NONPARAMETRIC ESTIMATION FOR COMPOUND POISSON PROCESSES ON COMPACT LIE GROUPS

S. Said, N. le Bihan
Gipsa-Lab
Department of Images and Signal
CNRS / INP Grenoble

C. Lageman*
Department of Electrical Engineering and Computer Science
Université de Liège

J. H. Manton
Department of Electrical and Electronic Engineering
University of Melbourne

ABSTRACT

Motivated by applications in multiple scattering, we study the problem of decompounding on compact Lie groups. Employing tools from harmonic analysis, we give a nonparametric approach to this problem. The case of the special orthogonal group $SO(3)$ is discussed in detail.

Index Terms— Estimation on Lie groups, compound Poisson process, harmonic analysis, nonparametric estimation

1. INTRODUCTION

Compound Poisson processes (CPP) in Lie groups directly generalize their classical real-valued counterparts. These processes are of current interest in mathematics [1] and certain versions of them have found applications in multiple scattering problems, *e.g.* scattering of charged particles [2] and of surface waves around the earth [3]. In this paper, we present an estimation problem related to CPP on compact Lie groups. This problem is known as *decompounding*. It is defined in section 4, where a non parametric approach to its solution is given. The case of decompounding on the special orthogonal group $SO(3)$ is studied in section 5. It is seen that this case is of particular interest to multiple scattering problems. Numerical examples of the approach of section 4 are presented in subsection 5.4. The main mathematical tool used in the following is harmonic analysis on compact Lie groups. Its importance to stochastics [4, 5] and to more applied estimation problems [4, 6, 7] has been noted by several authors.

2. HARMONIC ANALYSIS ON LIE GROUPS

In this section we shortly recall some facts on harmonic analysis on compact Lie groups. We refer the reader to standard literature on representation theory [8]. In the remainder of this paper G will denote a compact Lie group and μ the biinvariant Haar measure on G. $L^2(G, \mathbb{C})$ will denote the space of square integrable functions with respect to μ. As a compact Lie group, G has a countable number of equivalence classes of irreducible complex representations which we will denote by $\mathrm{Irr}(G)$. Let us fix for each $\delta \in \mathrm{Irr}(G)$ an irreducible representation $U^\delta \colon G \to \mathbb{C}^{d_\delta \times d_\delta}$, and note d_δ the dimension of δ. By the Peter-Weyl theorem, any $f \in L^2(G, \mathbb{C})$ has the so-called Fourier expansion

$$f(g) = \sum_{\delta \in \mathrm{Irr}(G)} d_\delta \operatorname{tr}(A_\delta^\dagger U^\delta(g))$$

with $A_\delta = \int_G f(g) U^\delta(g) d\mu$ and $A^\dagger$ the Hermitian conjugate of $A \in \mathbb{C}^{d_\delta \times d_\delta}$. For a group-valued random variable X with a probability density function (pdf) $p \in L^2(G, \mathbb{C})$ we call the mapping $\phi_X \colon \delta \mapsto A_\delta$, which assigns the coefficient A_δ of the Fourier expansion of p to each irreducible representation $\delta \in \mathrm{Irr}(G)$, the *characteristic function* of X [4, 9]. This is a straightforward generalization of characteristic functions for real-valued random variable. Characteristic functions are a significant tool for our estimation problem since they allow us in the scalar case to transform convolutions into simple multiplications of the (matrix) Fourier coefficients. We call a function $f \in L^2(G, \mathbb{C})$ conjugate invariant if $f(ghg^{-1}) = f(h)$ for all $g, h \in G$. For conjugate invariant functions the Fourier expansion simplifies to

$$f(g) = \sum_{\delta \in \mathrm{Irr}(G)} d_\delta a_\delta^\dagger \chi^\delta(g) \tag{1}$$

where $a_\delta = \operatorname{tr}(A_\delta)/d_\delta$ and $\chi^\delta(g) = \operatorname{tr}(U^\delta(g))$ is the character function of the representation U^δ. We call a G-valued random variable *conjugate invariant* if it has a conjugate invariant $L^2(G, \mathbb{C})$ density.

3. COMPOUND POISSON PROCESSES

We will now introduce CPP on compact Lie groups. Let $N(t)$ be a Poisson process and $(X_i)_{i \geq 1}$ a sequence of i.i.d. random

*In part, this paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.The research of the second author was also supported by the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems.

variables which take values in G. We assume that $N(t)$ and $(X_i)_{i\geq 1}$ are independent and that the X_i have a $L^2(G,\mathbb{C})$ pdf. A *compound Poisson process* is defined as the random product

$$Y(t) = \prod_{i=1}^{N(t)} X_i$$

see [4, 1]. Using the properties of characteristic functions (cf. [1]) one derives the following formula for the characteristic function of $Y(t)$.

Proposition 1 *Let the Poisson process $N(t)$ have parameter λ. The characteristic function of $Y(t)$ is given by*

$$\phi_{Y(t)}(\delta) = \exp\left(\lambda t\left(\phi_{X_1}(\delta) - I_{d_\delta}\right)\right)$$

where exp *denotes matrix exponential and I_{d_δ} the $d_\delta \times d_\delta$ identity matrix. If the X_i are conjugate invariant then*

$$\phi_{Y(t)}(\delta) = \exp\left(\lambda t\left(\frac{\mathrm{tr}\left(\phi_{X_1}(\delta)\right)}{d_\delta} - 1\right)\right) I_{d_\delta}$$

4. DECOMPOUNDING

In applications we are faced with the problem to determine the distribution of the X_i from measurements of instances of the process $Y(t)$ at a fixed time T. This problem is known as *decompounding* on the group G. The case of decompounding scalar or vector-valued CPP has been considered in [10, 11]. For decompounding CPP on a compact Lie group we propose a generalization of the approach for scalar processes in [11]. As in the scalar case we base it on inversion of the formula for the characteristic function of $Y(t)$.

In general, the measurements of the process $Y(t)$ will be corrupted by noise. To take this into account we choose a multiplicative noise model. We assume that instances $\hat{Z}_1,\ldots,\hat{Z}_n$ of the process $Z(t) = MY(t)$ at the time T, where M is an independent, G-valued random variable representing the noise, are measured. The noise M is a Gaussian group-valued random variable [6], i.e.

$$\phi_M(\delta) = \exp\left(-\frac{\sigma^2\rho_\delta}{2}\right) I_{d_\delta}$$

with σ^2 the variance parameter and ρ_δ the eigenvalues of the Laplace-Beltrami operator on G, see [5]. Note that the ρ_δ are well-known for classical groups like $SO(n)$. Here, we will discuss only the case that the variance parameter σ^2 of the noise distribution is known. To simplify our calculations we will further assume that the X_i are conjugate invariant. Since M is also conjugate invariant, this implies that $Y(t)$ is conjugate invariant for all $t > 0$. A discussion of more general distributions would be beyond the scope of this paper.

We first provide an estimator for the characteristic function of $Y(T)$. To deal with the noise we use a deconvolution of empirical density of the $\hat{Z}_i$ with the density of M. We multiply the *empirical characteristic function* $1/n\sum_{i=1}^n U^\delta(\hat{Z}_i)$ of the measurements $\hat{Z}_i$ with the inverse of the characteristic function of M. This corresponds to deconvolution approaches as introduced in [6, 7] for density estimation on $SO(n)$ and spheres. We get the following estimator for the characteristic function of $Y(T)$

$$\begin{aligned}\hat{\phi}_{Y(T)}(\delta) &= \mathrm{tr}\left(\frac{1}{n}\sum_{i=1}^n \phi_M(\delta)^{-1} U^\delta(\hat{Z}_i)\right) I_{d_\delta} \\ &= \left(\frac{\exp(\frac{\sigma^2\rho_\delta}{2})}{n}\sum_{i=1}^n \mathrm{tr}\left(U^\delta(\hat{Z}_i)\right)\right) I_{d_\delta}\end{aligned}$$

The trace operator ensures that the our a priori assumption of conjugation invariance of $Y(t)$ is satisfied by the estimates.

For decompounding itself we invert the formula for the characteristic function of $Y(T)$. Taking into account the invariance properties of $Y(T)$ yields the following equation

$$\mathrm{tr}\left(\phi^n_{X_i}(\delta)\right) = \frac{1}{\lambda T}\log\left(\mathrm{tr}\left(\phi_{Y(T)}(\delta)\right)\right) + 1$$

Note that by (1) this is sufficient to determine the conjugate invariant density of X_i. Our estimate for $\mathrm{tr}(\phi_{Y(T)}(\delta))$ is not necessarily a positive real number. However, since the group-valued random variables have a $L^2(G,\mathbb{C})$ pdf, the estimate for $\mathrm{tr}(\phi_{Y(T)}(\delta))$ is with probability 1 guarantueed not to be a real, non-positive number. Hence, we can use in the formula above the principal branch of the complex logarithm. Using our estimate for the characteristic function of $Y(T)$ we get the following estimator for the characteristic function of X_i

$$\begin{aligned}&\hat{\phi}^n_{X_i}(\delta) = \\ &\quad\left(\frac{1}{\lambda T}\log\left(\frac{1}{n}\sum_{i=1}^n \mathrm{tr}\left(U^\delta(Z_i)\right)\right) + \frac{\sigma^2\rho_\delta}{2\lambda T} + 1\right) I_{d_\delta} \quad (2)\end{aligned}$$

To obtain an estimate for the density itself we use (1) to get

$$p^n_S(g) = \sum_{\delta\in S} d_\delta\,\mathrm{tr}\left(\hat{\phi}^n_{X_1}(\delta)\right)\chi^\delta(g) \quad (3)$$

where S is any finite subset of $\mathrm{Irr}(G)$. While the estimates are not real-valued functions, they converge to the pdf of the X_i. If a real-valued estimate is required, one can use the real part of p^n_S; the convergence properties of this estimate are the same as for the complex estimates.

Theorem 1 *Let S_m be an increasing sequence in $\mathrm{Irr}(G)$ such that $\bigcup_{m=1}^\infty S_m = \mathrm{Irr}(G)$. For $n, m \to \infty$ the random functions $p^n_{S_m}(g)$, (3), converge to the density of the X_i.*

Here, convergence means the convergence in probability of the $L^2(G,\mathbb{C})$ deviation between estimated and original densities. Due to constraints on the size of this paper we omit the proof. It is basically an application of the weak law of large numbers.

5. DECOMPOUNDING ON THE SPECIAL ORTHOGONAL GROUP

Our starting point is the relation between decompounding on the special orthogonal group $SO(3)$ and multiple scattering. The direction of propagation (DOP) of a wave/particle in a random medium is determined by a vector belonging to the unit sphere $S^2 \subset \mathbb{R}^3$. The effect of a scatterer encountered at a random time during the propagation is the transitive action of a random element of $SO(3)$ on the unit vector. It is classically admitted that the time between two scattering events obeys an exponential law [12]. The DOP, noted $x(t)$ after a time of propagation t, can thus be expressed as

$$x(t) = \prod_{i=1}^{N(t)} \mathbf{R}_i x_0$$

where the i.i.d. $SO(3)$-valued random variables $(\mathbf{R}_i)_{i\geq 1}$ represent the random action of identical scatterers on the DOP and x_0 is the original DOP. $N(t)$ is the random number of scatterers encountered after a time t, which can be supposed independent of the scatterers themselves. Here, we consider that the wave packet enters the medium at a fixed angle of incidence. The distribution of x_0 is then concentrated at one point, say the north pole, of S^2. As a consequence, the scattering problem reduces to the study of a CPP on $SO(3)$. Physically, decompounding –*i.e.* estimating the distribution of the $\mathbf{R}_i$– means estimating the characteristics of the scatterers, which is of great interest in the study of random media. We will use decompounding to perform this characterisation. First, we remind some results about harmonic analysis on $SO(3)$.

5.1. Harmonic analysis on $SO(3)$

$SO(3)$ is often considered as the archetype of compact Lie groups. Here we precise the formalism of section 2 in the case $G = SO(3)$. The classes of irreducible complex representations of $SO(3)$ are labelled by a natural index $\delta = 0, 1, \ldots$ where the dimension of the reprentation of order δ is equal to $2\delta + 1$. It is usual to choose the representations whose matrices $U^\delta = \{U^\delta_{ab} | -\delta \leq a, b \leq \delta\}$ are given in the ZXZ convention for the Euler angles by

$$U^\delta_{ab}(\alpha, \beta, \gamma) = e^{\mathtt{i}a\alpha} d^\delta_{ab}(\cos\beta) e^{\mathtt{i}b\gamma}$$

where (cf. [6]) the d^δ_{ab} note the Wigner d-functions with appropriate normalizations and $\alpha, \gamma \in [0, 2\pi)$ and $\beta \in [0, \pi]$ are the Euler angles. The Haar measure of $SO(3)$ is written in the Euler angles as $d\mu(\alpha, \beta, \gamma) = (16\pi^2)^{-1} \sin\beta d\alpha d\beta d\gamma$. The Fourier expansion of a function $f \in L^2(SO(3), \mathbb{C})$ has the matrix coefficient $A_\delta = \{A^{ab}_\delta\}$ where

$$A^{ab}_\delta = \int_\alpha \int_\beta \int_\gamma f(\alpha, \beta, \gamma) U^\delta_{ab}(\alpha, \beta, \gamma) d\mu(\alpha, \beta, \gamma)$$

5.2. Symmetries

Returning to the scattering situation described at the beginning of this section, assume the scatterers have a spherical shape. This amounts to the distribution of the DOP $x(t)$ having an axial symmetry. It can be inferred from this symmetry that the distribution of the "scatterers" $\mathbf{R}_i$ on $SO(3)$ is conjugate invariant. In terms of the Euler angles, this distribution will depend only on β and not on α and γ, its Fourier expansion reducing to an expansion in Legendre polynomials $d^\delta_{00}(\cos\beta) = P_\delta(\cos\beta)$. This is an instance of the general formula (1) for conjugate invariant functions. This symmetry assumption is retained for the simulations in subsection 5.4.

5.3. Compound Poisson Process on $SO(3)$

We consider a CPP $Y(t)$ taking values on $SO(3)$ defined as

$$Y(t) = \prod_{i=1}^{N(t)} X_i \tag{4}$$

with a similar notation to that explained above and where the X_i have a conjugate invariant pdf $p \in L^2(SO(3), \mathbb{C})$. Given n noisy measurements of $Y(t)$ at time T, the following estimator for p can be evaluated

$$\hat{p}(\alpha, \beta, \gamma) = \sum_{\delta=0}^{B} \sum_{ab=-\delta}^{\delta} (2\delta + 1) \hat{A}^{ab}_\delta U^\delta_{ab}(\alpha, \beta, \gamma)$$

where B is an appropriately chosen cut-off and for all $0 \leq \delta \leq B$ the matrix $\{\hat{A}^{ab}_\delta\}$ is the Hermitian conjugate of the (diagonal) matrix obtained following the prescription of equation (2).

5.4. Simulations

We have simulated CPP on $SO(3)$ with conjugate invariant distributions for the X_i. We have considered values of λT from units to a few tens. We avoided higher values in order for our simulations to correspond to a multiple scattering regime and note a diffusive one, in the physical interpretation proposed above. Three types of distributions for the X_i have been studied, Gaussian, Von Mises and Exponential. The Legendre coefficients of the original (*i.e.* of the X_i) distributions and those estimated using the approach of section 4 are presented in figure (1). The reconstruction error between the estimated and original Gaussian pdf is presented in figure (2) as a function of the order of Legendre coefficients. This error is the mean square error between reconstructed and original empirical Fourier (here Legendre) coefficients for different orders δ. Figure (2) concerns only the Gaussian case, where the error is presented for different values of n, the size of the data set. From figure (1), it is possible to see that the Von Mises and Gaussian cases are well recovered using our decompounding approach, while in the case of exponential distribution, results are poorer. This is due to the fact that the exponential distribution is more dispersed on $SO(3)$ than the other two distributions. In figure (2), one can see that the convergence of the Fourier coefficients is faster for low orders (small δ). This is due the lower order functions U_{ab}^{δ} having lower frequencies.

6. CONCLUSION

The problem of decompounding on compact Lie groups was presented. Under some simplifying assumptions, a non parametric approach to this problem based on harmonic analysis was given. The importance of the case of the special orthogonal group $SO(3)$ to multiple scattering problems was discussed. Numerical simulations implemented for the case of $SO(3)$ seem to validate the general approach proposed.

7. REFERENCES

[1] D. Applebaum, "Compound Poisson processes and Lévy processes in groups and symmetric spaces," *Journal of Theoretical Probability*, vol. 13, no. 2, pp. 383–425, 2000.

[2] X. Ning, L. Papiez, and G. Sandison, "Compound-Poisson-process method for the multiple scattering of charged particles," *Physical Review E*, vol. 52, no. 5, pp. 5621–5633, 1995.

[3] H. Sato and M. Nishino, "Multiple isotropic-scattering model on the spherical earth for the synthesis of rayleigh-wave envelopes," *J. Geophys. Res.*, vol. 107, no. B02, pp. 2343, 2002.

[4] Ulf Grenander, *Probabilities on Algebraic Structures*, John Wiley & Sons Inc., 1963.

[5] M. Liao, *Lévy processes on Lie groups*, Cambridge University Press, Cambridge, 2004.

[6] Ja-Yong Koo and P.T. Kim, "Asymptotic minimax bounds for stochastic deconvolution over groups," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 289–298, Jan. 2008.

[7] P. T. Kim and D. St. P. Richards, "Deconvolution density estimation on compact lie groups," in *Algebraic Methods in Statistics and Probability*. 2001, pp. 155–171, AMS.

[8] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Springer, New York, 1985.

[9] P. Kim, "Deconvolution density estimation on so(n)," *Annals of Statistics*, vol. 26, no. 3, pp. 1083–1102, 1998.

[10] B. Buchmann and R. Grübel, "Decompounding: An estimation problem for Poisson random sums," *The Annals of Statistics*, vol. 31, no. 4, pp. 1054–1074, 2003.

[11] B. van Es, S. Gugushvili, and P. Spreij, "A kernel type nonparametric density estimator for decompounding," *Bernoulli*, vol. 13, no. 3, pp. 672–694, 2007.

[12] X. Ning, L. Papiez, and G. Sandinson, "Compound poisson process method for the multiple scattering of charged particles," *Physical Review E*, vol. 52, no. 5, pp. 5621–5633, 1995.

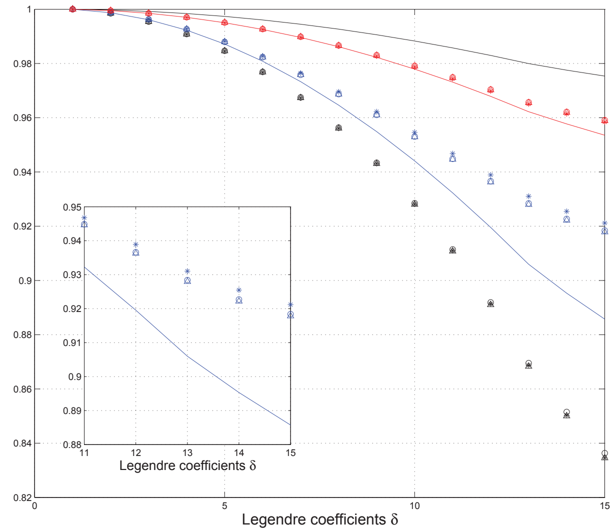

Fig. 1. Original (solid line) and estimated ($*$,$\circ$ and $\triangle$) Legendre coefficients for Gaussian (Blue), Von Mises (Red) and Exponential (black) for differents sizes of data sample ($* = 10^3$; $\circ = 5.10^4$; $\triangle = 10^6$).

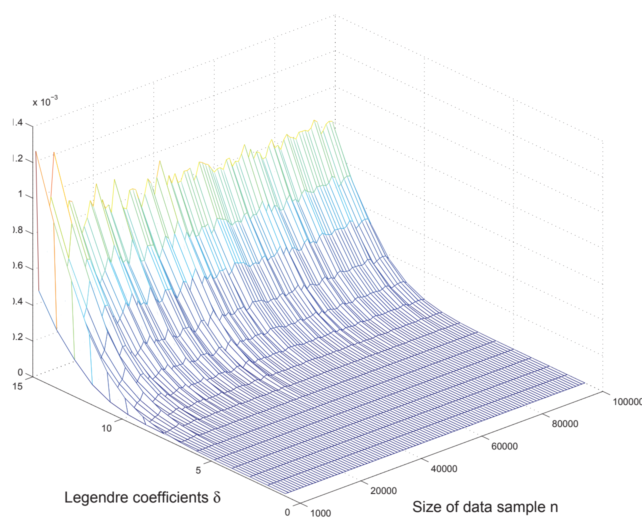

Fig. 2. Reconstruction error (MSE) in the Gaussian case for different data sample sizes as a function of the Legendre coefficient order δ.