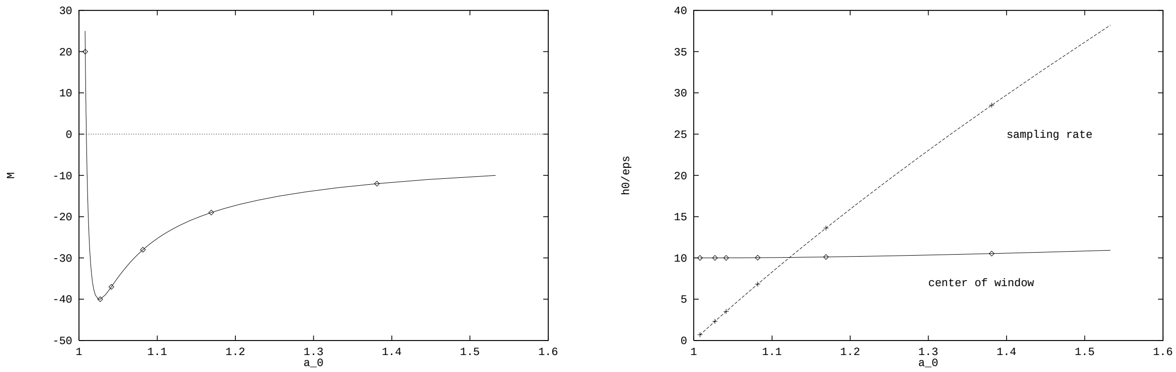

Figure 2 :

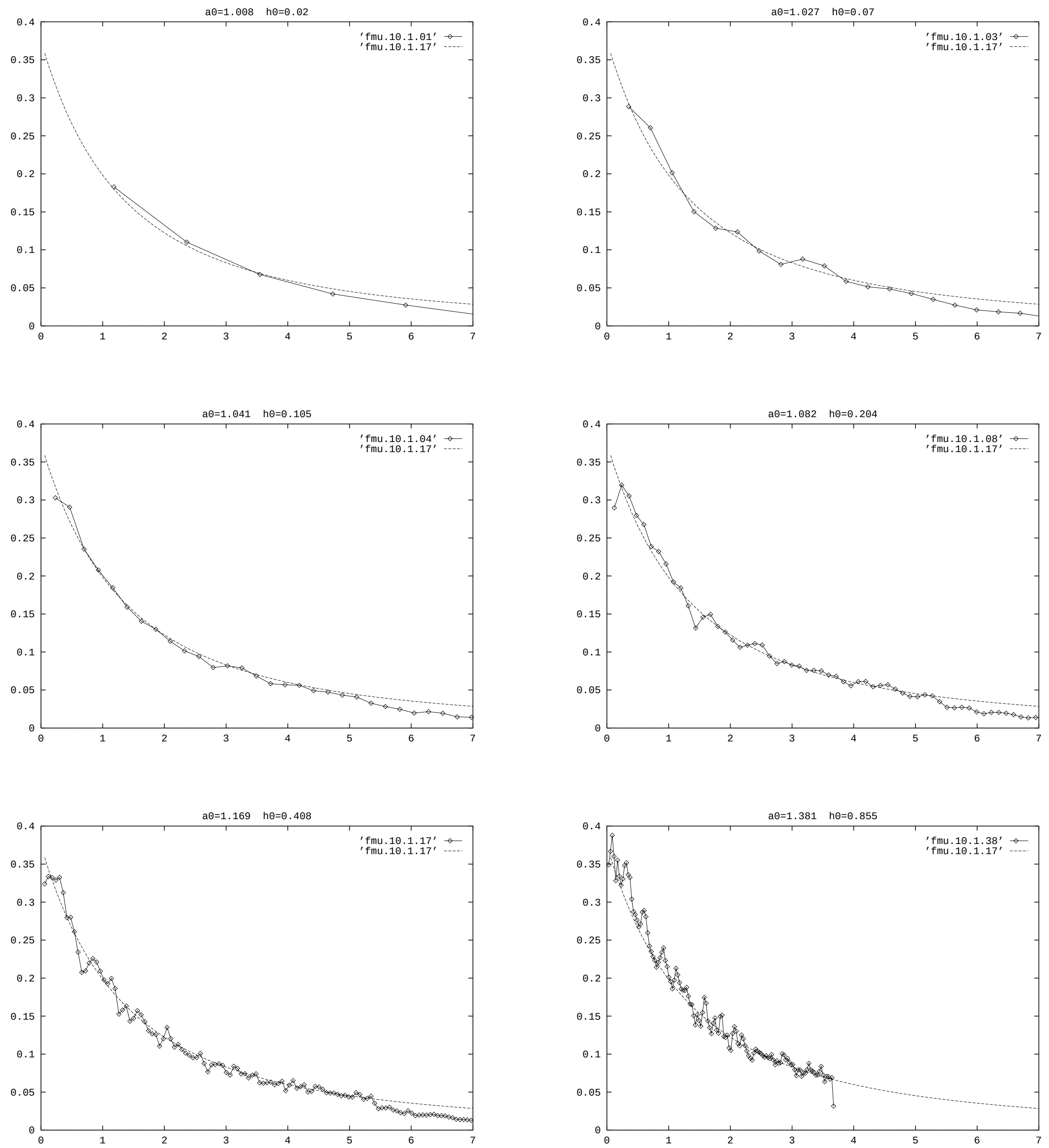
a0=1.008 h0=0.02
'fmu.10.1.01'
'fmu.10.1.17'
a0=1.027 h0=0.07
'fmu.10.1.03'
'fmu.10.1.17'
a0=1.041 h0=0.105
'fmu.10.1.04'
'fmu.10.1.17'
a0=1.082 h0=0.204
'fmu.10.1.08'
'fmu.10.1.17'
a0=1.169 h0=0.408
'fmu.10.1.17'
'fmu.10.1.17'
a0=1.381 h0=0.855
'fmu.10.1.38'
'fmu.10.1.17'

Figure 3 :

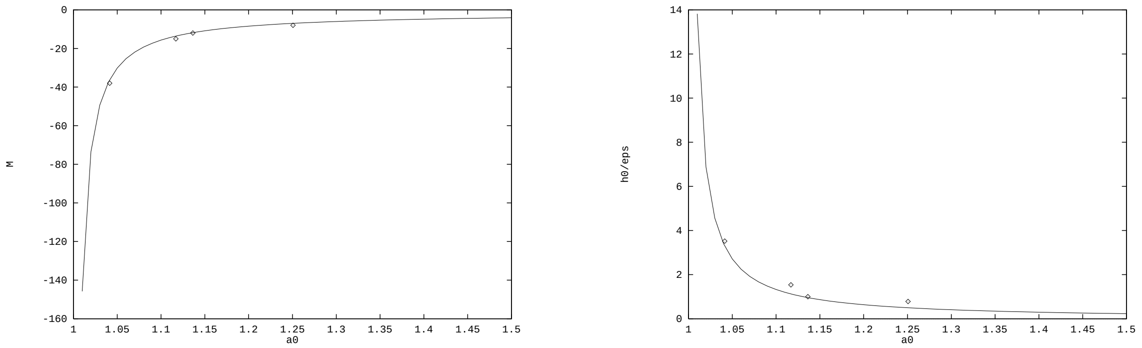

Figure 4 :

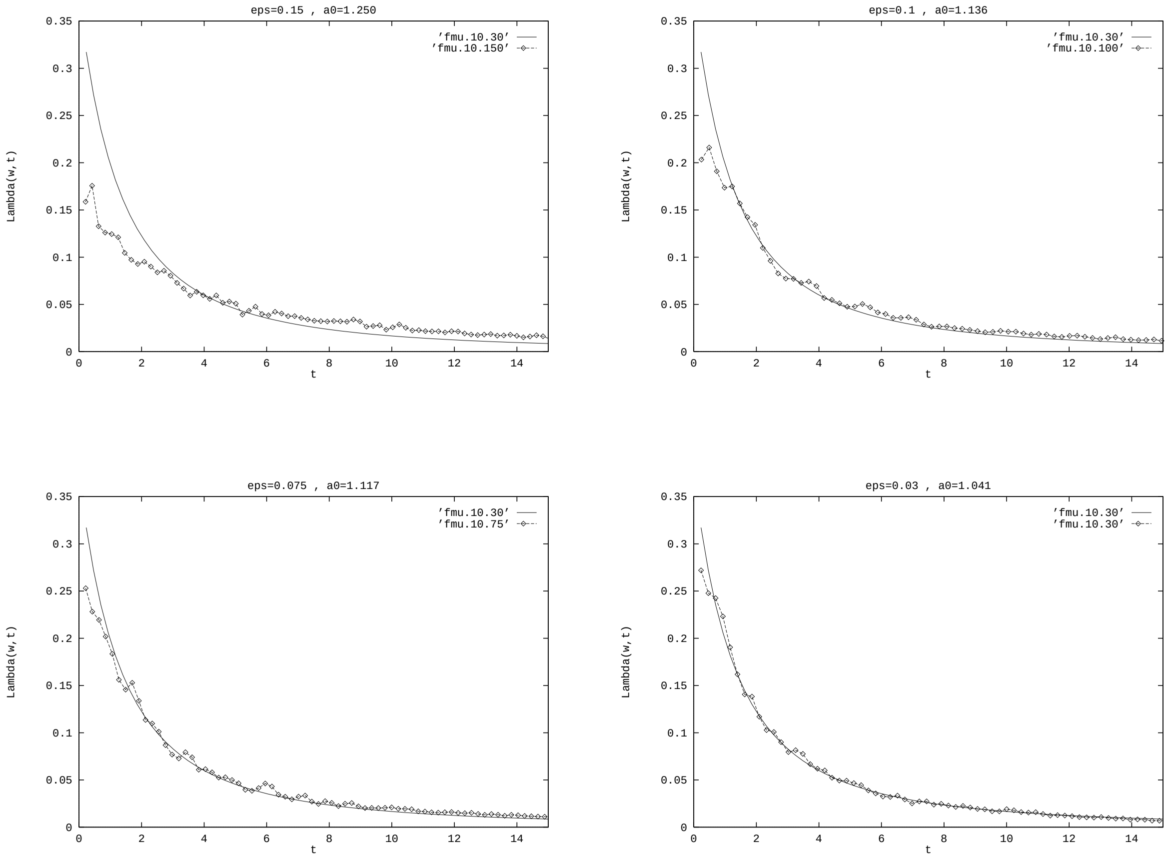

Figure 5 :

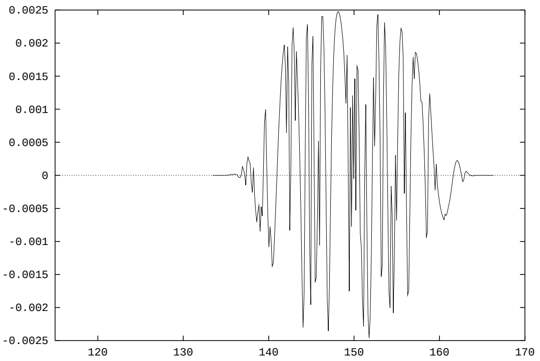

Figure 6 :

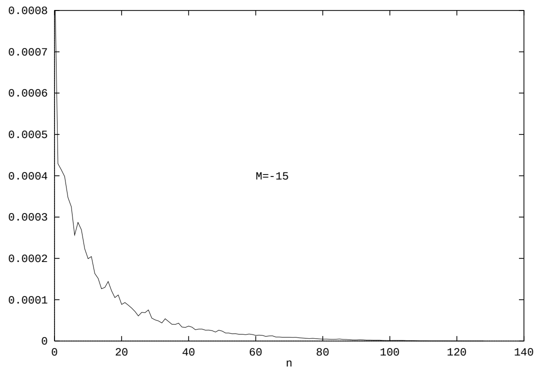

Figure 7 :

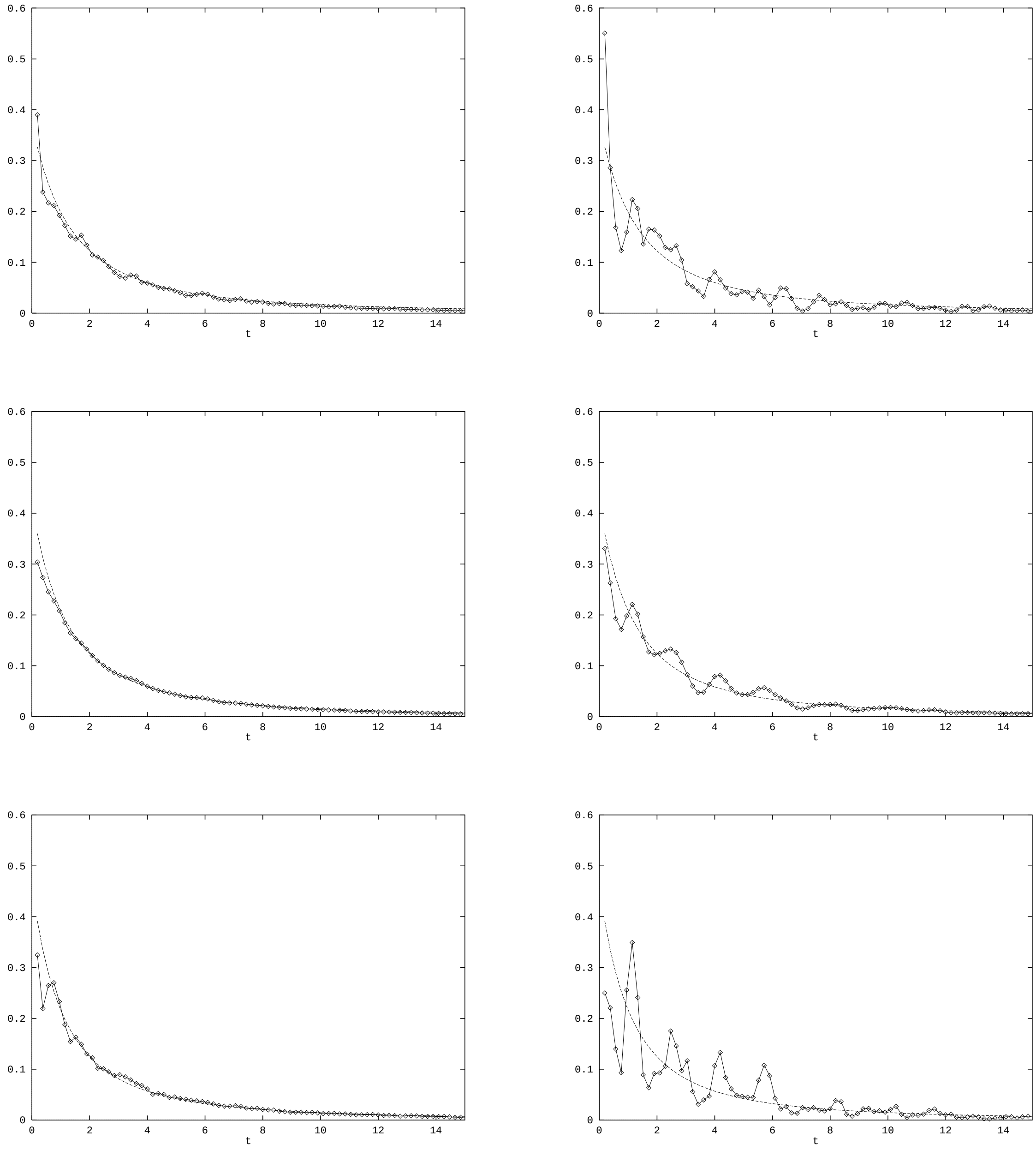

Figure 8 :

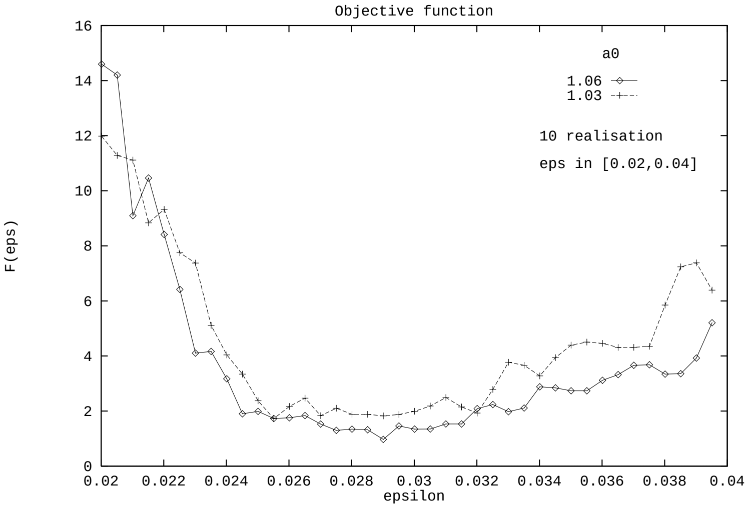

Figure 9 :

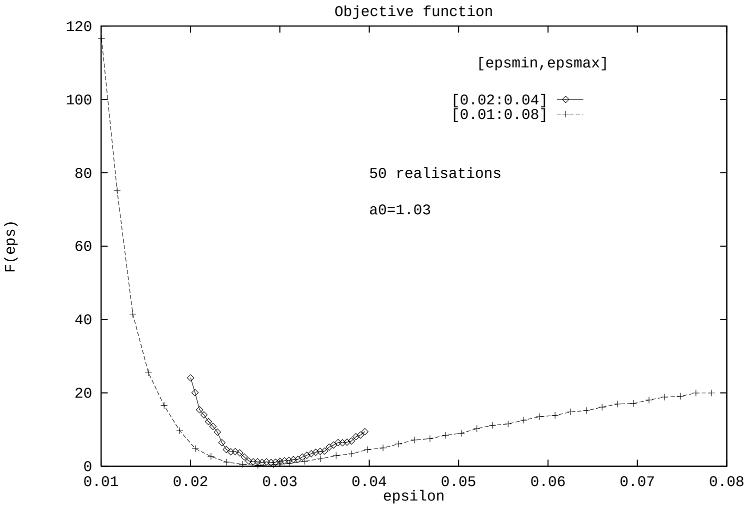

Figure 10 :

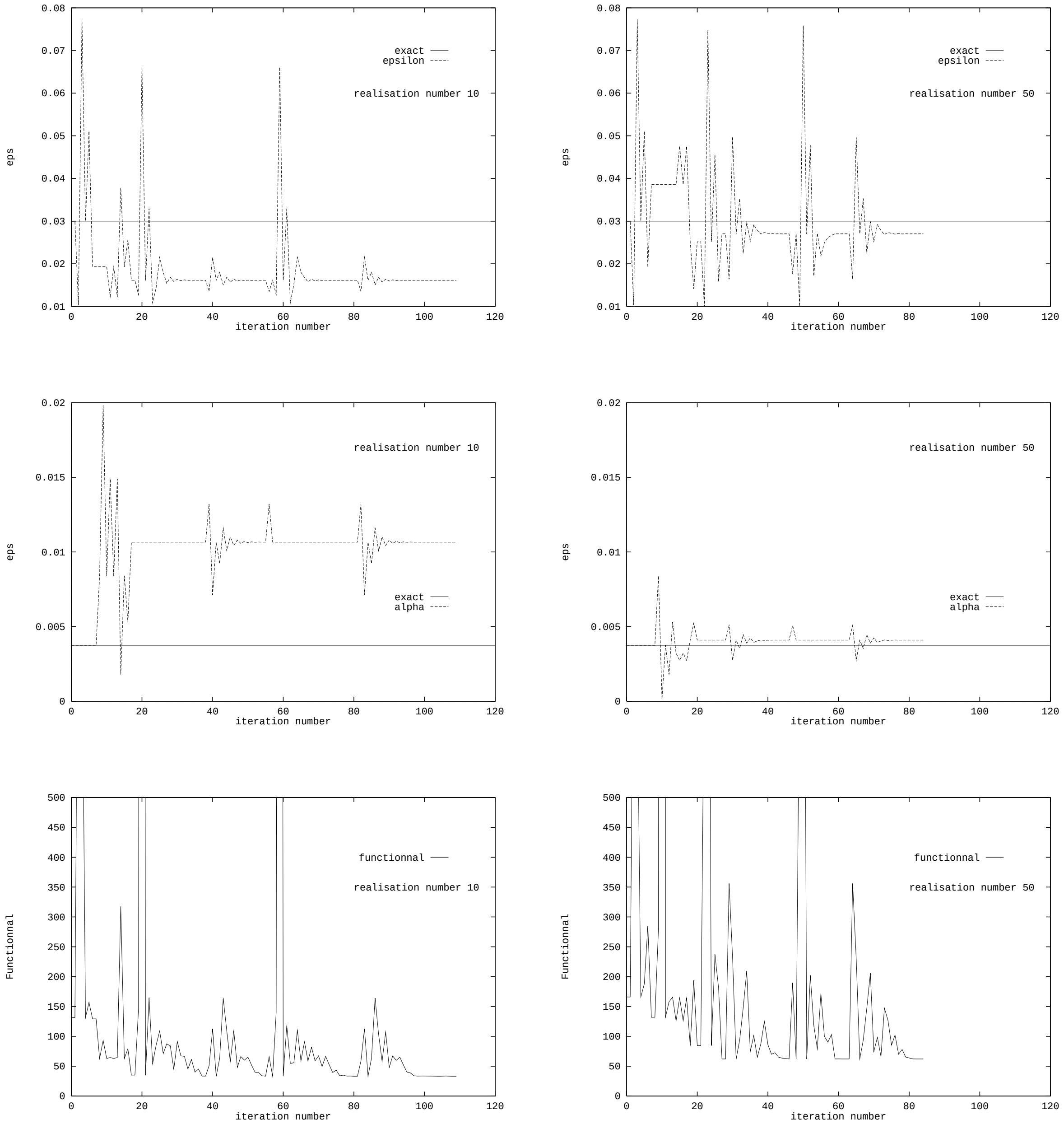
eps
iteration number
exact
epsilon
realisation number 10
realisation number 50
exact
alpha
realisation number 10
realisation number 50
Functionnal
functionnal
realisation number 10
realisation number 50

Figure 11 :

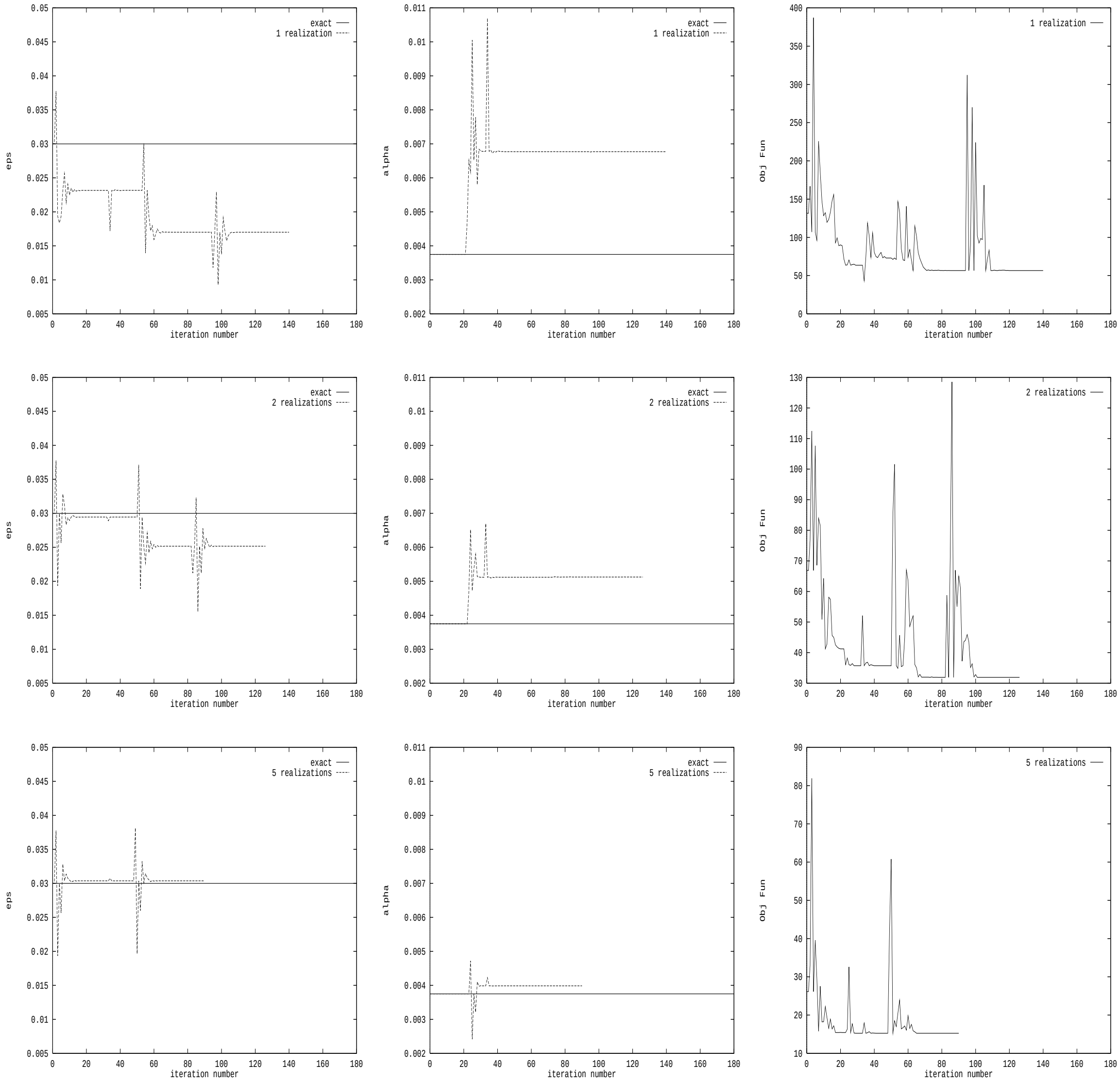
exact
1 realization
2 realizations
5 realizations
eps
alpha
Obj Fun
iteration number

Figure 12 :

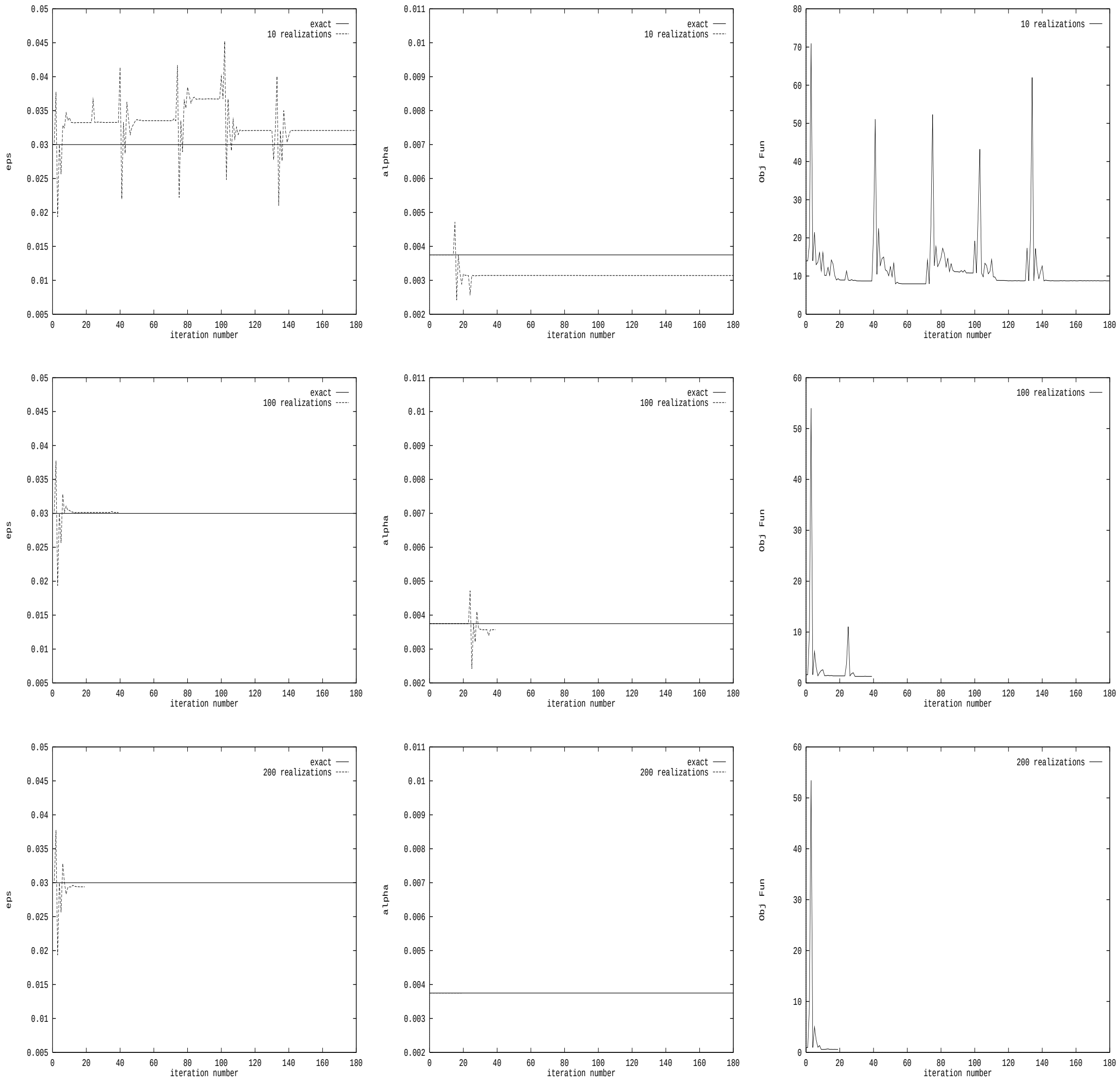
exact
10 realizations
100 realizations
200 realizations
eps
alpha
Obj Fun
iteration number

Figure 13 :

FIGURE 1:$\hat{G}_{mn}(\omega, t=0)$ for $m = M+1 = -16$ (right) and $m = M+2 = -15$ (left) and 2 frequencies: $\omega = 10 = l\epsilon a_0^{-17}$ and $\omega = 10.58$

FIGURE 2:Variations of M, h_0 as functions of a_0

FIGURE 3:$\Lambda(\omega, t)$ as a function of t, $\omega = 10$. 3000 layers

FIGURE 4:Variations of M, h_0/ϵ as functions of a_0

FIGURE 5:$\Lambda(\omega, t)$ as a function of t, $\omega = 10$. 6000 layers

FIGURE 6:Real part of $\hat{R}\hat{f}\hat{h}$ used to compute $R^{\epsilon}_{-16,n}$ (left) and $R^{\epsilon}_{-15,n}$ (right) as a function of frequency

FIGURE 7:$\mathbb{E}\{|R^{\epsilon}_{mn}|^2\}$ averaged over 300 realizations. $m = M+1 = -16$ on the left hand side and $m = M+2 = -15$ on the right hand side

FIGURE 8:$\Lambda(\omega, t)$ as a function of t averaged over 300 (left) and 10(right) realizations. $a_0 = 1.117$, $\omega = 10, a = a_0$ (top), $\omega = 10.58, a = 1.0556$ (middle) $\omega = 11.168, a_0 = 1$ (bottom), $\epsilon = 0.075$, $\omega = l\epsilon a_0^{-14}/a$

FIGURE 9:influence of a_0 on the functional

FIGURE 10:influence of ϵ range on the functional

FIGURE 11:Inversion of two single realizations

FIGURE 12:Solution obtained using 1, 2 or 5 independent realizations

FIGURE 13:Solution obtained using 10, 100 or 200 independent realizations

Bibliography

[1] M. Asch, W. Kohler, G. Papanicolaou, M. Postel and B. White. Frequency content of randomly scattered signals. *SIAM Review* 33, 526-629, (1991).

[2] R. Burridge, G. Papanicolaou, P. Sheng and B. White. Probing a random medium with a pulse. *SIAM J. Appl. Math.* 49, 582-607, (1989).

[3] G. Papanicolaou and S. Weinryb. *A functional limit theorem for waves reflected by a random medium. Appl. Math. Optimiz.* 30, 307-334, (1991).

[4] M.B. Priestley. *Non-linear and non-stationary time series analysis.* Academic Press, New York,, 1988.

[5] M. Asch, G. Papanicolaou, M. Postel, P. Sheng and B. White. Frequency content of randomly scattered signals. Part I. *Wave Motion* 12, 429-450, (1990).

[6] G. Papanicolaou, P. Sheng, B. White and Z.Q. Zhang. Multiple scattering noise in one dimension: universality through localization length scales. *Phys. Rev. Lett.* 57, 1000-1003, (1986).

[7] G. Papanicolaou, P. Sheng, B. White and Z.Q. Zhang. Wave localization characteristics in the time domain. *Phys. Rev. Lett.* 59, 1918-1921, (1987).

[8] I. Daubechies. The wavelet transform, time frequency localization and signal analysis. *IEEE Transactions on Information Theory* 36, (1990). No 5.

[9] J. Lowengrub. *Ph.D. dissertation.* Courant Institute, New York University, 1989.

[10] G. Papanicolaou, M. Postel, P. Sheng and B. White. Frequency content of randomly scattered signals. Part II: Inversion. *Wave Motion* 12, 527-549, (1990).

[11] B.P. Flannery, W.H. Press, S.A. Teukolsky and W.T. Vetterling. *Numerical Recipes, The Art of Scientific Computing.* Cambridge University Press, 1986.