

# Inférence statistique sur la fonction de phase des milieux aléatoires par diffusion multiple des ondes

Nicolas LE BIHAN<sup>1</sup>, Ludovic MARGERIN<sup>2</sup>, Salem SAID<sup>1</sup>

<sup>1</sup>GIPSA-Lab, DIS, CNRS UMR 5216  
961 Rue de le Houille Blanche, Domaine Universitaire, BP 46, 38402 Saint Martin d'Hères, France

<sup>2</sup>CEREGE, CNRS UMR 6635  
Europôle Méditerranéen de l'Arbois, BP 80, 13545 Aix En Provence cedex 04, France  
nicolas.le-bihan@gipsa-lab.inpg.fr, margerin@cerège.fr  
salem.said@gipsa-lab.inpg.fr

**Résumé** – Nous proposons une technique d'estimation paramétrique pour inférer sur un milieu aléatoire à l'aide du signal de sortie issu de la propagation d'une onde dans le milieu. Utilisant un modèle de processus de Poisson composé pour le signal de sortie, nous proposons un estimateur non-paramétrique de la fonction de phase du milieu. Les résultats de l'estimateur sont validés sur simulation Monte Carlo et le domaine de validité du modèle est également démontré.

**Abstract** – We propose a non parametric estimation technique for random media phase function estimation using the output signal collected after the propagation of a wave in the medium. Using a compound Poisson process model for the output signal, we derive a non-parametric estimation technique for the medium phase function identification. Results are validated on Monte Carlo simulations and the validity range of the model is investigated.

## 1 Introduction

Dans l'étude de la propagation des ondes en milieux aléatoires (milieux possédant des hétérogénéités réparties aléatoirement), un problème important consiste à essayer d'inférer sur le milieu à l'aide du signal de sortie obtenu après propagation d'une onde dans le milieu. Dans un tel milieu, l'onde qui se propage rencontre des diffuseurs qui affectent sa propagation (déviation de sa direction). En Physique, la théorie du transfert radiatif [1] permet de prédire l'intensité en sortie d'un tel milieu aléatoire constitué de diffuseurs, mais n'autorise pas une résolution simple problème inverse, *c.à.d* l'estimation des paramètres du milieu à l'aide du signal de sortie. En effet, dans beaucoup de cas réalistes, la théorie du transfert radiatif nécessite la mise en œuvre de solutions numériques lourdes, ce qui ne permet pas un résultat analytique du problème inverse.

Nous présentons donc une approche permettant l'estimation des paramètres du milieu. Nous proposons un modèle de type Processus de Poisson Composé (PPC) sur le groupe des rotations dans  $\mathbb{R}^3$  pour décrire le signal recueilli en sortie d'un milieu aléatoire. Cette approche permet, grâce à une relation directe entre la fonction caractéristique de la distribution en sortie et celle de la *fonction de phase* des diffuseurs, d'obtenir un estimateur non-paramétrique pour cette fonction qui caractérise le milieu aléatoire. Nous présentons donc, après quelques résultats d'analyse harmonique sur le groupe des rotations, le modèle PPC pour la diffusion multiple. Après une validation du modèle

par comparaison avec la solution Monte Carlo du transfert radiatif, nous présentons l'estimateur de la fonction de phase et illustrons son comportement sur des signaux issus de simulation numérique de diffusion multiple.

## 2 Analyse harmonique sur $SO(3)$

Dans cette section, nous présentons rapidement quelques résultats d'analyse harmonique non-commutative dans le cas particulier du groupe de rotation  $SO(3)$ . Plus de détails peuvent être trouvés dans [3, 10].

### 2.1 Transformée de Fourier sur $SO(3)$

Nous nous intéressons aux fonctions de  $L^2(SO(3), \mathbb{C})$ . Ces fonctions sont intégrables au sens de la mesure de Haar de  $SO(3)$ . Nous utilisons pour paramétrer  $SO(3)$  la convention XZX pour les angles d'Euler  $\psi, \varphi \in [0, 2\pi)$  et  $\theta \in [0, \pi]$ . Le théorème de Peter-Weyl [3] permet d'écrire une fonction  $f(\psi, \theta, \varphi) \in L^2(SO(3), \mathbb{C})$  comme :

$$f(\psi, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m,n=-l}^{+l} F_l^{mn} D_{mn}^l(\psi, \theta, \varphi) \quad (1)$$

Les fonctions de base  $D_{mn}^l(\psi, \theta, \varphi)$  prennent la forme :

$$D_{mn}^l(\psi, \theta, \varphi) = e^{i(\psi m + \varphi n)} d_{mn}^l(\cos \theta) \quad (2)$$

où les  $d_{mn}^l(\cos \theta)$  sont les fonctions d-Wigner (ou polynômes de Jacobi). Les “coefficients de Fourier”  $F_l^{mn}$  sont donnés par :

$$F_l^{mn} = \int_{\psi} \int_{\theta} \int_{\varphi} f(\psi, \theta, \varphi) D_{mn}^{l\dagger}(\psi, \theta, \varphi) d\mu(\psi, \theta, \varphi) \quad (3)$$

Les  $F_l^{mn}$  peuvent être également représentés sous forme la forme de matrices  $\mathbf{F}_l$  de taille  $(2l + 1) \times (2l + 1)$ .

## 2.2 Variables aléatoires sur $SO(3)$

Nous considérons les variables aléatoires à valeurs sur  $SO(3)$ . On peut les représenter par des matrices aléatoires  $G$  de  $\mathbb{R}^{3 \times 3}$  qui satisfont :  $GG^T = I$  et  $\det(G) = 1$ . Nous nous intéressons ici aux variables  $R$  qui possèdent une densité de probabilité (*pdf*)  $p \in L^2(SO(3), \mathbb{C})$ . Il est alors possible de développer cette *pdf* comme dans (1). La donnée des coefficients de ce développement (de “Fourier”) correspond à la fonction caractéristique. Dans la Section 4, nous utiliserons ces développements pour proposer une technique d’estimation non-paramétrique pour les *pdf* à valeurs sur  $SO(3)$ .

## 3 Modèle du signal de diffusion multiple

Nous souhaitons modéliser un processus de diffusion multiple d’une onde plane dans un milieu hétérogène aléatoire à géométrie de type “tranche”, illustrée sur la Figure 1. Nous utilisons ici une extension au cas de  $SO(3)$  du modèle de processus de Poisson composé proposé initialement par Ning *et al.* [5]. La grandeur physique qui nous intéresse est la direction de propagation de l’onde  $\mu(t)$  qui évolue au cours du temps de propagation dans le milieu.  $\mu(t)$  est un vecteur unitaire de  $\mathbb{R}^3$ , c’est à dire un élément de  $\mathcal{S}^2$ , la sphère unité.

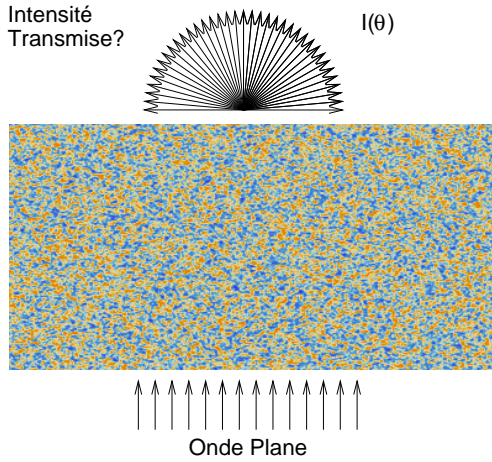


FIG. 1 – Géométrie de type “tranche” considérée pour la propagation en milieu aléatoire, avec une source de type onde plane.

## 3.1 Processus de Poisson composé sur $SO(3)$

La direction de propagation de l’onde  $\mu(t)$  est affectée lors de la propagation par la rencontre avec un diffuseur dont l’effet est de dévier aléatoirement  $\mu(t)$ . Le temps entre deux diffusions est également aléatoire et suit une loi exponentielle [4]. On peut alors écrire  $\mu(t)$  dans le milieu aléatoire comme :

$$\mu(t) = \prod_{i=1}^{N(t)} \mathbf{r}_i \mu_0 \quad (4)$$

où  $\mu_0$  est la direction de propagation initiale (à l’entrée dans le milieu), les  $\mathbf{r}_i$  sont des variables aléatoires *i.i.d.* à valeurs sur  $SO(3)$  et  $N(t)$  est un processus de Poisson de paramètre  $1/\tau = c/\ell$ , avec  $\tau$  le temps libre moyen,  $c$  la célérité des ondes dans le milieu et  $\ell$  le libre parcours moyen (distance moyenne entre deux diffusions). On prendra  $c = 1$  pour simplifier les expressions. L’expression (4) est celle d’un Processus de Poisson Composé (PPC) sur  $SO(3)$  dont on observe l’action sur  $\mathcal{S}^2$ . Dans la suite, nous allons nous intéresser au lien entre la *pdf* de  $\mu(t)$  et celle des  $\mathbf{r}_i$ . Avant cela, nous présentons comment le modèle PPC décrit la physique de la diffusion multiple.

## 3.2 Fonction de phase d’un milieu aléatoire

Pour modéliser comment les diffuseurs agissent sur  $\mu(t)$ , on utilise une *fonction de phase* qui est en fait la densité de probabilité des variables  $\mathbf{r}_i$ , notée  $p_{\mathbf{r}}$ . Cette fonction de phase est une fonction des angles d’Euler dans le cas général (diffuseurs à géométrie quelconque, anisotropie des diffuseurs). Dans le cas de diffuseurs statistiquement isotropes, la symétrie du problème entraîne que  $p_{\mathbf{r}}$  n’est fonction que du cosinus de l’angle de diffusion  $\theta$ . Dans notre étude, nous utilisons la fonction de phase Von Karman [6] pour modéliser le milieu aléatoire :

$$p_{\mathbf{r}}(\theta) = \frac{2ka^2(s-1)}{(1 - (4ka^2 + 1)^{1-s})(1 - 2ka^2(\cos \theta - 1))^s} \quad (5)$$

avec  $k$  le nombre d’onde de l’onde incidente,  $a$  la longueur de corrélation des hétérogénéités et  $s$  l’exposant décrivant la “rugosité” du milieu. Cette fonction inclut comme cas particuliers les fonctions Henyey-Greenstein ( $s = 3/2$ ) et exponentielle ( $s = 2$ ).

Pour de telles fonctions de phase, la décomposition (1) devient une décomposition en polynômes de Legendre de la forme :

$$p_{\mathbf{r}}(\theta) = \sum_{l=0}^{\infty} f_l P_l(\cos \theta) \quad (6)$$

où  $P_l(\cos \theta)$  est le polynôme de Legendre d’ordre  $l$  et où les  $f_l$  sont scalaires (cas particulier des  $F_l^{mn}$  donnés dans (3)). La forme compacte des  $f_l$  pour la fonction de phase Von Karman est :

$$f_l = \frac{2\sqrt{\pi}(s-1)(1+2ka^2)^{-(l+s)}\Gamma(l+s)}{\Gamma\left(\frac{3}{2}+l\right)\Gamma(s)(1-(1+4ka^2)^{1-s})} \times {}_2F_1\left(\frac{l+s}{2}, \frac{1+l+s}{2}; \frac{3}{2}+l; \frac{4ka^4}{(1+2ka^2)^2}\right) \quad (7)$$

où  ${}_2F_1(a; b; c; z)$  est la fonction hypergéométrique de Gauss (Chap. 15 dans [7]). Dans le cas Henyey-Greenstein, les coefficients  $f_l$  prennent la forme particulièrement simple :

$$f_l = \left( \int_{-1}^1 \cos \theta p_r(\theta) \sin \theta d\theta \right)^l$$

D'un point de vue Physique, il est intéressant d'identifier la fonction de phase d'un milieu, et cela peut être fait via l'estimation des  $f_l$ . Avant de proposer une méthode pour effectuer cette estimation, nous nous intéressons aux limites de validité du modèle PPC.

### 3.3 Validité du modèle

L'approche PPC pour la description de la diffusion multiple ne fait aucune hypothèse sur le milieu, et peut donc sembler assez générique. Mais, de fait, l'absence de la localisation des diffuseurs dans le modèle (pas de variable position dans le PPC) limite les cas de figure où le modèle s'applique. En effet, il est nécessaire de prendre en compte toute l'énergie présente dans le milieu à chaque temps  $t$  pour décrire n'importe quel type de diffusion multiple. Ce n'est pas le cas pour le PPC. Mais dans la géométrie tranche considérée ici (Fig. 1), et dans le cas de la diffusion vers l'avant (fonctions de phase très piquée autour de  $\cos \theta = 1$ ), l'énergie est propagée vers l'avant et le PPC décrit correctement le signal de sortie.

Pour quantifier la perte de l'information sur sa direction de propagation originale, on introduit le libre parcours moyen de transport  $\ell^* = \ell / (1 - \langle \cos \theta \rangle)$ , avec  $\langle \cos \theta \rangle$  la valeur moyenne du cosinus de l'angle de diffusion. Le modèle PPC permet de décrire de façon correcte le signal de sortie pour une épaisseur de tranche de l'ordre de  $\ell^*/2$ . Le PPC est donc pertinent dans le cas où  $\cos \theta \simeq 1$ . Dans ce régime on a  $\ell^* \gg \ell$ , ce qui correspond à de la diffusion multiple d'ordre élevé.

Sur la figure (2) nous comparons la solution (semi-analytique) donnée par le PPC avec la solution obtenue par le code de simulation Monte Carlo de diffusion multiple développé dans [2]. Les différentes courbes sont présentées pour des valeurs allant de  $\ell^*/24$  (courbes inférieures) et  $\ell^*/2$  (courbes supérieures). La fonction de phase utilisée est une Henyey-Greenstein de paramètre  $\langle \cos \theta \rangle = 0.95$ .

On peut voir que les courbes noires données par le PPC sont en accord avec les courbes rouges données par la solution Monte Carlo du transfert radiatif, pour une gamme importante d'angles de diffusion. Pour les grands angles ( $\mu = \cos \theta \rightarrow 0$ ), le PPC ne décrit pas bien l'énergie de sortie du milieu. Les courbes supérieures montrent que le domaine de validité du PPC est moins important quand la longueur de propagation (ou profondeur de pénétration de l'onde dans le milieu) se rapproche de  $\ell^*$ .

L'intérêt du PPC réside donc dans le fait qu'il autorise une très bonne description du phénomène de diffusion vers l'avant et surtout qu'il va nous permettre la résolution du problème inverse : estimer la fonction de phase à l'aide de la distribution de sortie de la direction de propagation  $\mu(t)$ . C'est ce que nous présentons dans la Section suivante.

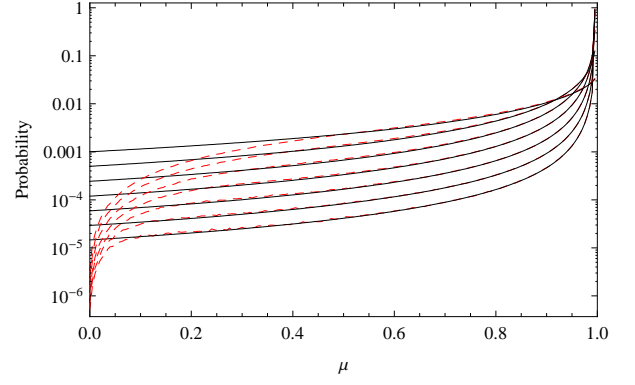


FIG. 2 – Comparaison entre la distribution en sortie du milieu aléatoire par le PPC (rouge) et la solution Monte Carlo du transfert radiatif (noir), pour différentes épaisseurs de milieu. Cas de la fonction de phase Henyey-Greenstein de paramètre  $\langle \cos \theta \rangle = 0.95$ .

## 4 Estimation non-paramétrique

Nous souhaitons inférer sur le milieu aléatoire à l'aide de la distribution des angles de diffusion en sortie du milieu, c'est à dire la distribution de  $\mu(t)$ , notée  $p_{\mu(t)}$ . Nous nous plaçons dans le cas où  $\mu(t)$  est de la forme (4) et où  $p_r$  peut s'écrire comme dans (5). En prenant que la distribution de  $\mu_0$  est un Dirac au pôle nord de  $S^2$ , on obtient directement à partir de (4) que :

$$\begin{aligned} p_{\mu(t)}(\theta) &= \sum_{k=0}^{\infty} \frac{(t/\tau)^k}{k!} e^{-t/\tau} p_r^{\otimes k}(\theta) \\ &= \sum_{k=0}^{\infty} \frac{(t/\tau)^k}{k!} e^{-t/\tau} \sum_{l \geq 0} (2l+1) (f_l)^k P_l(\cos \theta) \\ &= \sum_{l \geq 0} (2l+1) e^{\frac{t}{\tau}(f_l-1)} P_l(\cos \theta) \end{aligned} \quad (8)$$

avec  $f_l$  les coefficients de Legendre la fonction de phase. Il vient alors directement, en posant que  $p_{\mu(t)}$  a une expansion en polynômes de Legendre de la forme :

$$p_{\mu(t)} = \sum_{l \geq 0} g_l P_l(\mu(t)),$$

que l'on a :

$$f_l = \frac{\tau}{t} \ln g_l + 1 \quad (9)$$

Cette relation permet donc d'obtenir les coefficients de Legendre de la fonction de phase à l'aide de ceux de la distribution de  $\mu(t)$ . Cela permet de décomposer le PPC, c.à.d. dans (4) d'estimer la distribution des  $r_i$  à partir de celle de  $\mu(t)$ . Cette décomposition a été initialement proposée pour les PPC à valeurs réelles dans [8], et étendue au cas des groupes de Lie compacts dans [9].

Nous considérons ici le cas où les données à notre disposition sont des valeurs de  $\mu(T)$  recueillies à un temps  $T > 0$  après le début de la propagation de l'onde dans le milieu. Nous faisons l'hypothèse que nous possédons  $K$  échantillons dénotés  $\mu_\kappa$ ,  $\kappa = 1, \dots, K$ . On peut alors écrire l'estimateur empirique des coefficients de Legendre de  $p_{\mu(T)}$  qui prend la forme :

$$\hat{g}_l = \frac{1}{K} \sum_{\kappa=1}^K P_l(\mu_\kappa) \quad (10)$$

Cette expression est dérivée de (3) dans le cas particulier où les  $F_l^{mn}$  dégénèrent en  $P_l$ . L'estimateur (10) a une variance de la forme :  $K^{-1} (\mathbb{E}[P_l^2(\mu)] - g_l^2)$ .

En injectant l'estimateur (10) dans (9), on obtient un estimateur  $\hat{f}_l$  des coefficients de Legendre de la fonction de phase. Il est toutefois nécessaire de connaître  $\tau$  pour obtenir les  $\hat{f}_l$ , ce qui, expérimentalement, est possible via des mesures sur les ondes balistiques (non affectées par les diffuseurs). La connaissance des  $\hat{f}_l$  permet donc d'identifier le milieu de propagation et ce à l'aide seulement de la distribution des  $\mu_\kappa$ .

Afin d'illustrer le comportement de la technique de *décomposition* du PPC, nous présentons des résultats d'estimation des  $\hat{f}_l$ . Les échantillons  $\mu_\kappa$  ont été générés par le code Monte Carlo développé dans [2]. Nous présentons sur la figure 3 les  $\hat{f}_l$  obtenus avec (9) et (10) pour deux épaisseurs de propagation différentes :  $6.25\ell$  (bleu) et  $12.5\ell$  (rouge). La courbe noire représente les coefficients de Legendre théoriques utilisés pour générer les données. On peut voir que la technique de *décomposition* permet de retrouver une grande partie des coefficients de la fonction de phase, ce qui assure que la modélisation PPC est un outil intéressant pour l'étude des données de diffusion multiple. On voit également que lorsque les ondes se sont propagées plus longtemps dans le milieu (courbe rouge), l'estimation de la fonction de phase est nettement moins bonne, en particulier sur les coefficients de Legendre d'ordres élevés. Ceci est dû à l'uniformisation de la distribution des  $\mu_\kappa$  qui ne permet pas d'obtenir une bonne *décomposition*. D'une manière générale, les coefficients de Legendre d'ordre élevé sont mal estimés par la méthode, et ceci à cause de la variance de l'estimateur  $\hat{f}_l$ .

Néanmoins, la technique de *décomposition* du PPC permet d'obtenir une information importante sur la fonction de phase des diffuseurs et permet donc une caractérisation des milieux diffusants qu'il n'est pas possible d'obtenir avec une approche "transfert radiatif".

## 5 Conclusion

La modélisation de la diffusion multiple des ondes à l'aide du PPC permet une identification de la fonction de phase du milieu, dans le cas où les angles de diffusion sont petits (diffusion vers l'avant). Cette approche est valable pour des épaisseurs de matériau de l'ordre de  $\ell^*/2$ , pouvant correspondre à un régime de diffusion multiple à plusieurs  $\ell$ . Dans ce cadre, une approche "fréquentielle" basée la relation entre la fonction ca-

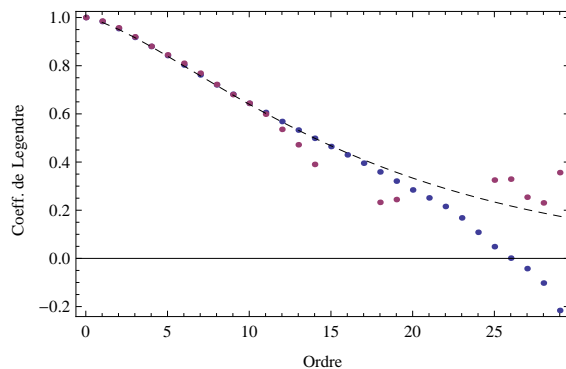


FIG. 3 – Coefficients de Legendre de la fonction de phase Von Karman, obtenus par *décomposition* du PPC pour les épaisseurs  $6.25\ell$  (bleu) et  $12.5\ell$  (rose), et fonction de phase exacte (noir).

ractéristique de la fonction de phase et celle de la distribution mesurée  $\mu(T)$ . Au delà de limite  $\ell^*/2$ , la méthode de *décomposition* du PPC n'est rapidement plus valable pour les coefficients de Legendre d'ordre élevé, du fait de l'uniformisation de la distribution de sortie. Les perspectives de ce travail sont l'inclusion de l'information de polarisation dans le modèle ainsi que le développement d'approche d'estimation paramétrique se basant sur la fonction de phase Von Karman.

## Références

- [1] S. Chandrasekhar, *Radiative TRansfer*, Dover, 1960.
- [2] L. Margerin *et al.*, *Monte Carlo simulation of multiple scattering of elastic waves*, J. Geophys. Res., 105(B4), 78737892, 2000.
- [3] J. Dieudonné, *Special functions and linear representations of Lie groups*, AMS, 1980.
- [4] P. Sheng, *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*, Springer, 2006.
- [5] X. Ning, L. Papiez, & G. Sandison, *Compound Poisson process method for the multiple scattering of charged particles*, Physical Review E, Vol. 52, No. 5, pp. 5621-5633, 1995.
- [6] L. Klimes, *Correlation functions of random media*. Pure appl. geophys., Vol. 159, pp. 1811-1831, 2002.
- [7] M. Abramowitz, & I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, 1970.
- [8] B. Buchmann, & R., Grübel, *Decomposing : An estimation problem for Poisson random sums*. The Annals of Statistics, Vol. 31, No. 4, pp. 1054-1074, 2003.
- [9] S. Said. *et al.*, *Nonparametric estimation for Compound Poisson Processes on compact Lie groups*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Taïwan, 2009.
- [10] G.S. Chirikjian & A.B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis*. CRC Press, 2000.