Nonparametric estimation of the heterogeneity of a random medium using compound Poisson process modeling of wave multiple scattering

Nicolas Le Bihan*

*GIPSA-Lab/CNRS, Department of Images and Signal, 961 Rue de la Houille Blanche,
Domaine Universitaire, 38402 Saint Martin d'Hères Cedex, France*

Ludovic Margerin†

CEREGE, Europôle Méditerranéen de l'Arbois, BP 80, 13545 Aix En Provence Cedex 4, France

In this paper, we present a nonparametric method to estimate the heterogeneity of a random medium from the angular distribution of intensity of waves transmitted through a slab of random material. Our approach is based on the modeling of forward multiple scattering using compound Poisson processes on compact Lie groups. The estimation technique is validated through numerical simulations based on radiative transfer theory.

I. INTRODUCTION

Small-angle scattering of electromagnetic, acoustic, or elastic waves constitutes the basic tool to characterize porous media [1], solid suspensions [2], or Earth's lithosphere [3] to cite a few examples only. In many cases, the probing wave is assumed to undergo single scattering only and multiple scattering is often considered as a nuisance. The main reason is that in the single-scattering regime, the diffraction pattern of intensity, i.e., the *phase function* is related to the power spectrum of the medium heterogeneities in a simple manner [4]. In the field of optics, some solutions have been proposed to invert for the size of particles in the multiple-scattering regime [5] based on various small-angle approximations of the radiative transfer equation [6].

In this work, we develop an alternative approach based on the compound Poisson process (CPP) model of wave multiple scattering in random media. Our model is generic and makes little assumption on the underlying wave equation. The random medium is described by its heterogeneity power spectrum that refers to the Fourier transform of the spatial two-point correlation function. This enables us to treat on the same footing turbid media such as earth's crust or the atmosphere, aggregates of particles, and porous media [7]. The finer details of the microstructure encapsulated in higher moments of the random field (>2) are not considered in our approach.

We consider waves propagating in a turbid medium described by the Von Karman correlation function, which is commonly used to represent geophysical media. We demonstrate the possibility to estimate the power spectrum of the random fluctuations from the angular distribution of intensity transmitted through a slab of material. We find that the CPP model is effective before the diffusive regime sets in, i.e., when the slab thickness is less than one transport mean free path.

The organization of the paper is as follows. In Sec. II, the necessary mathematical background on noncommutative har-

monic analysis is provided. In Sec. III, we develop the Poisson process approach to multiple scattering. In Sec. IV, we validate the model through comparisons with Monte Carlo simulations of the radiative transfer equation. The inverse problem is examined in Sec. V, where we propose an estimator of the power spectrum of heterogeneities and discuss its statistical properties. In Sec. VI, the inverse method is validated for realistic examples of random media through numerical simulations.

II. NONCOMMUTATIVE HARMONIC ANALYSIS

In this section, we summarize noncommutative harmonic analysis results that are important for the present work. We are interested in the harmonic analysis of probability density functions (pdfs) over compact Lie groups to develop a nonparametric approach for the estimation of their characteristic functions [8], i.e., their Fourier transform, in Sec. V.

Consider the direction of propagation of a particle or a wave represented by the normalized vector $\mu \in \mathcal{S}^2$, where $\mathcal{S}^2$ is the unit sphere in $\mathbb{R}^3$. We denote by μ_N the direction of propagation of the particle after N scattering events. Such scattering events are considered random so that μ_N is a random variable on $\mathcal{S}^2$. The relation between the direction after $N-1$ scattering events, i.e., μ_{N-1}, and the direction after N scattering events μ_N can be written as

$$\mu_N = \mathbf{r}_N \mu_{N-1}, \tag{1}$$

where $\mathbf{r}_N$ is a random element of the rotation group SO(3). This follows from the transitive action of SO(3) on $\mathcal{S}^2$. Assuming that a particle enters the random medium with initial direction of propagation μ_0, its direction of propagation after N scattering events in the medium is given by

$$\mu_N = \mathbf{r}_N \mathbf{r}_{N-1} \cdots \mathbf{r}_2 \mathbf{r}_1 \mu_0 = \prod_{n=1}^{N} \mathbf{r}_n \mu_0, \tag{2}$$

where each $\mathbf{r}_n$ represents the random action of the nth scatterer, i.e., the $\mathbf{r}_n$ are SO(3)-valued random variables. In the sequel, we present some results about the pdf of random

*nicolas.le-bihan@gipsa-lab.inpg.fr
†margerin@cerege.fr

variables of form (2). We make use of known results in harmonic analysis on compact Lie groups [9–11].

A. Harmonic analysis on SO(3) and $\mathcal{S}^2$

Let us denote by $\mathcal{L}^2(\mathrm{SO}(3), d\mathbf{r})$ the space of square integrable functions on SO(3) with respect to the Haar measure $d\mathbf{r}$ [9]. A pdf on SO(3) is a function $f \in \mathcal{L}^2[\mathrm{SO}(3), d\mathbf{r}]$ that obeys the constraint $\int_{\mathrm{SO}(3)} f(\mathbf{r}) d\mathbf{r} = 1$. A pdf on SO(3) can be decomposed over the complete orthonormal basis of Wigner-D functions $D_l^{pq}(\mathbf{r})$, with $\mathbf{r} \in \mathrm{SO}(3)$, $l \in \mathbb{Z}^+$, and $p, q \in \mathbb{Z}$. This means that given a random variable $\mathbf{r} \in \mathrm{SO}(3)$, its pdf $f(\mathbf{r})$, can be expressed as the infinite series

$$f(\mathbf{r}) = \sum_{l \geq 0} \sum_{p,q=-l}^{l} (2l+1) \hat{f}_l^{pq} \overline{D_l^{pq}(\mathbf{r})}, \qquad (3)$$

where

$$\hat{f}_l^{pq} = \int_{\mathrm{SO}(3)} f(\mathbf{r}) D_l^{pq}(\mathbf{r}) d\mathbf{r} \qquad (4)$$

and the overbar means complex conjugation. The set of coefficients $\hat{f}_l^{pq}$ are sometimes called the "Fourier transform" of $f(\mathbf{r})$ [12]. Note that it is possible to use a matrix notation, namely, $\hat{\mathbf{f}}_l$, in which case, the elements of the $(2l+1) \times (2l+1)$ matrix $\hat{\mathbf{f}}_l$ are $\{\hat{\mathbf{f}}_l\}_{pq} = \hat{f}_l^{pq}$.

Similarly, let us denote by $\mathcal{L}^2(\mathcal{S}^2, d\mu)$ the space of square integrable functions on the unit sphere $\mathcal{S}^2$ with respect to the invariant measure on the sphere $d\mu$. A pdf $w \in \mathcal{L}^2(\mathcal{S}^2, d\mu)$ also satisfies $\int_{\mathcal{S}^2} w(\mu) d\mu = 1$. Similar to Eq. (3), a pdf w on $\mathcal{S}^2$ can be decomposed into an infinite series using the well-known spherical harmonics $Y_l^p(\mu)$ with $\mu \in \mathcal{S}^2$, $l \in \mathbb{Z}^+$, and $p \in \mathbb{Z}$. Thus, given a random variable $\mu \in \mathcal{S}^2$, its pdf $w(\mu)$ can be written as

$$w(\mu) = \sum_{l \geq 0} \sum_{p=-l}^{l} (2l+1) \hat{w}_l^p \overline{Y_l^p(\mu)}, \qquad (5)$$

with

$$\hat{w}_l^p = \int_{\mathcal{S}^2} w(\mu) Y_l^p(\mu) d\mu. \qquad (6)$$

In this case, it is also possible to use a vector representation, where the elements of the $(2l+1)$ vectors $\hat{\mathbf{w}}_l$ are $\hat{w}_l^p (p = -l, \ldots, 0, \ldots, l)$. Again, $\hat{w}_l^p$ is called the Fourier transform of $w(\mu)$.

We now make use of the following fundamental property: the action of SO(3) on $\mathcal{S}^2$ is transitive [10]. As a consequence, for any $\mathbf{r} \in \mathrm{SO}(3)$ and $\mu \in \mathcal{S}^2$, we have $\mathbf{r}\mu \in \mathcal{S}^2$. This Lie group action of the rotation group on the unit sphere implies that, given two pdfs $f(\mathbf{r})$ and $w(\mu)$ taking, respectively, values on SO(3) and $\mathcal{S}^2$, their convolution is a pdf over $\mathcal{S}^2$ and reads as

$$h(\mu) = (f * w)(\mu) = \int_{\mathrm{SO}(3)} f(\mathbf{r}) w(\mathbf{r}^{-1}\mu) d\mathbf{r}, \qquad (7)$$

for any $\mu \in \mathcal{S}^2$, where $*$ means convolution. A very important and interesting consequence of the Peter-Weyl theorem

[9] is that this convolution equation is transformed into a multiplication in the "frequency domain" as follows:

$$h(\mu) = \sum_{l \geq 0} \sum_{p=-l}^{l} (2l+1) \hat{h}_l^p \overline{Y_l^p(\mu)}. \qquad (8)$$

Using the matrix or vector notation introduced previously, we obtain the following relation:

$$\hat{\mathbf{h}}_l = \hat{\mathbf{f}}_l \hat{\mathbf{w}}_l \ \forall \ l, \qquad (9)$$

which is a matrix-vector multiplication taking place for each degree l. As explained in [11], the pdf of $\mathbf{r} = \mathbf{r}_1 \mathbf{r}_2$, is the convolution of the pdfs of $\mathbf{r}_1$ and $\mathbf{r}_2$, namely, $f_1(\mathbf{r}_1)$ and $f_2(\mathbf{r}_2)$, where convolution is again defined with respect to the group action:

$$f(\mathbf{r}) = \int_{\mathrm{SO}(3)} f_2(\mathbf{t}) f_1(\mathbf{t}^{-1}\mathbf{r}) d\mathbf{t}. \qquad (10)$$

In Eq. (10), $\mathbf{t}$ denotes an element of SO(3) and $d\mathbf{t}$ stands for the Haar measure. Consequently, an N-times product of independent random elements of SO(3), i.e., N consecutive random rotations, consists in a random rotation which pdf is the N-times convolution of the pdfs of the elementary rotations. Thus, if we denote by $\mathbf{r} = \mathbf{r}_N \mathbf{r}_{N-1} \cdots \mathbf{r}_2 \mathbf{r}_1$ the outcome of N successive rotations, the Fourier transform of $\mathbf{r}$ can be expressed as follows:

$$\hat{\mathbf{f}}_l = \prod_{n=1}^{N} (\hat{\mathbf{f}}_n)_l \ \forall \ l. \qquad (11)$$

which is a simple product of matrices. In Eq. (11), $(\hat{\mathbf{f}}_n)_l$ denotes the matrix of Fourier coefficients of degree l of f_n, the pdf of $\mathbf{r}_n$.

B. Symmetries

Using Eqs. (9) and (11), it is possible to obtain the Fourier coefficients of any distribution on $\mathcal{S}^2$ that was multiply convolved by pdfs on SO(3). In what follows, we will derive more specific results for distributions on $\mathcal{S}^2$ which are functions of $\cos \theta$ only, with θ as the scattering angle. Such functions are called *zonal* functions on $\mathcal{S}^2$ [9]. This simplification relies on the property of statistical isotropy of the random medium. As a consequence, the Fourier transform over $\mathcal{S}^2$ reduces to a Legendre expansion over $[-1; 1]$ for zonal functions. In anisometric random media where the correlation length depends on space direction, this simplification is not permitted.

To obtain simple results, we further impose that the initial distribution of directions of propagation μ_0 is a zonal function. Such an assumption is not too restrictive and covers for example the case of an incident plane wave, i.e., μ_0 has a Dirac distribution. Using the $\mathcal{XZX}$ parametrization of SO(3) [12], any rotation $\mathbf{r} \in \mathrm{SO}(3)$ can be parametrized using Euler angles φ, θ and ψ. Elements of $\mathcal{S}^2$ are parametrized by φ (azimuth) and θ (coelevation). In the sequel, we use the Euler angle in the parametrization of pdfs.

The symmetry of the scattering process around the direction of propagation implies that the pdf $w_\mu(\mu)$ of the direc-

tion of propagation μ, parametrized using coelevation θ (with respect to direction of propagation) and azimuth φ, is invariant with respect to φ. This allows us to expand $w_\mu(\varphi, \theta)$ as follows:

$$w_\mu(\varphi, \theta) = \sum_{l \geq 0} (2l+1) \hat{w}_l^0 Y_l^0(\varphi, \theta) = \sum_{l \geq 0} (2l+1) \hat{w}_l P_l(\cos \theta),$$

(12)

where $\hat{w}_l$ are scalar quantities. The rotational symmetry of the scattering process also implies that the pdf $f_{\mathbf{r}}(\varphi, \theta, \psi)$ of random elements $\mathbf{r} \in SO(3)$ must be bi-invariant, i.e., invariant by left and right actions on $SO(3)$ (see [9] for details). Such a bi-invariant pdf $f_{\mathbf{r}}$ depends on the single variable θ, i.e., $f_{\mathbf{r}}(\varphi, \theta, \psi) = f_{\mathbf{r}}(\theta)$, and can therefore be expanded as

$$f_{\mathbf{r}}(\theta) = \sum_{l \geq 0} (2l+1) \hat{f}_l^{00} \overline{D_l^{00}(\varphi, \theta, \psi)} = \sum_{l \geq 0} (2l+1) \hat{f}_l P_l(\cos \theta).$$

(13)

Considering the successive action of N independent and identically distributed random rotations $\mathbf{r}_n \in SO(3)$ with bi-invariant pdf on the initial distribution of propagation directions μ_0 leads to the following expression of the pdf $h_{\mu_N}(\theta)$ of the random variable $\mu_N = \mathbf{r}_N \mathbf{r}_{N-1} \cdots \mathbf{r}_2 \mathbf{r}_1 \mu_0$:

$$h_{\mu_N}(\theta) = \sum_{l \geq 0} (2l+1)(\hat{f}_l)^N \hat{w}_l P_l(\cos \theta).$$

(14)

Note that the pdf of the direction of propagation μ_N after N scattering events h_{μ_N} is a function of θ, only.

In the special case of an incident plane wave or a highly collimated beam, the initial direction obeys the following simple probability distribution $w_{\mu_0}(\theta) = \delta(\cos \theta - \cos \theta_0)$, where θ_0 is the direction of propagation of the incident wave. For convenience, we set $\theta_0 = 0$, which yields

$$h_{\mu_N}(\theta) = \sum_{l \geq 0} (2l+1)(\hat{f}_l)^N P_l(\cos \theta).$$

(15)

The pdf of μ_N possesses a simple Legendre expansion with coefficients equal to the Nth power of the Legendre coefficients $\hat{f}_l$ of the $SO(3)$ random variables.

Note that Eq. (15) could have been obtained using the well-known summation formula for Jacobi polynomials (see Chapter 2 in [9]) in the case of *zonal* functions on the sphere. The approach developed in this section is more general because it could be applied to less symmetrical distributions, thereby allowing the modeling of more complicated scattering processes.

III. COMPOUND POISSON PROCESS MODEL FOR MULTIPLE SCATTERING

In this section, we develop the multiple-scattering model. We begin with a summary of useful results about CPPs on compact Lie groups [13]. Note that we use the term pdf when we refer to the function describing the angular pattern of scattering. In Sec. IV, we will substitute it with *phase function* which is the usual terminology in the field of scattering in random media.
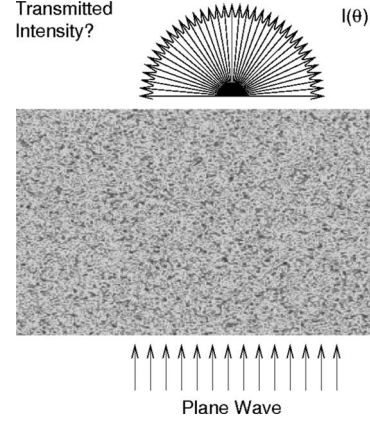

FIG. 1. Position of the problem. A plane wave or a narrow beam are normally incident on a slab of random heterogeneous material with thickness z expressed in transport mean-free-path units. The transmitted intensity probability distribution is $I(\theta)$. The gray level fluctuations reflect the inhomogeneity of the random medium.

A. Model

The implementation of the CPP model for the study of multiple scattering of particles or waves has already been described in [14]. In [14], the authors were mostly interested in formulating the direct problem, i.e., in the way a CPP can be used to predict the output distribution of scattering angles in a scattering experiment. They demonstrated the ability of the CPP to model the multiple scattering of electrons and used some recursive integral equations to obtain the pdf of the CPP. The CPP model introduced in [14] is based on the cumulative scattering angle, which is a real-valued random process defined as a sum of real-valued random variables. In this work, we introduce a CPP on a compact Lie group, namely, the rotation group $SO(3)$. Our approach is more general and does not rely on an *a priori* small-angle approximation.

Let us consider a particle or wave entering a random medium at time $t = 0$, with initial direction of propagation $\mu_0 \in \mathcal{S}^2$, as illustrated in Fig. 1. For the moment we neglect the possibility that the wave may escape the random medium. We consider a slab of random material with mean free time τ and velocity $c = 1$ for simplicity. The normalization of the velocity implies $\tau = \ell$ where ℓ is the mean free path. We seek to model the time evolution of the direction of propagation $\mu(t)$. Denoting by $N(t)$ the random number of scattering events that have occurred after propagation during time t in the random medium, the direction of propagation $\mu(t)$ can be written as

$$\mu(t) = \prod_{n=0}^{N(t)} \mathbf{r}_n \mu_0.$$

(16)

To model wave scattering, $N(t)$ is chosen as a Poisson process with parameter Λ [15] and is independent of the $\mathbf{r}_n$. To be fully consistent, we must impose $\mathbf{r}_0 = \mathbf{I}$, where $\mathbf{I}$ denotes the identity in $SO(3)$. This simply means that the unscattered energy propagates in the direction imposed by the source. Clearly, the $\mathbf{r}_n$, $n \geq 1$ are independent and identically distrib-

uted random rotations described by the pdf $f_{\mathbf{r}}$. Our choice for $N(t)$ implies the usual exponential distribution of times of flight between two scattering events, with parameter $\Lambda = 1/\tau$.

B. Probability density function of the CPP

To complete our task, we need to relate the pdf of $\mu(t)$, denoted by $p_{\mu(t)}$, to the scattering properties of the medium. Keeping in mind the *iid* assumption for $\mathbf{r}_n$ and using conditional probability decomposition with respect to the Poisson process $N(t)$, one gets

$$p_{\mu(t)}(\theta) = \sum_{n=0}^{\infty} e^{-\Lambda t}\frac{(\Lambda t)^n}{n!}(f_{\mathbf{r}}^{\otimes n} * w_{\mu_0})(\theta), \quad (17)$$

where $\otimes n$ denotes the *n*-times convolution, $f_{\mathbf{r}}$ is the common distribution of the $\mathbf{r}_n$, and w_{μ_0} is the distribution of μ_0. Making use of the notation introduced in Sec. II A for the multiple action of $\mathbf{r}_i$: $\mu_n = \mathbf{r}_n \mathbf{r}_{n-1} \cdots \mathbf{r}_2 \mathbf{r}_1 \mathbf{r}_0 \mu_0$, the pdf of μ_n can be rewritten as

$$f_{\mu_n}(\theta) = (f_{\mathbf{r}}^{\otimes n} * w_{\mu_0})(\theta). \quad (18)$$

Assuming again that the distribution of μ_0 is a Dirac at the north pole, the pdf of $\mu(t)$ simplifies to

$$p_{\mu(t)}(\theta) = \sum_{n=0}^{\infty} \frac{(\Lambda t)^n}{n!} e^{-\Lambda t} f_{\mathbf{r}}^{\otimes n}(\theta). \quad (19)$$

The assumption about μ_0 is not very strong since in many practical cases the direction of propagation of the incoming wave is either known or controlled. For an isotropic random medium, we recall that the pdfs of $\mathbf{r}_i$ can be considered zonal. Therefore, the pdf $f_{\mathbf{r}}$ can be expanded in a Legendre series with coefficients denoted by $\hat{f}_k$. Using Eq. (15) the pdf of $\mu(t)$ takes the form

$$p_{\mu(t)}(\theta) = \sum_{n=0}^{\infty} \frac{(\Lambda t)^n}{n!} e^{-\Lambda t} \sum_{l \geq 0} (2l+1)(\hat{f}_l)^n P_l(\cos\theta)$$

$$= e^{-\Lambda t} \sum_{l \geq 0} (2l+1) e^{\Lambda t \hat{f}_l} P_l(\cos\theta). \quad (20)$$

This provides an expression of the distribution of $\mu(t)$ of the form

$$p_{\mu(t)}(\theta) = \sum_{l \geq 0} (2l+1) e^{\Lambda t(\hat{f}_l - 1)} P_l(\cos\theta). \quad (21)$$

Expanding $p_{\mu(t)}$ in a Legendre series

$$p_{\mu(t)}(\theta) = \sum_{l \geq 0} (2l+1)\hat{\mu}_l P_l(\cos\theta), \quad (22)$$

a term-by-term identification in Eqs. (21) and (22) yields

$$\hat{\mu}_l = e^{\Lambda t(\hat{f}_l - 1)}. \quad (23)$$

Result (23) relates directly the Legendre coefficients $\hat{\mu}_l$ of the pdf of the scattering process $\mu(t)$ to the Legendre coefficients $\hat{f}_l$ of the pdf of the random rotations caused by the

scatterers. This close link between the Fourier coefficients of these pdfs is the cornerstone of our approach as it enables us to develop a statistical estimator of the Fourier coefficients $\hat{f}_l$ from the observations of the pdf of $\mu(t)$.

IV. VALIDITY OF THE CPP MODEL

A. Models of random media

To show the applicability of our statistical approach, we consider the following experiment (see Fig. 1): a plane wave or narrow beam is normally incident on a slab of random material described by Von Karman or Gaussian spectra, which are commonly used to describe geophysical media [16]. The angular pattern of intensity transmitted through the slab is measured on output. In the slab geometry with strong forward scattering, we expect that most of the energy will be collected in the forward direction and therefore the CPP should give some useful predictions of the pattern of transmitted intensity. In fact, this idea is confirmed by the agreement between CPP formula (21) and small-angle scattering approximations of the radiative transfer equation [6]. It is to be noted that formula (21) was obtained under the hypothesis that we are able to collect all the energy at a given time t but is totally free from a small-angle scattering assumption.

In the CPP approach, the building block for multiple scattering is the pdf of the random rotations that scatter the initial direction of the incoming beam. In the language of scattering theory, the function that describes the angular scattering pattern is called the phase function and will be denoted by $f(\mu)$, where $\mu = \cos\theta$ is the cosine of the scattering angle and does not represent a vector in this section. For sufficiently weak fluctuations, the phase function can be obtained from the Born approximation [4] and depends of the power spectrum of heterogeneities Φ as follows:

$$f(\mu) = \Xi\Phi[2k\sin(\theta/2)], \quad (24)$$

where k denotes the central wave number of the probing wave and Ξ is a normalization constant. For the sake of clarity, we review basic results on random media of special interest.

1. Gaussian media

The phase function of a Gaussian random medium is defined by the following formula:

$$f(\mu) = \frac{k^2 a^2}{2(1 - e^{-k^2 a^2})} e^{-k^2 a^2[(1-\mu)/2]}, \quad (25)$$

where a is the correlation length of heterogeneities. The anisotropy parameter or mean cosine of the scattering angle is defined as

$$g = \int_{-1}^{1} f(\mu)\mu \, d\mu. \quad (26)$$

This parameter plays a crucial role in the definition of the transport mean free path $l^* = l/(1-g)$. The transport mean free path is the typical length scale beyond which multiple scattering can be described by a diffusion equation. In the

common case of strong forward scattering it can be much larger than the mean free path. As will be shown below, our approach is useful in this regime. For Gaussian random media, one obtains $g = \coth(\frac{k^2 a^2}{2}) - \frac{2}{k^2 a^2}$. The Legendre coefficients of the phase function are given by

$$\hat{f}_l = \int f(\mu) P_l(\mu) d\mu. \tag{27}$$

In the high-frequency limit (large ka), the Legendre coefficients of the Gaussian phase function can be approximated as

$$\hat{f}_l \approx g^{l(l+1)/2}, \tag{28}$$

with $g \approx 1 - 1/2k^2 a^2$.

2. Von Karman media

The Von Karman spectrum implies the following angular dependence of the phase function:

$$f(\mu) = \frac{2k^2 a^2 (\alpha - 1)}{[1 - (4k^2 a^2 + 1)^{1-\alpha}][1 - 2k^2 a^2 (\mu - 1)]^\alpha}, \tag{29}$$

where α is an exponent which controls the small-scale roughness of the medium. Note that the phase function is normalized $[\int_{-1}^{1} f(\mu) d\mu = 1]$. The special cases $\alpha = 3/2, 2$ correspond to the well-known Henyey-Greenstein (HG) and exponential phase functions, respectively.

Using integration by parts, the coefficients of the expansion can be expressed as

$$\hat{f}_l = \frac{1}{2^n n!} \int_{-1}^{1} (1 - \mu^2)^l f^{(l)}(\mu) d\mu, \tag{30}$$

where $f^{(l)}(\mu)$ denotes the lth derivative of the phase function. Using tables of integrals, Eq. (30) yields the following compact form for the Legendre coefficients of the Von Karman correlation function:

$$\hat{f}_l = \frac{2\sqrt{\pi}(\alpha - 1)(1 + 2ka^2)^{-(l+\alpha)} \Gamma(l + \alpha)}{\Gamma\left(\frac{3}{2} + l\right) \Gamma(\alpha)[1 - (1 + 4ka^2)^{1-\alpha}]}$$
$$\times \, _2F_1\left[\frac{l+\alpha}{2}, \frac{1+l+\alpha}{2}; \frac{3}{2} + l; \frac{4ka^4}{(1 + 2ka^2)^2}\right], \tag{31}$$

where the definition of the hypergeometric function $_2F_1$ can be found in [17]. As noted before, in the case $\alpha = 3/2$, we recover the HG phase function which is a classically used approximation to the Mie theory for spherical scatterers. The Legendre coefficients can be put in the form $\hat{f}_l = g^l$, with $g = (1 + 2k^2 a^2 - \sqrt{1 + 4k^2 a^2})/2k^2 a^2$, i.e., the Legendre coefficients are simply powers of the anisotropy parameter, valid for any $g \in [0, 1[$. A remarkable property of HG random media is that the convolution on the sphere of unit directions of n HG phase functions with same parameter g is a itself a HG phase function with parameter g^n:

$$f_{HG}^{\otimes n}(\mu) = \sum_{l=0}^{\infty} (2l + 1) g^{ln} P_l(\mu). \tag{32}$$
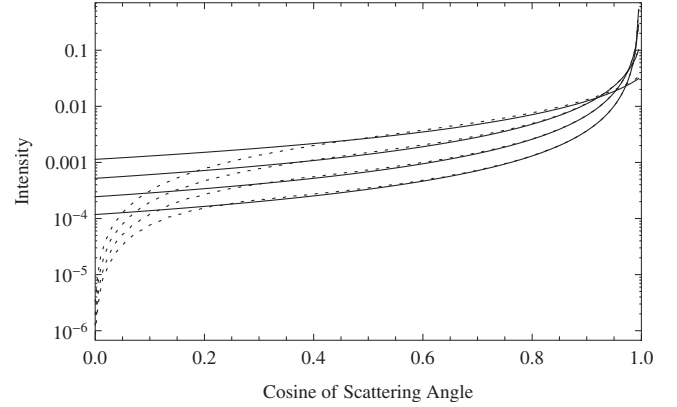

FIG. 2. Comparison between Monte Carlo simulations (dots) and CPP model for the intensity transmitted through a slab of random medium with Heynyey-Greenstein phase function. The four curves correspond to increasing slab thickness $H = 25/8, 25/4, 25/2, 25\ell$, with ℓ as the mean free path. The agreement is very good in a cone around the forward direction and degrades as the scattering angle approaches grazing angles.

B. Numerical results

To appreciate the strengths and weaknesses of our statistical model we show direct comparisons between the CPP and the radiative transfer equation for a Henyey-Greenstein random medium in the strong forward-scattering regime. The anisotropy factor is $g = 0.98$ and the slab thickness takes the values $H = 25/8, 25/4, 25/2, 25\ell$, where ℓ is the mean free path of the waves in the random medium (the transport mean free path ℓ^* equals 50ℓ). For simplicity, we assume that the index of the slab matches that of the surrounding medium. The equation of radiative transfer is solved using the Monte Carlo code developed by [18] in the framework of statistically anisotropic media. Figure 2 shows that the CPP model gives a nonuniform approximation to the exact distribution of intensity transmitted through the slab of random material. It is noticeable that the CPP model is always wrong in a cone of directions perpendicular to the direction of incidence of the wave. The width of the cone of direction increases with the slab thickness H. When H becomes of the order of the transport mean free path, typically, the CPP model fails to give reliable predictions of the angular distribution of intensity. One obvious reason is that in the CPP model, we do not prescribe any boundary condition, whereas in the slab geometry, the intensity must of course vanish in the lower hemisphere of propagation directions.

We further remark that in the slab geometry the distribution of scattering events does not follow a simple Poisson law, which also limits the validity of our model. To illustrate this point, we examine the validity of the model in Fig. 3, where we show the comparison between the CPP and radiative transfer theory for a Gaussian medium with anisotropy parameter $g = 0.98$. The slab thickness takes the values $H = 25/16, 25/8, 25/4, 25/2\ell$. The distribution of transmitted intensity differs markedly from the HG case (see Fig. 2 for comparison). Because the Gaussian medium is extremely smooth, the single scattering pattern is concentrated in a small cone of directions and large-angle scattering is ex-
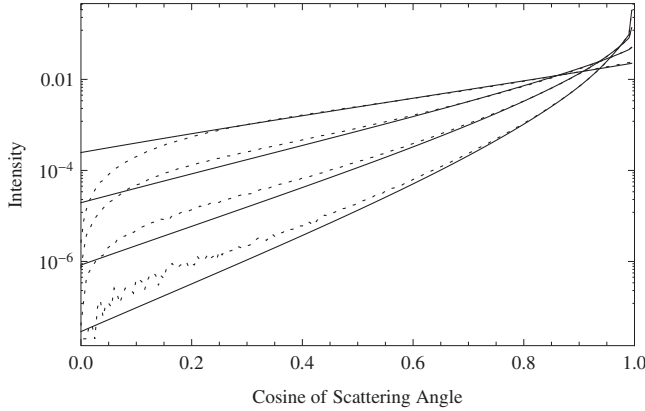
FIG. 3. Comparison between Monte Carlo simulations (dots) and CPP model for the intensity transmitted through a slab of random medium with Gaussian phase function. The four curves correspond to increasing slab thicknesses $H=25/16, 25/8, 25/4, 25/2\ell$, with ℓ as the mean free path. The agreement is very good in a cone around the forward direction and degrades as the scattering angle approaches grazing angles.

cluded. Propagation at large angle can only occur for sufficiently high-order multiple scattering. This entails a significant deviation of the statistics of the number of scattering events from the simple Poisson distribution. This point is illustrated in Fig. 4, where we plot the distribution of scattering events around the forward direction $(0.8<\mu<1)$ and at large angles $(0.3<\mu<0.5)$ for particles propagating through a slab of thickness $H=25/8\ell$. While in the forward direction the agreement between the calculated distribution of number of scattering events and Poisson statistics is excellent, at large angles a clear discrepancy is observed. The distribution is biased toward larger number of scattering events, and we verified empirically that it is in fact much better fitted by a normal distribution. The standard deviation and mean of the Gaussian distribution have been adjusted by trial and error. The main point is to illustrate the large devia-
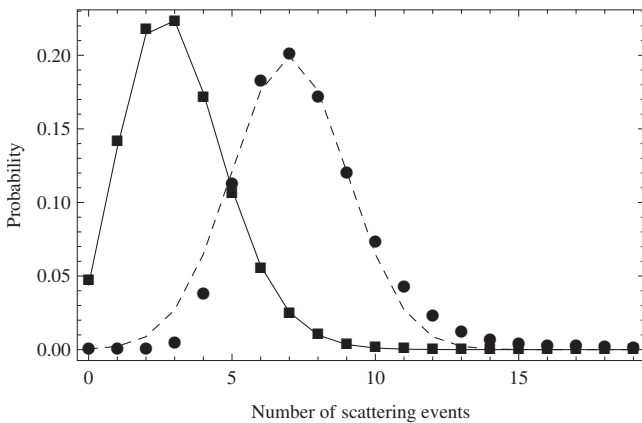

FIG. 4. Distribution of the number of scattering events for the intensity transmitted through a slab of Gaussian random medium with thickness $H=3.125\ell$. Squares: $(0.8<\mu<1)$; dots: $(0.3<\mu<0.5)$. Solid line: Poisson distribution with parameter $\Lambda=3.125$. Dashed line: Gaussian distribution with mean $\mu=7$ and standard deviation $\sigma=2$.
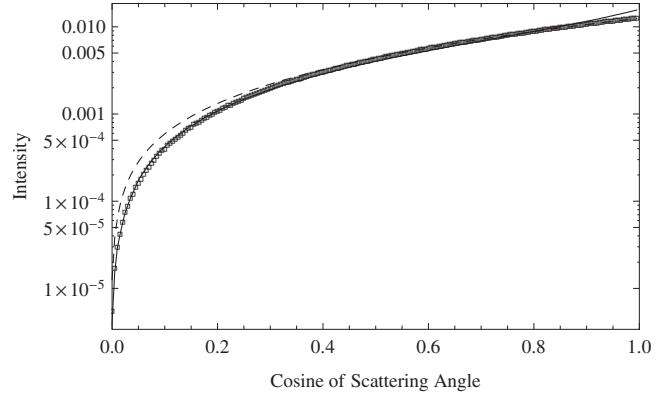

FIG. 5. Transmitted intensity profile through a slab of thickness $H=\ell^*/2$ and anisotropy parameter $g=0.98$. Squares: Gaussian random medium. Solid line: HG random medium. Dashed line: diffusion approximation. The difference between the Gaussian and HG medium is only noticeable close to the forward direction $(\mu=1)$.

tion from the Poisson distribution. Note however that the Poisson law will always be well approximated by a normal distribution (with equal mean and variance) if the number of scattering events becomes very large.

We now discuss briefly the transition toward the diffusive regime. The case of strongly anisotropic scattering has been treated analytically in [19], where it is shown that the pattern of transmitted intensity is almost universal, independent of the type of scatterers. In Fig. 5, we show the profile of transmitted intensity obtained with the radiative transfer approach with the slab thickness $H=\ell^*/2$ for the Gaussian and HG phase functions with the same anisotropy parameter $g=0.98$. The curves for the two different phase functions are hardly distinguishable. For comparison, we have also plotted the simple outcome of the diffusion approximation $I(\mu)=\mu(1/2+3\mu/2)$. Our numerical study confirms the general conclusions of [19] and also demonstrates that the diffusive regime sets in for a slab thickness of the order of ℓ^* but not more.

In spite of all its deficiencies, the CPP model gives a simple, almost analytical solution of the multiple-scattering problem and explains accurately the angular distribution of most of the forward-scattered energy.

V. ESTIMATION: THE INVERSE PROBLEM

In this section, we consider the problem of estimating the Legendre coefficients of the phase function f of the scatterers from the measurement of the output angular distribution $p_{\mu(T)}$ at a given depth in the random medium. Using the CPP model, we can construct an estimator for the Legendre coefficients of the phase function $\hat{f}_l$, using a "decompounding" formula. Note that the mean free time τ or equivalently (thanks to velocity normalization) the mean free path ℓ has to be known in the decompounding approach. Techniques to estimate the mean free path of waves in transmission geometry are described for instance in [20]. The estimators we derive below are extensions of the work done in [21,22] for real valued random variables to the specific case of random

variables taking values over the rotation group and having zonal pdfs.

A. Estimator for the Legendre coefficients

After proper normalization, the transmitted intensity profile can be interpreted as a pdf $p_{\mu(T)}$, where T is the time required for ballistic waves to propagate through the slab. Let us denote by $\bar{p}_{\mu(T)}$ the *sample* pdf that corresponds to the data. The sampling resolution of $\bar{p}_{\mu(T)}$ depends on the acquisition system. Typically, a detector has finite aperture and therefore averages the intensity over some finite solid angle $d\Omega$. This implies that $\bar{p}_{\mu(T)}$ is in fact discrete and should be more appropriately termed *probability mass function* with the usual normalization $\sum_{m=1}^{M}\bar{p}_{\mu_m(T)}=1$ if M samples are available.

We are interested in estimating the Legendre coefficients of the pdf $p_{\mu(T)}$ from the observation of $\bar{p}_{\mu_m(T)}$. This can be achieved in two different ways. One can either define directly the estimator on $\bar{p}_{\mu(T)}$ or use $\bar{p}_{\mu(T)}$ to generate some realizations of $\mu(T)$ and define an estimator with these realizations. In this paper, we adopt the second approach. To do so, we use a simple linear interpolation to evaluate numerically the cumulative distribution function (cdf) and its inverse from the data $\bar{p}_{\mu(T)}$. This allows us to generate as many realizations of $\mu(T)$ as desired by evaluating the inverse cumulative distribution function for a collection of uniformly distributed random numbers in $[0,1]$.

Let us assume that we have generated K values of $\mu(T)$ using this technique, and let us denote them by $\mu_\kappa(T)$ with $\kappa=1,\ldots,K$. The empirical Legendre coefficients of $p_{\mu(T)}$ denoted by $\widetilde{\hat{\mu}}_l$ are given by

$$\widetilde{\hat{\mu}}_l = \frac{1}{K}\sum_{\kappa=1}^{K}P_l(\mu_\kappa(T)). \qquad (33)$$

Note that $\mu_\kappa(T)$ denotes the cosine of scattering angle θ_κ. It can be verified that the estimator defined in Eq. (33) is unbiased, i.e., $\mathbb{E}[\widetilde{\hat{\mu}}_l-\mu_l]=0,\forall l$. Its variance can be expressed as

$$\mathbb{E}[(\widetilde{\hat{\mu}}_l-\mathbb{E}[\widetilde{\hat{\mu}}_l])^2] = \frac{1}{K}(\mathbb{E}[P_l^2(\mu)]-\hat{\mu}_l^2), \qquad (34)$$

where $\mathbb{E}[P_l^2(\mu)]=\int_{-1}^{1}p_{\mu(T)}(\mu)P_l^2(\mu)d\mu$ and $\mu=\cos\theta$.

In the Gaussian case, the variance takes a simple analytical form, in the limit case where $(g-1)\ll 1$. Using a Taylor series expansion of the Legendre polynomials near $\mu=1$ yields $\mathrm{Var}(\widetilde{\hat{\mu}}_l)=\frac{1}{K}\frac{l^2(l+1)!^2}{4}(g-1)^2$.

For other random media, such as Henyey-Greenstein or exponential, an analytical expression of the variance is not easily obtained. This also applies to Gaussian media away from the limiting case mentioned above. Nevertheless, the behavior of the variance can be obtained by numerical integration. In Fig. 6, we present the variance for three different phase functions: exponential, Gaussian, and Henyey-Greenstein. The estimation has been made using $K=1000$ samples. It can be seen on Fig. 6 that the variance of the estimator increases monotonically with l for the three differ-
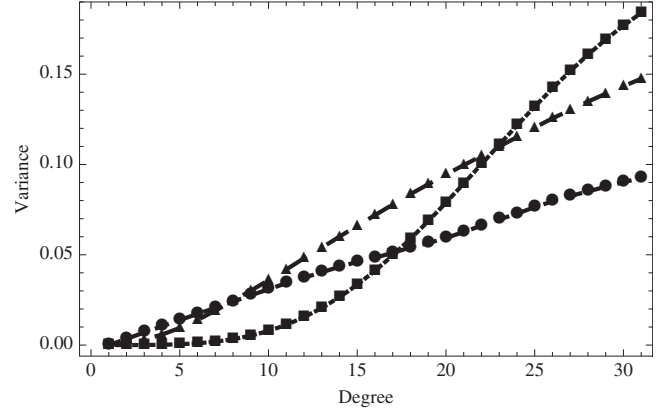

FIG. 6. Variance of the Legendre coefficients empirical estimator for different phase functions: Gaussian (squares), exponential (triangles), and Henyey-Greenstein (circles) with anisotropy parameter $g=0.995$.

ent phase functions. While the Gaussian case follows a fourth degree polynomial growth, Henyey-Greenstein, and exponential cases follow a more or less linear increase with the degree of the Legendre coefficients l. The Gaussian case is favorable for low degree coefficients while Henyey-Greenstein exhibits the lowest variance for higher degrees. Note that the variance has consequences on the ability of the proposed approach to evaluate correctly the high-degree Legendre coefficients. This will be illustrated in Sec. V B.

Back to the decompounding problem, we wish to estimate the Legendre coefficients of f from the samples $\mu_\kappa(T)$. This is possible by inverting Eq. (23) and using our empirical estimator [Eq. (33)]. We obtain an estimate of the Legendre coefficients of the phase function of the medium:

$$\widetilde{\hat{f}}_l = \frac{1}{\Lambda T}\ln\widetilde{\hat{\mu}}_l + 1. \qquad (35)$$

Equation (35) is the decompounding formula. It implies that the set of Legendre coefficients $\widetilde{\hat{f}}_l$ can be estimated from the set $\widetilde{\hat{\mu}}_l$ that is directly computed from the data thanks to Eq. (33). Note that it is necessary that $\widetilde{\hat{\mu}}_l > 0$. This is verified for small l ($\mu_0 > \mu_l, \forall l \geq 1$) but may not be verified for higher degrees, mainly because of the increase in the variance of the estimator $\widetilde{\hat{\mu}}_l$ with l (see Fig. 6). As a consequence, it will be necessary to truncate the number of Legendre coefficients used in the decompounding approach. The truncation degree depends on the acquisition setup and the kind of medium investigated. This will be further illustrated in the following Section. It must be emphasized that Eq. (35) is central as it allows direct estimation of the heterogeneity spectrum of the medium from the output intensity distribution measurement.

B. Inverting for the random medium power spectrum

As we will now show, the CPP approach constitutes a powerful tool to invert the power spectrum of a random medium in the regime: $\ell < H < \ell^*$, particularly in the regime $\ell \ll \ell^*$. To demonstrate this, we consider a slab of material with Von-Karman correlation function. The cases $\alpha=2$ (ex-
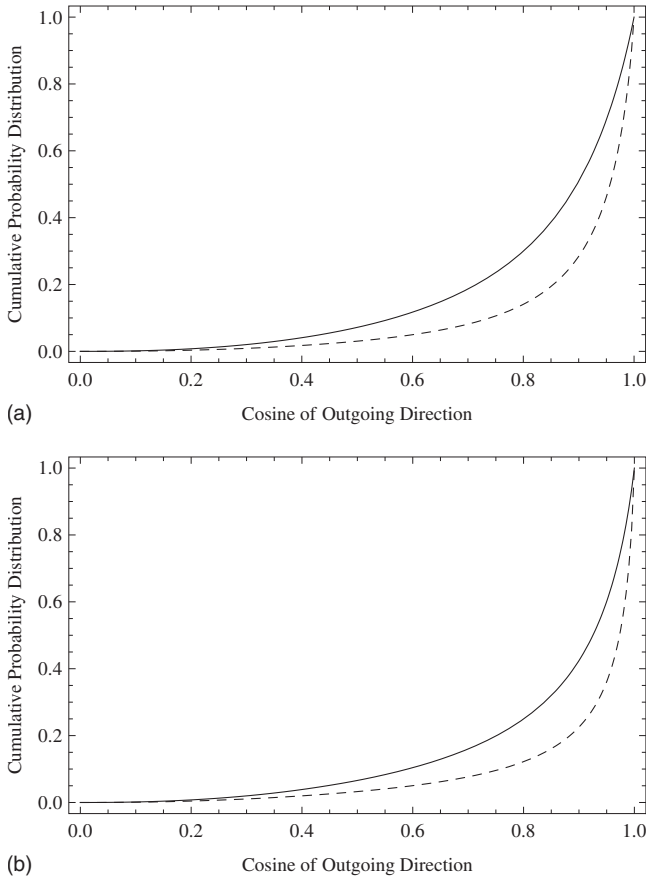
FIG. 7. Cumulative distribution function of the intensity transmitted through a slab of random material with thicknesses $H=6.25\ell$ (dashed line), 12.5ℓ (solid line), and anisotropy parameter $g=0.98$. Top: exponential medium. Bottom: "Kolmogorov" medium.

ponential correlation) and $\alpha=5/3$ (Kolmogorov spectrum) are considered. Synthetic data are generated using Monte Carlo simulations of the scalar radiative transfer equation in 3D. It is to be noted that the simulations provide the intensity distribution averaged over some finite solid angle, as will be the case in a practical experiment. In the numerical simulations, we imposed $d\Omega=1/200$ to obtain sufficiently smooth results. Due to the limited averaging, the synthetic data can be noisy, particularly where the intensity is small but such phenomenon is also expected in practice. To obtain the Legendre coefficients of the intensity distribution, we process the data as follows. From the numerical simulations, we determine the cumulative distribution of the intensity. Although the original distribution can be noisy the cumulative distribution is much smoother. Examples of empirical cdf for exponential and Kolmogorov spectra are shown in Fig. 7. From the cumulative distribution, we can evaluate the coefficients of the Legendre expansion of the intensity by drawing a sufficiently large set of random numbers X according to

$$X = \mathrm{icdf}(U), \qquad (36)$$

where U is a uniformly distributed random number and icdf denotes the intensity inverse cumulative distribution function
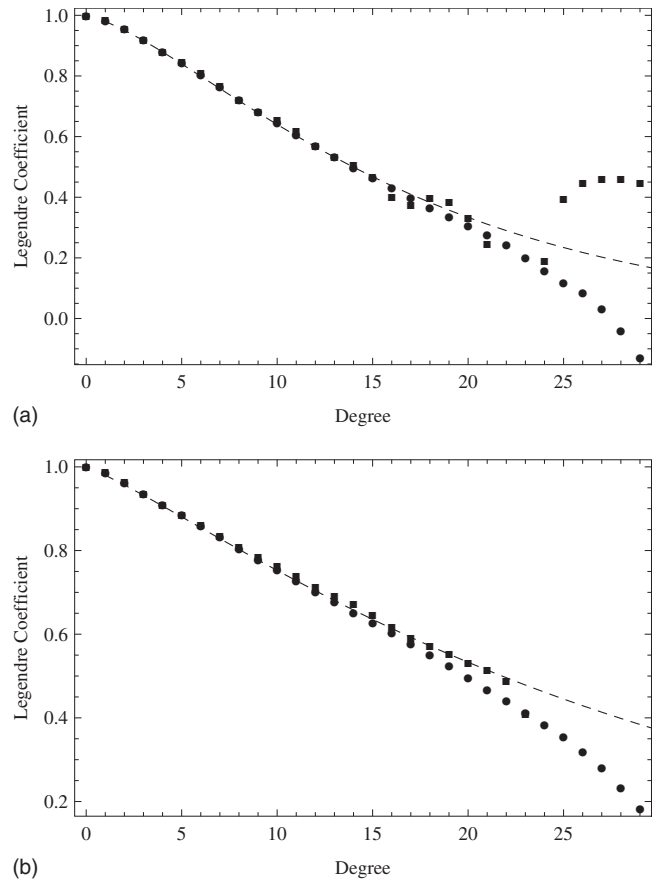

FIG. 8. Estimated Legendre coefficients of the heterogeneity power spectrum for two slabs of random material with thicknesses $H=6.25\ell$ (dots), 12.5ℓ (squares), and anisotropy parameter $g=0.98$. Top: exponential medium. Bottom: Kolmogorov medium. The dashed line shows the exact Legendre coefficients of the distributions.

obtained by linear interpolation of the original cdf. Successive application of formulas (33) and (35) yields the coefficients $\widetilde{\mu}_l$ and the desired Legendre coefficients $\widetilde{\hat{f}}_l$ of the power spectrum of heterogeneities. To obtain accurate estimates we need to draw typically 10^5 random samples. The comparison between the estimated coefficients and the exact Legendre expansion for two different Von-Karman random media with $g=0.98$ is shown in Fig. 8 and shows excellent agreement. To check the accuracy of the decompounding formula, it is important to analyze at least two sets of data corresponding to two different slab thicknesses. First, this offers a consistency check for the only parameter that enters in the decompounding, i.e., the mean free path of the waves. If there is a significant error on this quantity (typically more than 20%), the estimated Legendre coefficients of the power spectrum will differ significantly at low degree l. Second, it provides a test of validity of the inferred heterogeneity power spectrum at larger degree l. Figure 8 shows that the estimated $\widetilde{\hat{f}}_l$ for two different slab thicknesses split up beyond some degree l_0. For $l>l_0$, it is not possible to estimate reliably the coefficients of the original distribution. It is to be noted that this test is independent of any assumption on the random medium.

VI. CONCLUSION

We have developed a nonparametric method to infer the properties of random media. Our approach relies on a generic mathematical model and could in principle be used to probe random media with acoustic, elastic, or electromagnetic waves. The method also provides tests of consistency and accuracy of the results. We show that the angular distribution of intensity in a random medium can be decompounded to estimate the power spectrum of heterogeneities in a random medium. An extension of the theory to incorporate polarization measurements is currently being investigated.

[1] S. Sinha, in *Methods in the Physics of Porous Media*, Experimental Methods in the Physical Sciences Vol. 35, edited by P.-Z. Wong (Academic, San Diego, 1999).

[2] R. Xu, *Particle Suspensions: Light Scattering Methods* (Kluwer Academic, New York, 2002).

[3] H. Sato and M. C. Fehler, *Seismic Wave Propagation and Scattering in the Heterogeneous Earth* (Springer, Berlin, 1998).

[4] S. Rytov, Y. Kravtsov, and V. Tatarskii, *Principles of Statistical Radiophysics: Elements of Random Fields* (Springer, Berlin, 1989), Vol. 3.

[5] E. D. Hirleman, Appl. Opt. **30**, 4832 (1991).

[6] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978), Vol. 1 and 2.

[7] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*, Interdisciplinary Applied Mathematics Vol. 16 (Springer, New York, 2002).

[8] U. Grenander, *Probabilities on Algebraic Structures* (Wiley, New York, 1963).

[9] J. Dieudonné, *Special Functions and Linear Representations of Lie Groups* (American Mathematical Society, Providence, RI, 1980).

[10] R. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups* (Springer, New York, 1985).

[11] S. Said and N. Le Bihan, Waves Random Complex Media **18**, 275 (2008).

[12] P. Kostelec and D. Rockmore, J. Fourier Anal. Appl. **14**, 145 (2008).

[13] S. Said, C. Lageman, N. Le Bihan, and J. Manton, *Eighth IMA Conference on Mathematics in Signal Processing* (IMA, Cirencester, UK, 2008).

[14] X. Ning, L. Papiez, and G. Sandison, Phys. Rev. E **52**, 5621 (1995).

[15] S. Karlin and H. Taylor, *A Second Course in Stochastic Processes* (Academic, New York, 1981).

[16] L. Klimeš, Pure Appl. Geophys. **159**, 1811 (2002).

[17] N. Temme, *Special Functions: An Introduction to the Classical Functions of Mathematical Physics* (Wiley, New York, 1996).

[18] L. Margerin, Tectonophysics **416**, 229 (2006).

[19] E. Amic, J. M. Luck, and T. M. Nieuwenhuizen, J. Phys. A **29**, 4915 (1996).

[20] M. L. Cowan, K. Beaty, J. H. Page, Z. Liu, and P. Sheng, Phys. Rev. E **58**, 6626 (1998).

[21] B. van Es, S. Gugushvili, and P. Spreij, Bernoulli **13**, 672 (2007).

[22] B. Buchmann and R. Grübel, Ann. Stat. **31**, 1054 (2003).