



**HAL**  
open science

# Estimation of coherent scattering in dynamic random media

Clément J. Roussel, Arnaud Coatanhay, Alexandre Baussard

► **To cite this version:**

Clément J. Roussel, Arnaud Coatanhay, Alexandre Baussard. Estimation of coherent scattering in dynamic random media. *Waves in Random and Complex Media*, 2021, 31 (6), pp.2301-2326. 10.1080/17455030.2020.1740824 . hal-02520048

**HAL Id: hal-02520048**

**<https://hal.science/hal-02520048>**

Submitted on 18 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# *Estimation of coherent scattering in dynamic random media*

C. J. Roussel<sup>a\*</sup>, A. Coatanhay<sup>a</sup>, A. Baussard<sup>a</sup>

<sup>a</sup>*ENSTA Bretagne, Lab-STICC (UMR CNRS 6285), 2 rue François Verny, 29806 Brest  
Cedex 9, France*

*(submitted October 2018)*

We address the question of estimating coherent scattering (weak scattering or coherent scatterer) by a time-evolving random medium. We consider two models of coherent scattering: homodyned K (HK) and generalized K (GK), and in both cases we derive stochastic differential equations for the scattered field. Approximate transition probabilities are computed using Euler-Maruyama scheme and the parameters for coherent scattering are estimated by maximum likelihood (ML). Using numerical simulations, we show that in the HK case, maximum likelihood estimation does not provide a significant advantage over a simplistic estimator based on the ergodicity property of the scattered field. On contrary, in GK scattering the transition probabilities carry significant information about the parameters, resulting in much better performance of the ML estimator.

**Keywords:** dynamic random media, electromagnetic scattering, weak scattering, target detection, transition probabilities

## 1. Introduction

The random walk model (see [1]) represents the field scattered by a random medium as the sum of contributions from scattering centers, which may be discrete scatterers (particles in suspension for example) or focusing elements from continuous media (rough surfaces, turbulent fluids). The random walk is very general and therefore applicable to a wide variety of scattering configurations. The scattered complex amplitude can be assumed to be the sum of the complex amplitudes of the scattering centers, and a distribution is usually assigned to the phase of the scatterers. In case of *strong scattering*, the phase is spread uniformly over  $[0, 2\pi[$ . If the number  $N$  of scattering centers is increased to infinity (after proper normalization of the amplitudes), the resulting scattered field has exponentially-distributed intensity (modulus squared). Also for strong scattering, if the number  $N$  of scatterers fluctuates according to a negative binomial distribution of mean  $\bar{N}$ , the scattered field becomes  $K$ -distributed [2], [3] in intensity after letting  $\bar{N}$  go to infinity.

As a particular example, scattering of radar waves by the sea surface has been extensively studied (see [4]) due to its practical implications for maritime surveillance. It has been shown empirically that for not too low grazing angles, the  $K$  distribution adequately represents the statistics of the scattered field [4].

The  $K$  distribution can be generalized by at least two slightly different models:

---

\*Corresponding author. Email: clement.roussel@ensta-bretagne.org

the homodyned  $K$  model (HK) and the generalized  $K$  model (GK) ([5] chapter 9). The HK model corresponds to the addition of a constant to a  $K$ -distributed complex amplitude, while the GK model corresponds to the addition of a bias common to all scatterers in the random walk model (with negative binomial population of scatterers). Alternatively, the GK model is obtained for a random walk with non uniform (over  $[0, 2\pi[$ ) phase distribution (see [6], [7]). In both HK and GK scattering, and contrary to the random walk with uniform scatterer's phases (*e.g.* the  $K$  distribution), the phase of the total scattered field is not uniformly distributed over  $[0, 2\pi[$ . We refer to this situation as *coherent scattering*.

At least two interpretations exist for coherent scattering: weak scattering and presence of a coherent scatterer (designated target in the following). *Weak scattering* refers to the situation where the scatterers in the random walk do not spread in phase uniformly over  $[0, 2\pi[$ . Though GK seems to be the model of choice for weak scattering, HK has also been proposed for it ([5] chapter 9, [7]). The other interpretation is that the field scattered by the random medium can be described by an isotropic random walk ( $K$  distribution), but that there is in addition a *target* embedded in the random medium and which contributes to the total field. HK seems to be best adapted to describe this situation, but both HK and GK have been proposed to model it [4]. By convention, we refer to a target in the remaining of this paper. However, it should be remembered that the HK and GK models are general to coherent scattering, such that our results are equally valid for weak scattering. The isotropic contribution of the random medium is termed *clutter*.

A major limitation of the HK and GK models presented in the literature before [5], is their static nature: the relation between the scattered field at two subsequent times is not explicit, and usually independence is assumed. If the time interval is long, this assumption is essentially valid. However, it breaks down when the time interval is short, which occurs if the sampling frequency sufficiently high. In order to overcome this problem, Field's model [5] represents the clutter in a fully dynamic way as a stochastic process solution to a stochastic differential equation. Stochastic differential equations (SDE) are the stochastic equivalent of ordinary differential equations. Field's model is a dynamic generalization of the isotropic ( $K$  distribution) random walk model. It leads to the concept of SDE, which is powerful for describing the transitions of the clutter, however small the time interval is. Indeed, when a SDE describes the dynamics of a process (or signal), one can compute the transition probabilities of the process, *i.e.* the probability that it takes some value in the future given an observation made at the present time. In chapter 9 of [5], Field expresses the multidimensional SDE for HK and GK scattering (interpreted as being for weak scattering) in polar coordinates.

In this paper, we propose to reexpress the SDE for HK and GK scattering in cartesian coordinates to estimate the transition probabilities of the target plus clutter signal (total scattered field), and to use them for maximum likelihood estimation of the target parameters.

In section 2, we explicit the isotropic random walk model, give a brief account of the concepts of SDE and transition probabilities, and explain Field's model in connection with the random walk model. In section 3, we use Field's model and Ito's calculus to derive the SDE of the clutter plus target signal, for the HK and GK "target models". In section 4, we derive analytical approximation for the transition probabilities in HK and GK scattering. We then show that maximum likelihood estimation of the target parameters is possible using these transition probabilities and leads to explicit formula. In section 5, we assess the performance

of the maximum likelihood estimator using numerical simulations of the SDE we derived in section 3. Section 6 is a discussion about the limits and advantages of our approach, as well as its implications in terms of decision theory. Finally, we conclude in section 7.

## 2. Theoretical background

### 2.1. The random walk model

The reflectivity of a dynamic random medium is best represented mathematically by a stochastic process, *i.e.* a family  $\{X_t, t \geq 0\}$  of random variables. The most well-known statistical models [3], [7], [8], [1], [4] for the clutter (no coherent scatterer) are derived by starting from the fundamental idea that the reflectivity is a sum of contributions over a population of independent scatterers, *i.e.*  $\forall t \geq 0$ :

$$X_t = \sum_{n=1}^{N_t} a_t^{(n)} e^{i\phi_t^{(n)}}. \quad (1)$$

$a_t^{(n)}$  and  $\phi_t^{(n)}$  are respectively the amplitude and phase of the  $n$ -th scatterer. It is assumed that for fixed  $t$ , the amplitudes  $a_t^{(n)}$  are independent and identically distributed (i.i.d). Similarly, the phases  $\phi_t^{(n)}$  are independent and uniformly distributed over  $[0, 2\pi[$ , and all phases and amplitudes are independent. Finally, the number of scatterers  $N_t$  is itself a random variable which follows a negative binomial distribution. Moreover, all processes are stationary.

Let  $\bar{N} = \mathbb{E}[N_t]$  be the average number of scatterers. If we normalize the amplitudes by  $\bar{N}$  and let  $\bar{N} \rightarrow +\infty$ , we obtain:

$$Y_t = \lim_{\bar{N} \rightarrow +\infty} \sum_{n=1}^{N_t} \frac{a_t^{(n)}}{\bar{N}} e^{i\phi_t^{(n)}}. \quad (2)$$

$\forall t$ ,  $Y_t$  is then a complex random variable whose properties are known. For example,  $|Y_t|^2$  (the intensity) follows the famous K distribution. However, something crucial is missing in this formulation: the relation between  $Y_{t_1}$  and  $Y_{t_2}$  for two subsequent times  $t_1, t_2$  is not specified. Only the static distributions (for a fixed  $t$ ) are known. One can partially dodge the problem by specifying correlation times and designing *ad hoc* numerical procedures for simulating individual trajectories of the process (see [4]). However, there is a much more satisfying framework: that of stochastic differential equations.

### 2.2. Stochastic differential equations

A stochastic differential equation (abbreviated SDE), is the proper generalization of a differential equation to stochastic processes. The solution, instead of being simply a function of time, is a stochastic process, *i.e.* at time  $t$  one has a random variable or vector, instead of a single value. A homogeneous  $n$  dimensional ( $n$ -D) SDE is in the following form:

$$\begin{cases} dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t \\ Y_0 = \xi_0 \end{cases} \quad (3)$$

$W_t = [W_t^{(1)}, W_t^{(2)}, \dots, W_t^{(n)}]^T$  is a  $n$ -D brownian motion, which means that the  $W_t^{(k)}$  are independent 1-D brownian motions.  $\xi_0$  is the initial condition (random vector).  $\mu$  is called the ‘drift’ and  $\sigma$  is called the ‘volatility’ (matrix valued). Under the conditions of Ito’s theorem of existence and unicity of the solutions, the unique solution, denoted  $(Y_t)_{t \geq 0}$  is a homogeneous Markov process (e.g. [9]). Observe that if one sets the volatility  $\sigma = 0$  and  $\xi_0 \in \mathbb{R}$  (deterministic initial condition), one gets back to an ordinary differential equation.

Let us denote  $y \mapsto p(Y_t = y | Y_0 = \xi_0)$  the distribution of  $Y_t$  with initial condition  $Y_0 = \xi_0$ . In the 1-D case,  $p(Y_t = y | \xi_0)$  is a function of  $(t, y) \in \mathbb{R}^2$  and it is solution to the so-called Fokker-Planck (a.k.a Kolmogorov forward) equation:

$$\frac{\partial p(Y_t = y | Y_0 = \xi_0)}{\partial t} = \frac{1}{2} \frac{\partial^2 \sigma(y)^2 p(Y_t = y | Y_0 = \xi_0)}{\partial y^2} - \frac{\partial \mu(y) p(Y_t = y | Y_0 = \xi_0)}{\partial y}. \quad (4)$$

The Fokker-Planck equation (FPE) is a partial derivative equation in the variables  $(y, t)$ . Its solution is the time-dependent distributions of  $Y_t$  and it depends on the initial condition. There are 2 special cases. First, if  $Y_0$  is distributed according to the stationary solution of the FPE, then so is  $Y_t$  for all  $t$ . In that case, the distribution is simply denoted  $p(Y_t = y)$ . This is a priori the case for the reflectivity of the random medium since it is stationary. Second, if one sets  $Y_0 = x$  where  $x \in \mathbb{R}$ , then the distributions given by the solution of the Fokker-Planck equation are the transition probabilities  $p(Y_t = y | Y_0 = x)$ . Note that since the process is homogeneous,  $\forall h \geq 0$ :

$$p(Y_{t+h} = y | Y_t = x) = p(Y_h | Y_0). \quad (5)$$

Following the explanations in [10], the transition probabilities can be used to ‘transport’ non-synchronous observations of the random medium to a common time. The transition probabilities were obtained by the analytical resolution of 1-D Fokker-Planck equations. Most importantly, it should be understood that  $p(Y_{t+h} = y | Y_t = x)$  is the distribution of  $Y_{t+h}$  knowing that you have observed the value  $x$  at time  $t$ . It is different from the stationary distribution and enables you to make more precise probabilistic inferences for what the value of the reflectivity should be a bit later given some observed value at the present time.

Another application of the transition probabilities is parameter estimation. First we assume that  $Y_t$  depends on some unknown parameter  $\lambda$ . Assume also that a particular trajectory of  $Y_t$  is observed at discrete times  $t_0 < t_1 < \dots < t_n$  and yields  $\tilde{Y}_0, \tilde{Y}_1, \dots, Y_{t_n}$ .

Since  $Y_t$  is a Markov process, its joint probability density function associated is:

$$\begin{aligned} & p_\lambda(Y_{t_0} = \tilde{Y}_0, \dots, Y_{t_n} = \tilde{Y}_n) \\ &= p_\lambda(Y_{t_0} = \tilde{Y}_0, \dots, Y_{t_{n-1}} = \tilde{Y}_{n-1}) p_\lambda(Y_{t_n} = \tilde{Y}_n | Y_{t_0} = \tilde{Y}_0, \dots, Y_{t_{n-1}} = \tilde{Y}_{n-1}) \\ &= p_\lambda(Y_{t_0} = \tilde{Y}_0, \dots, Y_{t_{n-1}} = \tilde{Y}_{n-1}) p_\lambda(Y_{t_n} = \tilde{Y}_n | Y_{t_{n-1}} = \tilde{Y}_{n-1}). \end{aligned}$$

By recurrence, the joint probability density function, or more simply the likelihood,

is:

$$\mathcal{L}(\lambda) = p_\lambda(Y_{t_0} = \tilde{Y}_0) \prod_{i=1}^n p_\lambda(Y_{t_i} = \tilde{Y}_i | Y_{t_{i-1}} = \tilde{Y}_{i-1}). \quad (6)$$

Improperly speaking, it is the ‘probability’ that we observe the sequence  $\tilde{Y}_0, \tilde{Y}_1, \dots, \tilde{Y}_n$  if the parameter is  $\lambda$ . Maximum likelihood estimation (ML) consists in maximizing  $\mathcal{L}$  with respect to  $\lambda$ , which yields the estimated parameter  $\hat{\lambda}$ . In practice, we often maximize the log of  $\mathcal{L}$  rather than  $\mathcal{L}$  directly. As evidenced in equation (6), all we need to estimate  $\lambda$  by ML is the stationary distribution  $p(Y_t = y)$  and the transition probabilities  $p(Y_{t+h} = y | Y_t = x)$ . Estimating parameters using this procedure (SDE  $\rightarrow$  FPE  $\rightarrow$  transition probabilities  $\rightarrow$  ML estimation) has been done in [11] for estimating the parameters of the clutter only.

The reader may have noticed that we provided the FPE only for 1-D processes. This is because we merely wanted to illustrate the general procedure to get transition probabilities. Most of the time, the Fokker-Planck equation cannot be solved analytically, especially for dimensions greater than 1. In the remainder of this paper, the procedure of going from a SDE to transition probabilities and ML estimation is done but the transition probabilities are approximated and not obtained through the FPE. As exemplified in section 4, the approximation for small time steps comes from Euler-Maruyama numerical scheme (see [12] and [13]). This scheme is the most simple to numerically solve a SDE. Let  $[0, T]$  be a finite time interval and  $t_0 = 0 < t_1 < \dots < t_n$  a partition of  $[0, T]$ . The Euler-Maruyama method applied to equation (3) reads:

$$Y_{t_i} = Y_{t_{i-1}} + \mu(Y_{t_{i-1}})(t_i - t_{i-1}) + \sigma(Y_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}). \quad (7)$$

Equation (7) states that the increment of  $Y$  between  $t_{i-1}$  and  $t_i$  is the sum of a term proportional to  $\Delta t_i = t_i - t_{i-1}$  and a term proportional to the increment of the multidimensional brownian motion  $\Delta W_{t_i} = W_{t_i} - W_{t_{i-1}}$ . This increment is a Gaussian random vector with law  $\mathcal{N}(0, \Delta t_i \sigma \sigma^T(Y_{t_{i-1}}))$ . Euler-Maruyama scheme is useful for at least 2 different reasons: for small  $\Delta t$ , it provides a gaussian approximation to the transition probabilities, and otherwise it can simply be used to simulate many trajectories (realizations) of the process which is useful to assess parameter estimation methods as exemplified in [11] and in section 5.

### 2.3. Field’s model of the clutter

In the most general situation, the reflectivity  $C\Psi(t)$  of the observed scene is defined as the ratio of the reflected wave to the incoming wave. We have:

$$E_s(t) = C\Psi(t)E_i(t), \quad (8)$$

where  $E_s(t)$  and  $E_i(t)$  are respectively the scattered and incident waves (one component of the electric or magnetic wave).  $C\Psi(t)$  is the complex reflectivity.  $C$  is a positive constant equal to the square root of the mean power.  $\Psi(t)$ , denoted as  $\Psi_t$  in stochastic process theory, is the normalized complex reflectivity (simply referred to as reflectivity). Actually,  $\Psi_t$  modulates both the amplitude and phase of the wave. Field, in [5] (chapters 6 to 8), built upon the random walk model for the clutter (no coherent scatterer) in equation (1) which, just to remind, led to the K distribution for the intensity. We noted that this model was static: no dynamics

are specified and so the relation between the process at two different times is undefined. With the same notations as in section 2.1, Field adds mainly 3 hypotheses. First, he assumes that the dynamics of the phases are determined by the SDE:

$$\begin{cases} d\phi_t^{(n)} = \mathcal{B}^{1/2} dW_t^{(n)} \\ \phi_0^{(n)} = \Delta^{(n)}, \end{cases} \quad (9)$$

where  $\forall n$ ,  $\Delta^{(n)}$  is uniformly distributed over  $[0, 2\pi[$  and  $\mathcal{B}$  is a positive constant. Second, he assumes that the amplitudes are deterministic constants, *i.e.*  $\exists a > 0$  such that

$$\forall t, n : a_t^{(n)} = a. \quad (10)$$

Finally, he assumes that the number of scatterers  $N_t$  is a linear Birth-Death-Immigration population model, which is consistent with the fact that  $\forall t$ ,  $N_t$  is negative binomial distributed (see for example [14] for an introduction to population models). Under such hypotheses, he shows that the normalized reflectivity  $\Psi_t$  of the random medium only (*i.e.* the clutter) can be expressed as the product:

$$\Psi_t = x_t^{1/2} \gamma_t = x_t^{1/2} \left( \gamma_t^{(R)} + i\gamma_t^{(I)} \right), \quad (11)$$

where  $x_t$ ,  $\gamma_t^{(R)}$  and  $\gamma_t^{(I)}$  are stochastic processes solutions to the following stochastic differential equations ([5], chapter 8):

$$\begin{cases} dx_t = \mathcal{A}(1 - x_t)dt + \left(2\frac{\mathcal{A}}{\alpha}x_t\right)^{\frac{1}{2}} dW_t^{(x)} \\ d\gamma_t^{(R)} = -\frac{1}{2}\mathcal{B}\gamma_t^{(R)}dt + \frac{1}{\sqrt{2}}\mathcal{B}^{\frac{1}{2}}dW_t^{(R)} \\ d\gamma_t^{(I)} = -\frac{1}{2}\mathcal{B}\gamma_t^{(I)}dt + \frac{1}{\sqrt{2}}\mathcal{B}^{\frac{1}{2}}dW_t^{(I)}. \end{cases} \quad (12)$$

$W_t^{(x)}$ ,  $W_t^{(R)}$ ,  $W_t^{(I)}$  are 3 independent brownian motions. It should be noted that  $\gamma_t^{(R)}$  and  $\gamma_t^{(I)}$  are assumed to be independent.

Field's model of the clutter is valid for any wavelength: optical, infrared, microwaves etc. However, for the hypothesis of uniform phase over  $[0, 2\pi[$  to be true, the size of the random medium should be large compared to the wavelength of the incoming wave.

The first equation of (12) is equation (8.9) of [5] (in [5]  $\alpha x_t$  has been replaced by  $x_t$ ) while the last two equations of (12) constitute the complex-valued equation (8.4) of [5]. From (11), one should understand that if we know  $x_t, \gamma_t^{(R)}, \gamma_t^{(I)}$ , we know the reflectivity of the time-evolving random medium  $\Psi_t$ .  $x_t$  is the scattering cross-section normalized by its mean value, and will be referred to as SCS in the following.  $\gamma_t$  is the complex-valued speckle (with unit mean amplitude), expressed with its real and imaginary parts  $\gamma_t^{(R)}$  and  $\gamma_t^{(I)}$ .

Three constants parametrize the model:  $\mathcal{A}$  and  $\alpha$  for the SCS,  $\mathcal{B}$  for the speckle.  $\mathcal{A}$  and  $\alpha$  are from the underlying Birth-Death-Immigration population model for the number of scatterers and  $\mathcal{B}$  comes from the dynamics of the phases in equation (9).  $\mathcal{A}$  and  $\mathcal{B}$  are homogeneous to the inverse of a time (*i.e.* a frequency).  $\mathcal{A}$  can be understood as the inverse of a decorrelation time for the SCS, and  $\mathcal{B}$  as the inverse of a decorrelation time for the speckle.  $\alpha$  is the inverse of the variance of the SCS

$x_t$  for any time  $t$  as seen below. In Field's model,  $\mathcal{A} \ll \mathcal{B}$ , *i.e.*  $\gamma_t$  decorrelates much faster than  $x_t$ . The SCS should therefore be thought of as a slow process which modulates the power of the fast complex process  $\gamma_t$ . Also note that the carrier wave does not appear in equation (8). This is so because it is assumed that the carrier wave oscillates on a much shorter timescale than the reflectivity  $\Psi_t$ . To sum up, there are three timescales  $T$  at stake:

$$T_{carrier} \ll T_{speckle} \ll T_{SCS}. \quad (13)$$

Only  $T_{speckle}$  and  $T_{SCS}$  are relevant here, and are dictated by  $\mathcal{B}$  and  $\mathcal{A}$  respectively. It is shown in [11] that:

$$\begin{cases} p\left(\gamma_t^{(R)} = x\right) = p\left(\gamma_t^{(I)} = x\right) = \frac{1}{\sqrt{\pi}}e^{-x^2} \\ p\left(\gamma_t^{(R)2} + \gamma_t^{(I)2} = x\right) = e^{-x} \\ p(C^2x_t = x) = \frac{\left(\frac{\alpha}{C^2}\right)^\alpha x^{\alpha-1} e^{-\frac{\alpha}{C^2}x}}{\Gamma(\alpha)} \end{cases} \quad (14)$$

where  $\Gamma$  is the gamma function. From equation (14), the mathematical expectation of the squared amplitude of the reflectivity is:

$$\begin{aligned} \mathbb{E}[|C\Psi_t|^2] &= \mathbb{E}\left[C^2x_t\left(\gamma_t^{(R)2} + \gamma_t^{(I)2}\right)\right] \\ &= C^2\mathbb{E}[x_t]\mathbb{E}\left[\gamma_t^{(R)2} + \gamma_t^{(I)2}\right] = C^2 \times 1 \times 1 = C^2, \end{aligned} \quad (15)$$

where we have used independence of  $x_t$  and  $\gamma_t^{(R)2} + \gamma_t^{(I)2}$ . Therefore, the normalized reflectivity  $\Psi_t$  is normalized in the sense that  $\mathbb{E}[|\Psi_t|^2] = 1$ . We also say that the mean power is 1.

#### 2.4. Numerical application

In section 5, we need to use numerical values for the parameters  $\mathcal{A}, \mathcal{B}, C$  and  $\alpha$  to generate trajectories of the reflectivity and assess the performance of the estimators. Rather than setting the values randomly, we set  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz and  $\alpha = 1$  as in [11]. These particular values are relevant to radar scattering by the sea surface. An important constraint that they respect is:

$$\mathcal{A} \ll \mathcal{B}, \quad (16)$$

which corresponds to the difference of timescales between the SCS and the speckle. We also set  $C = 1$  in the remaining. The idea behind this assumption is that the performances of the estimators depend on the ratio of the target power to the clutter power: it is not necessary to tune both the target and clutter mean powers.

### 3. SDE of the clutter plus target

Section 2.3 summed up Field's approach to derive the SDE of the clutter (random medium only). In [5], chapter 9, the SDE of the clutter plus target are expressed in polar coordinates, which yields heavy expressions. In this section, we derive the SDE of the clutter plus target (coherent scatterer) in cartesian coordinates (real



and imaginary parts) since it is a more convenient coordinate system for target parameter estimation. Two models for the target reflectivity will be considered: Homodyned K scattering and Generalized K scattering (see [5] p 70-71 and see sections 3.2 and 3.3). It will be evidenced in equations (22) and (26) that HK and GK scattering encompass the cases of Rice, Rayleigh and K distribution scattering. Rice scattering is recovered when there is a target and the SCS  $x_t$  is constant, Rayleigh scattering is recovered when there is no target and the SCS is constant, and K distribution scattering is recovered when there is no target but still a varying SCS.

### 3.1. SDE of the real and imaginary parts of the clutter

As a preliminary step, we derive the SDE of the in-phase (real part) component, denoted  $R_t^{(cl)}$ , and quadrature phase (imaginary part) component, denoted  $I_t^{(cl)}$ , of the clutter only. Of course,  $R_t^{(cl)} = x_t^{1/2} \gamma_t^{(R)}$  and  $I_t^{(cl)} = x_t^{1/2} \gamma_t^{(I)}$ . Since from equation (12)  $\gamma_t^{(R)}$  and  $\gamma_t^{(I)}$  follow the same SDE with different brownian motions, we just need to derive the SDE for  $R_t^{(cl)}$  for example, and that of  $I_t^{(cl)}$  will follow immediately.

Let  $s_t = x_t^{1/2}$ . Using Ito's formula for  $f(x_t)$  with  $f(x) = x^{1/2}$ , one can show that  $s_t$  is solution to the following SDE:

$$ds_t = \frac{\mathcal{A}}{2} \left( \frac{1}{s_t} \left( 1 - \frac{1}{2\alpha} \right) - s_t \right) dt + \left( \frac{\mathcal{A}}{2\alpha} \right)^{1/2} dW_t^{(x)}. \quad (17)$$

From Ito's product law:

$$dR_t^{(cl)} = s_t d\gamma_t^{(R)} + \gamma_t^{(R)} ds_t + d[s\gamma^{(R)}]_t \quad (18)$$

$$\Leftrightarrow dR_t^{(cl)} = s_t d\gamma_t^{(R)} + \gamma_t^{(R)} ds_t, \quad (19)$$

where we have used that  $[s\gamma^{(R)}]_t$ , the quadratic variation at time  $t$  of the processes  $s_t$  and  $\gamma_t^{(R)}$ , is zero by independence of  $x_t$  and  $\gamma_t^{(R)}$ . Using equations (12), (17) and (18), we obtain after factorization:

$$\begin{aligned} dR_t^{(cl)} = & \left( -\frac{\mathcal{A} + \mathcal{B}}{2} s_t \gamma_t^{(R)} + \frac{\mathcal{A} \gamma_t^{(R)}}{2s_t} \left( 1 - \frac{1}{2\alpha} \right) \right) dt \\ & + \left[ \gamma_t^{(R)} \left( \frac{\mathcal{A}}{2\alpha} \right)^{1/2} \quad s_t \left( \frac{\mathcal{B}}{2} \right)^{1/2} \right] \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \end{bmatrix}. \end{aligned} \quad (20)$$

The system of SDE for the clutter in the coordinate system  $(x, R^{(cl)}, I^{(cl)})$  becomes:

$$\begin{aligned}
\begin{bmatrix} dx_t \\ dR_t^{(cl)} \\ dI_t^{(cl)} \end{bmatrix} &= \begin{bmatrix} \mathcal{A}(1-x_t) \\ -\frac{\mathcal{A}+\mathcal{B}}{2}R_t^{(cl)} + \frac{\mathcal{A}R_t^{(cl)}}{2x_t} \left(1 - \frac{1}{2\alpha}\right) \\ -\frac{\mathcal{A}+\mathcal{B}}{2}I_t^{(cl)} + \frac{\mathcal{A}I_t^{(cl)}}{2x_t} \left(1 - \frac{1}{2\alpha}\right) \end{bmatrix} dt \\
&+ \begin{bmatrix} \left(\frac{2\mathcal{A}x_t}{\alpha}\right)^{1/2} & 0 & 0 \\ \frac{R_t^{(cl)}}{x_t^{1/2}} \left(\frac{\mathcal{A}}{2\alpha}\right)^{1/2} & \left(\frac{\mathcal{B}x_t}{2}\right)^{1/2} & 0 \\ \frac{I_t^{(cl)}}{x_t^{1/2}} \left(\frac{\mathcal{A}}{2\alpha}\right)^{1/2} & 0 & \left(\frac{\mathcal{B}x_t}{2}\right)^{1/2} \end{bmatrix} \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \\ dW_t^{(I)} \end{bmatrix}. \tag{21}
\end{aligned}$$

### 3.2. Homodyned $K$ scattering

In Homodyned  $K$  (HK) scattering, the reflectivity of the target is a complex constant added to the reflectivity of the clutter. We denote  $\Psi_t^{(HK)}$  the total reflectivity and we have:

$$\Psi_t^{(HK)} = \Psi_c^{(R)} + i\Psi_c^{(I)} + x_t^{1/2}\gamma_t, \tag{22}$$

where  $\Psi_c = \Psi_c^{(R)} + i\Psi_c^{(I)}$  is the target constant. It is constant in both phase and amplitude. For simplicity, we denote  $R_t, I_t$  the real and imaginary parts of  $\Psi_t^{(HK)}$ . We omit voluntarily the superscript  $(HK)$  to lighten the notation.

It is very straightforward to obtain the SDE of  $R_t$  and  $I_t$  in the case of HK scattering. Indeed, from (22) we get:

$$\begin{cases} R_t = \Psi_c^{(R)} + R_t^{(cl)} ; dR_t = dR_t^{(cl)} \\ I_t = \Psi_c^{(I)} + I_t^{(cl)} ; dI_t = dI_t^{(cl)}. \end{cases} \tag{23}$$

Therefore, we have directly from (21):

$$\begin{aligned}
\begin{bmatrix} dx_t \\ dR_t \\ dI_t \end{bmatrix} &= \begin{bmatrix} \mathcal{A}(1-x_t) \\ -\frac{\mathcal{A}+\mathcal{B}}{2} \left(R_t - \Psi_c^{(R)}\right) + \frac{\mathcal{A}(R_t - \Psi_c^{(R)})}{2x_t} \left(1 - \frac{1}{2\alpha}\right) \\ -\frac{\mathcal{A}+\mathcal{B}}{2} \left(I_t - \Psi_c^{(I)}\right) + \frac{\mathcal{A}(I_t - \Psi_c^{(I)})}{2x_t} \left(1 - \frac{1}{2\alpha}\right) \end{bmatrix} dt \\
&+ \begin{bmatrix} \left(\frac{2\mathcal{A}x_t}{\alpha}\right)^{1/2} & 0 & 0 \\ \frac{(R_t - \Psi_c^{(R)})}{x_t^{1/2}} \left(\frac{\mathcal{A}}{2\alpha}\right)^{1/2} & \left(\frac{\mathcal{B}x_t}{2}\right)^{1/2} & 0 \\ \frac{(I_t - \Psi_c^{(I)})}{x_t^{1/2}} \left(\frac{\mathcal{A}}{2\alpha}\right)^{1/2} & 0 & \left(\frac{\mathcal{B}x_t}{2}\right)^{1/2} \end{bmatrix} \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \\ dW_t^{(I)} \end{bmatrix}. \tag{24}
\end{aligned}$$

The SDE obtained for HK scattering, equation (24), can also be written in the more compact form:

$$\begin{bmatrix} dx_t \\ dR_t \\ dI_t \end{bmatrix} = \beta_{\Psi_c}^{(HK)}(x_t, R_t, I_t) dt + \Sigma_{\Psi_c}^{(HK)}(x_t, R_t, I_t) \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \\ dW_t^{(I)} \end{bmatrix}. \tag{25}$$

Vector  $\beta_{\Psi_c}^{(HK)}$  and matrix  $\Sigma_{\Psi_c}^{(HK)}$  are respectively the drift and volatility in the case of HK scattering.

### 3.3. Generalized K scattering

In Generalized K (GK) scattering, the reflectivity of the target is modulated by the SCS  $x_t$  (see [5] p 71):

$$\Psi_t^{(GK)} = \left( \Psi_c^{(R)} + i\Psi_c^{(I)} \right) \eta x_t + x_t^{1/2} \gamma_t, \quad (26)$$

where  $\eta$  is a constant coupling factor and  $\Psi_c = \Psi_c^{(R)} + i\Psi_c^{(I)}$  is the target constant. Physically, it means that the reflectivity of the target,  $\left( \Psi_c^{(R)} + i\Psi_c^{(I)} \right) \eta x_t$ , varies proportionally as the number of scatterers which contribute to the clutter. Similarly to section 3.2,  $R_t$  and  $I_t$  denote now the real and imaginary parts of  $\Psi_t^{(GK)}$ , respectively. From (26), we get:

$$\begin{cases} R_t = \Psi_c^{(R)} \eta x_t + R_t^{(cl)} ; dR_t = \Psi_c^{(R)} \eta dx_t + dR_t^{(cl)} \\ I_t = \Psi_c^{(I)} \eta x_t + I_t^{(cl)} ; dI_t = \Psi_c^{(I)} \eta dx_t + dI_t^{(cl)}. \end{cases} \quad (27)$$

Again we derive the SDE for  $R_t$  and that of  $I_t$  will follow immediately. From equations (27), (21) and (12), we get:

$$\begin{aligned} dR_t = & \left( \eta \Psi_c^{(R)} \mathcal{A}(1 - x_t) - \frac{\mathcal{A} + \mathcal{B}}{2} R_t^{(cl)} + \frac{\mathcal{A} R_t^{(cl)}}{2x_t} \left( 1 - \frac{1}{2\alpha} \right) \right) dt \\ & + \left[ \frac{R_t^{(cl)}}{x_t^{1/2}} \left( \frac{\mathcal{A}}{2\alpha} \right)^{1/2} + \eta \Psi_c^{(R)} \left( \frac{2\mathcal{A}x_t}{\alpha} \right)^{1/2} \left( \frac{\mathcal{B}x_t}{2} \right)^{1/2} \right] \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \end{bmatrix}. \end{aligned} \quad (28)$$

Replacing  $R_t^{(cl)} = R_t - \Psi_c^{(R)} \eta x_t$  in equation (28) and applying the same procedure for  $I_t$ , we get the following system of coupled SDE that describe GK scattering:

$$\begin{aligned} \begin{bmatrix} dx_t \\ dR_t \\ dI_t \end{bmatrix} = & \begin{bmatrix} \mathcal{A}(1 - x_t) \\ \eta \Psi_c^{(R)} \mathcal{A}(1 - x_t) + \left( R_t - \Psi_c^{(R)} \eta x_t \right) \left( -\frac{\mathcal{A} + \mathcal{B}}{2} + \frac{\mathcal{A}}{2x_t} \left( 1 - \frac{1}{2\alpha} \right) \right) \\ \eta \Psi_c^{(I)} \mathcal{A}(1 - x_t) + \left( I_t - \Psi_c^{(I)} \eta x_t \right) \left( -\frac{\mathcal{A} + \mathcal{B}}{2} + \frac{\mathcal{A}}{2x_t} \left( 1 - \frac{1}{2\alpha} \right) \right) \end{bmatrix} dt \\ + & \begin{bmatrix} \left( \frac{2\mathcal{A}x_t}{\alpha} \right)^{1/2} & 0 & 0 \\ \frac{R_t - \Psi_c^{(R)} \eta x_t}{x_t^{1/2}} \left( \frac{\mathcal{A}}{2\alpha} \right)^{1/2} + \eta \Psi_c^{(R)} \left( \frac{2\mathcal{A}x_t}{\alpha} \right)^{1/2} \left( \frac{\mathcal{B}x_t}{2} \right)^{1/2} & 0 & 0 \\ \frac{I_t - \Psi_c^{(I)} \eta x_t}{x_t^{1/2}} \left( \frac{\mathcal{A}}{2\alpha} \right)^{1/2} + \eta \Psi_c^{(I)} \left( \frac{2\mathcal{A}x_t}{\alpha} \right)^{1/2} & 0 & \left( \frac{\mathcal{B}x_t}{2} \right)^{1/2} \end{bmatrix} \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \\ dW_t^{(I)} \end{bmatrix}. \end{aligned} \quad (29)$$

The SDE obtained for GK scattering, equation (29), can also be written in the more compact form:

$$\begin{bmatrix} dx_t \\ dR_t \\ dI_t \end{bmatrix} = \beta_{\Psi_c}^{(GK)}(x_t, R_t, I_t) dt + \Sigma_{\Psi_c}^{(GK)}(x_t, R_t, I_t) \begin{bmatrix} dW_t^{(x)} \\ dW_t^{(R)} \\ dW_t^{(I)} \end{bmatrix}. \quad (30)$$

Vector  $\beta_{\Psi_c}^{(GK)}$  and matrix  $\Sigma_{\Psi_c}^{(GK)}$  are respectively the drift and volatility in the case of GK scattering.

#### 4. Maximum likelihood estimation of $\Psi_c$

Let  $\Psi_c = \Psi_c^{(R)} + i\Psi_c^{(I)}$  be the target constant. We remind from section 3 that the target reflectivity is simply the target constant  $\Psi_c$  for HK scattering, and is the target constant times the SCS and the coupling factor, *i.e.*  $\Psi_c \eta x_t$  for GK scattering. For both cases, our aim is to estimate  $\Psi_c$ . It is also assumed that there is always a target, since even the absence of target can be seen as the special case  $\Psi_c = 0$ .

We assume in this section that we observe three discrete time series:  $(\tilde{x}, \tilde{R}, \tilde{I}) = \{(\tilde{x}_k, \tilde{R}_k, \tilde{I}_k), k = 0, \dots, n\}$ . The measurements are made at times  $t_k$  with  $k$  ranging from 0 to  $n$ . We also assume for simplicity that  $\forall k, t_k - t_{k-1} = \Delta t$  is a constant.

As explained in section 2.2, the maximum likelihood (ML) estimation consists in maximizing the likelihood function with respect to the parameter  $\Psi_c$ . The likelihood can be written:

$$\mathcal{L}(\tilde{x}, \tilde{R}, \tilde{I}; \Psi_c) = p_{\Psi_c}^{\infty} \prod_{k=1}^n p_{\Psi_c}^{(k)}, \quad (31)$$

with

$$\begin{aligned} p_{\Psi_c}^{\infty} &= p_{\Psi_c}((x_{t_0}, R_{t_0}, I_{t_0}) = (\tilde{x}_0, \tilde{R}_0, \tilde{I}_0)) \\ &= p_{\Psi_c}^{\infty}(\tilde{x}_0, \tilde{R}_0, \tilde{I}_0) \end{aligned} \quad (32)$$

and

$$p_{\Psi_c}^{(k)} = p_{\Psi_c} \left( (x_{t_k}, R_{t_k}, I_{t_k}) = (\tilde{x}_k, \tilde{R}_k, \tilde{I}_k) \mid (x_{t_{k-1}}, R_{t_{k-1}}, I_{t_{k-1}}) = (\tilde{x}_{k-1}, \tilde{R}_{k-1}, \tilde{I}_{k-1}) \right). \quad (33)$$

$p_{\Psi_c}^{\infty}$  is the stationary (asymptotic) distribution of the process  $(x_t, R_t, I_t)$ . It accounts for the initial value at time  $t_0$  and it is implicit that it is evaluated at  $(\tilde{x}_0, \tilde{R}_0, \tilde{I}_0)$ .

$p_{\Psi_c}^{(k)}$  is the transition probability between times  $t_{k-1}$  and  $t_k$ .  $\mathcal{L}(\tilde{x}, \tilde{R}, \tilde{I}; \Psi_c)$  is therefore the joint probability (density) of the observed initial value and of the  $n$  transitions occurring from  $t_0$  to  $t_n$ . We assume that the parameters of the clutter, *i.e.*  $\mathcal{A}, \mathcal{B}, \mathcal{C}$  and  $\alpha$  are known. In that case, for a given observed trajectory,  $\mathcal{L}(\tilde{x}, \tilde{R}, \tilde{I}; \Psi_c)$  depends only on  $\Psi_c$ . Maximizing it with respect to  $\Psi_c$  yields the estimated target constant  $\tilde{\Psi}_c$ .

Instead of maximizing directly the likelihood function  $\mathcal{L}$ , we choose to maximize

its logarithm:

$$\begin{aligned} l(\tilde{x}, \tilde{R}, \tilde{I}; \Psi_c) &= \ln(\mathcal{L}(\tilde{x}, \tilde{R}, \tilde{I}; \Psi_c)) \\ &= \ln(p_{\Psi_c}^\infty) + \sum_{k=1}^n \ln(p_{\Psi_c}^{(k)}). \end{aligned} \quad (34)$$

If  $\tilde{x}, \tilde{R}, \tilde{I}$  are fixed, *i.e.* a time series is observed, the necessary conditions of optimality are:

$$\begin{cases} \frac{\partial l}{\partial \Psi_c^{(R)}}(\tilde{\Psi}_c^{(R)}, \tilde{\Psi}_c^{(I)}) = 0 \\ \frac{\partial l}{\partial \Psi_c^{(I)}}(\tilde{\Psi}_c^{(R)}, \tilde{\Psi}_c^{(I)}) = 0. \end{cases} \quad (35)$$

#### 4.1. ML estimation of $\Psi_c$ in HK scattering

From equations (34) and (35), it is obvious that one must know  $p_{\Psi_c}^\infty$  and the transition probabilities  $p_{\Psi_c}^{(k)}$  to estimate  $\Psi_c$ .

##### 4.1.1. Asymptotic distribution

Let  $\Psi_t^{(HK)} = \Psi_c^{(R)} + i\Psi_c^{(I)} + x_t^{1/2}\gamma_t$ . At fixed SCS  $x_t = x$ , the random variable  $(R_t, I_t)$  is a bivariate gaussian distribution (see [4] equation (4.12)):

$$p_{\Psi_c}((R_t, I_t) = (R, I) | x_t = x) = \frac{1}{\pi x} \exp\left(-\frac{(R - \Psi_c^{(R)})^2 + (I - \Psi_c^{(I)})^2}{x}\right). \quad (36)$$

In that case,  $(R_t^2 + I_t^2)^{1/2}$  follows the so-called Rice distribution. From equations (14) and (36), we get:

$$p_{\Psi_c}^\infty(R, I, x) = \frac{1}{\pi x} \exp\left(-\frac{(R - \Psi_c^{(R)})^2 + (I - \Psi_c^{(I)})^2}{x}\right) \times \frac{\alpha^\alpha x^{\alpha-1} e^{-\alpha x}}{\Gamma(\alpha)}. \quad (37)$$

##### 4.1.2. Approximate transition probabilities

Deriving the transition probabilities requires some tedious calculus. Exact transition probabilities could be derived by solving the Fokker-Planck equation associated with (24) (see [15] for a detailed account of the Fokker-Planck equation). In practice, it is rarely possible to solve it analytically. Instead, if the time difference  $\Delta t$  between  $t_{k-1}$  and  $t_k$  is small enough, one can approximate the transition probability by a multivariate gaussian distribution. The starting point is Euler-Maruyama's scheme applied to the SDE (25):

$$\begin{aligned}
\begin{bmatrix} \Delta x_t \\ \Delta R_t \\ \Delta I_t \end{bmatrix} &= \beta_{\Psi_c}^{(HK)}(x_t, R_t, I_t) \Delta t + \Sigma_{\Psi_c}^{(HK)}(x_t, R_t, I_t) \begin{bmatrix} \Delta W_t^{(x)} \\ \Delta W_t^{(R)} \\ \Delta W_t^{(I)} \end{bmatrix} \\
\Leftrightarrow \begin{bmatrix} x_{t+\Delta t} \\ R_{t+\Delta t} \\ I_{t+\Delta t} \end{bmatrix} &= \begin{bmatrix} x_t \\ R_t \\ I_t \end{bmatrix} + \beta_{\Psi_c}^{(HK)}(x_t, R_t, I_t) \Delta t + \Sigma_{\Psi_c}^{(HK)}(x_t, R_t, I_t) \Delta t^{1/2} \begin{bmatrix} n_x \\ n_R \\ n_I \end{bmatrix} \quad (38)
\end{aligned}$$

where  $[n_x \ n_R \ n_I]^T$  is a vector of independent standard gaussian random variables. It follows that:

$$p_{\Psi_c}^{(k)} \approx \frac{1}{(2\pi)^{3/2} |\sigma_{\Psi_c}^{(HK)} \Delta t|^{1/2}} \exp\left(-\frac{1}{2}(v_k - \mu_k)^T (\sigma_{\Psi_c}^{(HK)} \Delta t)^{-1} (v_k - \mu_k)\right), \quad (39)$$

with

$$\begin{cases} v_k = [\tilde{x}_k \ \tilde{R}_k \ \tilde{I}_k]^T \\ \mu_k = [\tilde{x}_{k-1} \ \tilde{R}_{k-1} \ \tilde{I}_{k-1}]^T + \beta_{\Psi_c}^{(HK)}(\tilde{x}_{k-1}, \tilde{R}_{k-1}, \tilde{I}_{k-1}) \Delta t \\ \sigma_{\Psi_c}^{(HK)} = \Sigma_{\Psi_c}^{(HK)} \Sigma_{\Psi_c}^{(HK)T}. \end{cases} \quad (40)$$

Indeed, to get the transition probability, we fix  $[x_t \ R_t \ I_t]^T = [\tilde{x}_{k-1} \ \tilde{R}_{k-1} \ \tilde{I}_{k-1}]^T$ . Thus, the only random part in the right hand side of equation (38) is the gaussian vector  $[n_x \ n_R \ n_I]^T$ .  $\sigma_{\Psi_c}^{(HK)}$  is referred to as the squared volatility. It is a positive definite symmetric matrix in the non degenerate case. We can show that:

$$\sigma_{\Psi_c}^{(HK)} = \begin{bmatrix} \frac{2\mathcal{A}x_t}{\alpha} & \frac{\mathcal{A}(R_t - \Psi_c^{(R)})}{\alpha} & \frac{\mathcal{A}(I_t - \Psi_c^{(I)})}{\alpha} \\ \frac{\mathcal{A}(R_t - \Psi_c^{(R)})}{\alpha} & \frac{\mathcal{A}}{2\alpha} \frac{(R_t - \Psi_c^{(R)})^2}{x_t} + \frac{\mathcal{B}x_t}{2} & \frac{\mathcal{A}}{2\alpha} \frac{(R_t - \Psi_c^{(R)})(I_t - \Psi_c^{(I)})}{x_t} \\ \frac{\mathcal{A}(I_t - \Psi_c^{(I)})}{\alpha} & \frac{\mathcal{A}}{2\alpha} \frac{(R_t - \Psi_c^{(R)})(I_t - \Psi_c^{(I)})}{x_t} & \frac{\mathcal{A}}{2\alpha} \frac{(I_t - \Psi_c^{(I)})^2}{x_t} + \frac{\mathcal{B}x_t}{2} \end{bmatrix}. \quad (41)$$

We have very simply that  $|\sigma_{\Psi_c}^{(HK)} \Delta t| = |\Sigma_{\Psi_c}^{(HK)}|^2 \Delta t^3 = \frac{\Delta t^3 \mathcal{A} \mathcal{B}^2 x_t^3}{2\alpha}$ . We can then invert  $\sigma_{\Psi_c}^{(HK)}$  and we get after some calculations:

$$\left(\sigma_{\Psi_c}^{(HK)} \Delta t\right)^{-1} = \begin{bmatrix} \frac{(R_t - \Psi_c^{(R)})^2 + (I_t - \Psi_c^{(I)})^2}{2\mathcal{B}\Delta t x_t^3} + \frac{\alpha}{2\mathcal{A}\Delta t x_t} & -\frac{R_t - \Psi_c^{(R)}}{\mathcal{B}\Delta t x_t^2} & -\frac{I_t - \Psi_c^{(I)}}{\mathcal{B}\Delta t x_t^2} \\ -\frac{R_t - \Psi_c^{(R)}}{\mathcal{B}\Delta t x_t^2} & \frac{2}{\mathcal{B}\Delta t x_t} & 0 \\ -\frac{I_t - \Psi_c^{(I)}}{\mathcal{B}\Delta t x_t^2} & 0 & \frac{2}{\mathcal{B}\Delta t x_t} \end{bmatrix}. \quad (42)$$

Note that in equation (39),  $(\sigma_{\Psi_c}^{(HK)} \Delta t)^{-1}$  must be evaluated at  $t = t_{k-1}$ , *i.e.*  $[x_t, R_t, I_t] = [\tilde{x}_{k-1}, \tilde{R}_{k-1}, \tilde{I}_{k-1}]$  respectively in equation (41). Using (34), (37) and

(39), we have:

$$l(\tilde{x}, \tilde{R}, \tilde{I}; \Psi_c) = \ln \left( \frac{\alpha^\alpha \tilde{x}_0^{\alpha-1} e^{-\alpha \tilde{x}_0}}{\pi \tilde{x}_0 \Gamma(\alpha)} \right) - \frac{(\tilde{R}_0 - \Psi_c^{(R)})^2 + (\tilde{I}_0 - \Psi_c^{(I)})^2}{\tilde{x}_0} - n \ln \left( (2\pi)^{3/2} \left| \sigma_{\Psi_c}^{(HK)} \Delta t \right|^{1/2} \right) + \sum_{k=1}^n \Phi_k \quad (43)$$

with

$$\Phi_k = -\frac{1}{2} (v_k - \mu_k)^T \left( \sigma_{\Psi_c}^{(HK)} \Delta t \right)^{-1} (v_k - \mu_k). \quad (44)$$

If we express the first optimality condition in (35), we get after some calculations:

$$\begin{aligned} & \frac{\partial l}{\partial \Psi_c^{(R)}}(\tilde{\Psi}_c^{(R)}, \tilde{\Psi}_c^{(I)}) = 0 \\ \Leftrightarrow & -\frac{2(\tilde{\Psi}_c^{(R)} - \tilde{R}_0)}{\tilde{x}_0} + \sum_{k=1}^n \frac{\tilde{\Psi}_c^{(R)} - \tilde{R}_{k-1}}{\mathcal{B} \Delta t \tilde{x}_{k-1}} \left( -\frac{w_k^2}{2} - 2\gamma_k^2 + \frac{2\gamma_k w_k}{\tilde{x}_{k-1}} \right) \\ & + \sum_{k=1}^n \frac{\tilde{R}_k - \tilde{R}_{k-1}}{\mathcal{B} \Delta t \tilde{x}_{k-1}} \left( 2\gamma_k - \frac{w_k}{\tilde{x}_{k-1}} \right) = 0, \end{aligned} \quad (45)$$

with

$$\begin{cases} w_k = \tilde{x}_k - \tilde{x}_{k-1} - \mathcal{A} \Delta t (1 - \tilde{x}_{k-1}) \\ \gamma_k = \frac{\mathcal{A} + \mathcal{B}}{2} - \frac{\mathcal{A}}{2\tilde{x}_{k-1}} \left( 1 - \frac{1}{2\alpha} \right). \end{cases} \quad (46)$$

We see that equation (45) depends only on  $x_t$  and the real part  $R_t$  of the reflectivity. It gives a condition on  $\tilde{\Psi}_c^{(R)}$  only. It is then straightforward to get the estimation of  $\Psi_c^{(R)}$ :

$$\tilde{\Psi}_c^{(R)} = \frac{-\frac{2\tilde{R}_0}{\tilde{x}_0} + \sum_{k=1}^n \frac{\tilde{R}_{k-1}}{\mathcal{B} \Delta t \tilde{x}_{k-1}} \left( -\frac{w_k^2}{2} - 2\gamma_k^2 + \frac{2\gamma_k w_k}{\tilde{x}_{k-1}} \right) - \sum_{k=1}^n \frac{\tilde{R}_k - \tilde{R}_{k-1}}{\mathcal{B} \Delta t \tilde{x}_{k-1}} \left( 2\gamma_k - \frac{w_k}{\tilde{x}_{k-1}} \right)}{\sum_{k=1}^n \frac{1}{\mathcal{B} \Delta t \tilde{x}_{k-1}} \left( -\frac{w_k^2}{2} - 2\gamma_k^2 + \frac{2\gamma_k w_k}{\tilde{x}_{k-1}} \right) - \frac{2}{\tilde{x}_0}} \quad (47)$$

The estimator for  $\Psi_c^{(I)}$  follows the same equation with  $\tilde{I}$  replacing  $\tilde{R}$ . Note that  $w_k$  and  $\gamma_k$  are common to both the estimators of  $\Psi_c^{(R)}$  and  $\Psi_c^{(I)}$ .

#### 4.2. ML estimation of $\Psi_c$ in GK scattering

As seen below, the estimator for  $\Psi_c$  in the GK scattering case is obtained by the same procedure as for HK scattering.

#### 4.2.1. Asymptotic distribution

Let  $\Psi_t^{(GK)} = (\Psi_c^{(R)} + i\Psi_c^{(I)}) \eta x_t + x_t^{1/2} \gamma_t$ . At fixed SCS  $x_t = x$ , the random variable  $(R_t, I_t)$  is a bivariate gaussian distribution (see [4] equation (4.12)):

$$p_{\Psi_c}((R_t, I_t) = (R, I) | x_t = x) = \frac{1}{\pi x} \exp\left(-\frac{(R - \Psi_c^{(R)} \eta x)^2 + (I - \Psi_c^{(I)} \eta x)^2}{x}\right). \quad (48)$$

From equations (14) and (48), we get:

$$p_{\Psi_c}^{\infty}(R, I, x) = \frac{1}{\pi x} \exp\left(-\frac{(R - \Psi_c^{(R)} \eta x)^2 + (I - \Psi_c^{(I)} \eta x)^2}{x}\right) \times \frac{\alpha^\alpha x^{\alpha-1} e^{-\alpha x}}{\Gamma(\alpha)}. \quad (49)$$

#### 4.2.2. Approximate transition probabilities

To derive approximate transition probabilities, we dodge the complications of the Fokker-Planck equation again and compute gaussian approximations. If we apply Euler-Maruyama's scheme to (30), we get:

$$\begin{aligned} \begin{bmatrix} \Delta x_t \\ \Delta R_t \\ \Delta I_t \end{bmatrix} &= \beta_{\Psi_c}^{(GK)}(x_t, R_t, I_t) \Delta t + \Sigma_{\Psi_c}^{(GK)}(x_t, R_t, I_t) \begin{bmatrix} \Delta W_t^{(x)} \\ \Delta W_t^{(R)} \\ \Delta W_t^{(I)} \end{bmatrix} \\ \Leftrightarrow \begin{bmatrix} x_{t+\Delta t} \\ R_{t+\Delta t} \\ I_{t+\Delta t} \end{bmatrix} &= \begin{bmatrix} x_t \\ R_t \\ I_t \end{bmatrix} + \beta_{\Psi_c}^{(GK)}(x_t, R_t, I_t) \Delta t + \Sigma_{\Psi_c}^{(GK)}(x_t, R_t, I_t) \Delta t^{1/2} \begin{bmatrix} n_x \\ n_R \\ n_I \end{bmatrix} \end{aligned} \quad (50)$$

where  $[n_x \ n_R \ n_I]^T$  is a vector of independent standard gaussian random variables. It follows that:

$$p_{\Psi_c}^{(k)} \approx \frac{1}{(2\pi)^{3/2} |\sigma_{\Psi_c}^{(GK)} \Delta t|^{1/2}} \exp\left(-\frac{1}{2} (v_k - \mu_k)^T (\sigma_{\Psi_c}^{(GK)} \Delta t)^{-1} (v_k - \mu_k)\right), \quad (51)$$

with

$$\begin{cases} v_k = [\tilde{x}_k \ \tilde{R}_k \ \tilde{I}_k]^T \\ \mu_k = [\tilde{x}_{k-1} \ \tilde{R}_{k-1} \ \tilde{I}_{k-1}]^T + \beta_{\Psi_c}^{(GK)}(\tilde{x}_{k-1}, \tilde{R}_{k-1}, \tilde{I}_{k-1}) \Delta t \\ \sigma_{\Psi_c}^{(GK)} = \Sigma_{\Psi_c}^{(GK)} \Sigma_{\Psi_c}^{(GK)T}. \end{cases} \quad (52)$$

The squared volatility  $\sigma_{\Psi_c}^{(GK)}$  is also a positive definite symmetric matrix in the non degenerate case. We can show that:

$$\sigma_{\Psi_c}^{(GK)} = \begin{bmatrix} \frac{2\mathcal{A}x_t}{\alpha} & \frac{\mathcal{A}(R_t + \Psi_c^{(R)} \eta x_t)}{\alpha} & \frac{\mathcal{A}(I_t + \Psi_c^{(I)} \eta x_t)}{\alpha} \\ \frac{\mathcal{A}(R_t + \Psi_c^{(R)} \eta x_t)}{\alpha} & \frac{\mathcal{A}(R_t + \Psi_c^{(R)} \eta x_t)^2}{2\alpha} + \frac{\mathcal{B}x_t}{2} & \frac{\mathcal{A}(R_t + \Psi_c^{(R)} \eta x_t)(I_t + \Psi_c^{(I)} \eta x_t)}{2\alpha} \\ \frac{\mathcal{A}(I_t + \Psi_c^{(I)} \eta x_t)}{\alpha} & \frac{\mathcal{A}(R_t + \Psi_c^{(R)} \eta x_t)(I_t + \Psi_c^{(I)} \eta x_t)}{2\alpha} & \frac{\mathcal{A}(I_t + \Psi_c^{(I)} \eta x_t)^2}{2\alpha} + \frac{\mathcal{B}x_t}{2} \end{bmatrix}. \quad (53)$$



We have again that  $\left| \sigma_{\Psi_c}^{(GK)} \Delta t \right| = \left| \Sigma_{\Psi_c}^{(GK)} \right|^2 \Delta t^3 = \frac{\Delta t^3 \mathcal{A} \mathcal{B}^2 x_t^3}{2\alpha}$ . We can invert  $\sigma_{\Psi_c}^{(GK)}$  and we get after some calculations:

$$\left( \sigma_{\Psi_c}^{(GK)} \Delta t \right)^{-1} = \begin{bmatrix} \frac{(R_t + \Psi_c^{(R)} \eta x_t)^2 + (I_t + \Psi_c^{(I)} \eta x_t)^2}{2\mathcal{B} \Delta t x_t^3} + \frac{\alpha}{2\mathcal{A} \Delta t x_t} & -\frac{R_t + \Psi_c^{(R)} \eta x_t}{\mathcal{B} \Delta t x_t^2} & -\frac{I_t + \Psi_c^{(I)} \eta x_t}{\mathcal{B} \Delta t x_t^2} \\ -\frac{R_t + \Psi_c^{(R)} \eta x_t}{\mathcal{B} \Delta t x_t^2} & \frac{2}{\mathcal{B} \Delta t x_t} & 0 \\ -\frac{I_t + \Psi_c^{(I)} \eta x_t}{\mathcal{B} \Delta t x_t^2} & 0 & \frac{2}{\mathcal{B} \Delta t x_t} \end{bmatrix}. \quad (54)$$

In equation (51),  $\left( \sigma_{\Psi_c}^{(GK)} \Delta t \right)^{-1}$  must be evaluated at  $t = t_{k-1}$ , *i.e.*  $[x_t, R_t, I_t] = [\tilde{x}_{k-1}, \tilde{R}_{k-1}, \tilde{I}_{k-1}]$  in equation (53). Using (34), (49) and (51), we have:

$$l(\tilde{x}, \tilde{R}, I; \tilde{\Psi}_c) = \ln \left( \frac{\alpha \tilde{x}_0^{\alpha-1} e^{-\alpha \tilde{x}_0}}{\pi \tilde{x}_0 \Gamma(\alpha)} \right) - \frac{(\tilde{R}_0 - \Psi_c^{(R)} \eta \tilde{x}_0)^2 + (\tilde{I}_0 - \Psi_c^{(I)} \eta \tilde{x}_0)^2}{\tilde{x}_0} - n \ln \left( (2\pi)^{3/2} \left| \sigma_{\Psi_c}^{(HK)} \Delta t \right|^{1/2} \right) + \sum_{k=1}^n \Phi_k \quad (55)$$

with

$$\Phi_k = -\frac{1}{2} (v_k - \mu_k)^T \left( \sigma_{\Psi_c}^{(HK)} \Delta t \right)^{-1} (v_k - \mu_k). \quad (56)$$

If we express the first optimality condition in (35), we get after some calculations:

$$\frac{\partial l}{\partial \Psi_c^{(R)}} (\tilde{\Psi}_c^{(R)}, \tilde{\Psi}_c^{(I)}) = 0 \quad (57)$$

$$\Leftrightarrow -\frac{2\eta \left( \tilde{\Psi}_c^{(R)} \eta \tilde{x}_0 - \tilde{R}_0 \right)}{\tilde{x}_0} + \sum_{k=1}^n \lambda_k^{(1)} + \tilde{\Psi}_c^{(R)} \sum_{k=1}^n \lambda_k^{(2)} = 0 \quad (58)$$

with

$$\begin{aligned} \lambda_k^{(1)} &= \frac{-w_k^2 \eta \tilde{R}_{k-1}}{2\mathcal{B} \Delta t \tilde{x}_{k-1}^2} - \frac{2}{\mathcal{B} \Delta t \tilde{x}_{k-1}} \left( \tilde{R}_k - \tilde{R}_{k-1} - \gamma_k \Delta t \tilde{R}_{k-1} \right) \\ &\times (-\eta \mathcal{A} \Delta t (1 - \tilde{x}_{k-1}) + \gamma_k \Delta t \eta \tilde{x}_{k-1}) + \frac{w_k}{\mathcal{B} \Delta t \tilde{x}_{k-1}^2} \eta \tilde{x}_{k-1} \left( \tilde{R}_k - \tilde{R}_{k-1} - \gamma_k \Delta t \tilde{R}_{k-1} \right) \\ &+ \frac{w_k}{\mathcal{B} \Delta t \tilde{x}_{k-1}^2} \tilde{R}_{k-1} (-\eta \mathcal{A} \Delta t (1 - \tilde{x}_{k-1}) + \gamma_k \Delta t \eta \tilde{x}_{k-1}) \end{aligned} \quad (59)$$

and

$$\begin{aligned} \lambda_k^{(2)} &= \frac{-w_k^2 \eta^2}{2\mathcal{B} \Delta t \tilde{x}_{k-1}} - \frac{2}{\mathcal{B} \Delta t \tilde{x}_{k-1}} (-\eta \mathcal{A} \Delta t (1 - \tilde{x}_{k-1}) + \gamma_k \Delta t \eta \tilde{x}_{k-1}) \\ &\times (-\eta \mathcal{A} \Delta t (1 - \tilde{x}_{k-1}) + \gamma_k \Delta t \eta \tilde{x}_{k-1}) \\ &+ \frac{2w_k \eta}{\mathcal{B} \Delta t \tilde{x}_{k-1}} (-\eta \mathcal{A} \Delta t (1 - \tilde{x}_{k-1}) + \gamma_k \Delta t \eta \tilde{x}_{k-1}) \end{aligned} \quad (60)$$

Equation (57) depends only on  $x_t$  and the real part  $R_t$  of the reflectivity. It gives

a condition on  $\tilde{\Psi}_c^{(R)}$  only. It is straightforward to get the estimation of  $\Psi_c^{(R)}$ :

$$\tilde{\Psi}_c^{(R)} = \frac{-\frac{2\eta\tilde{R}_0}{\tilde{x}_0} - \sum_{k=1}^n \lambda_k^{(1)}}{-2\eta^2 + \sum_{k=1}^n \lambda_k^{(2)}} \quad (61)$$

The estimator for  $\Psi_c^{(I)}$  follows the same equation with  $\tilde{I}$  replacing  $\tilde{R}$ . Note that  $w_k$  and  $\gamma_k$  are common to both the estimators of  $\Psi_c^{(R)}$  and  $\Psi_c^{(I)}$ .

## 5. Performance of the ML estimation

### 5.1. A simple estimator for $\Psi_c$

As seen in the previous section, estimating  $\Psi_c$  with ML is straightforward but involves quite heavy expressions, at least as compared to the 'ergodicity' estimator. This very simple estimator arises naturally. Indeed from equations (22) and (26) we get:

$$\begin{cases} \mathbb{E} \left[ \Psi_c^{(HK)} \right] = \mathbb{E} \left[ \Psi_c + x_t^{1/2} \gamma_t \right] = \Psi_c + \mathbb{E}[x_t^{1/2}] \mathbb{E}[\gamma_t] = \Psi_c \\ \mathbb{E} \left[ \Psi_c^{(GK)} \right] = \mathbb{E} \left[ \Psi_c \eta x_t + x_t^{1/2} \gamma_t \right] = \eta \Psi_c \mathbb{E}[x_t] + \mathbb{E}[x_t^{1/2}] \mathbb{E}[\gamma_t] = \eta \Psi_c. \end{cases} \quad (62)$$

We have used that  $x_t$  and  $\gamma_t$  are independent, and that  $\mathbb{E}[x_t] = 1$  and  $\mathbb{E}[\gamma_t] = 0$  (see equation (14)). We assume that the reflectivity is ergodic, or more precisely we assume that:

$$\mathbb{E} \left[ \Psi_t^{(HK)} \right] = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \tilde{\Psi}_t^{(HK)} dt \quad (63)$$

for any particular trajectory  $\tilde{\Psi}_t$  of the reflectivity, in which case:

$$\Psi_c = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \tilde{\Psi}_t^{(HK)} dt \approx \frac{1}{t_n} \int_0^{t_n} \tilde{\Psi}_t^{(HK)} dt \approx \frac{1}{n\Delta t} \sum_{i=0}^{n-1} \tilde{\Psi}_i^{(HK)} \Delta t.$$

Of course, the same holds in the GK scattering case, such that the ergodicity estimator yields:

$$\begin{cases} \hat{\Psi}_c = \frac{1}{n\Delta t} \sum_{i=0}^{n-1} \tilde{\Psi}_i^{(HK)} \Delta t & (HK \text{ scattering}) \\ \hat{\Psi}_c = \frac{1}{n\Delta t} \sum_{i=0}^{n-1} \tilde{\Psi}_i^{(GK)} \Delta t & (GK \text{ scattering}). \end{cases} \quad (64)$$

### 5.2. Numerical simulations

In this section, we compare the ML estimator to the ergodicity estimator which serves as a reference. What we really want is to test numerically whether it is relevant to use the ML estimator, whose formalism is based on transition probabilities. To assess the abilities of the ML and ergodicity estimators, we simulate many trajectories of  $(x_t, R_t, I_t)$  with a given  $\Psi_c$  and then try to retrieve it. To do so, we solve numerically the SDE (24) and (29) using Euler-Maruyama's scheme for  $R_t, I_t$  and Milstein's scheme for  $x_t$  (see [12] for Milstein's scheme). The Euler-Maruyama

scheme is not used for  $x_t$  because it raises numerical issues. Indeed, it could generate negative values, which is absurd since the SCS is always positive. As explained in section 2.4, we set:  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz and  $\alpha = 1$ . A realistic value for  $\Delta t$  is 0.001 s since the Pulse Repetition Frequency is about 1 kHz for satellite and airborne applications [16]. In section 2.4, we also said that we could set  $C = 1$ , which given equation (15) means that the clutter has unit power. This is not restrictive, since if  $C \neq 1$ , we can work with  $\Psi_t^{(HK)}/C$  or  $\Psi_t^{(GK)}/C$ , in which case the target constant become  $\frac{\Psi_c}{C}$ . In our experiments, the simulation timestep,  $\hat{\Delta}t$ , should be much smaller than the measurement timestep  $\Delta t$  quoted above. Otherwise, the numerical trajectory will not replicate adequately the dynamics of the original SDE, with its time dependent drifts and volatilities. Consequently, the generated trajectories are evaluated at times  $\hat{t}_k$  for  $k = 0 \dots mn$ , where  $m$  is the decimation ratio. Then, for the estimation stage, they are downsampled (*i.e* decimated) to the times  $t_i$  for  $i = 0 \dots n$ , with  $t_i = \hat{t}_{mi}$ . In our numerical simulations,  $\Delta t$  being fixed at 0.001 s, we choose  $\hat{\Delta}t = 10^{-5}$  s leading to a decimation ratio  $m = 100$ . Note that we have verified that  $\hat{\Delta}t = 10^{-5}$  s is small enough: changing it to  $\hat{\Delta}t = 10^{-7}$  s does not alter the results presented here but it requires a much longer computing time.

The detection of a coherent scatterer is naturally more challenging if the coherent scatterer has a power similar to the clutter. We explore the dependence of the estimator performance to both the trajectory duration and target intensity (power). We define the target intensity simply as its squared modulus:

$$|\Psi_c|^2 = \Psi_c^{(R)2} + \Psi_c^{(I)2}. \quad (65)$$

By rotational symmetry and for simplicity, we also set  $\Psi_c^{(I)} = 0$  such that all the target power goes into the real part.

We denote  $\tilde{\Psi}_{c,i}^{(R)}$  the estimated  $\Psi_c^{(R)}$  from the  $i$ -th trajectory, and  $M$  the number of trajectories. We assess the performance of the estimators by computing their estimation bias  $b(\tilde{\Psi}_c^{(R)})$  and variance  $\sigma^2(\tilde{\Psi}_c^{(R)})$ :

$$b(\tilde{\Psi}_c^{(R)}) = \frac{1}{M} \sum_{i=1}^M (\tilde{\Psi}_{c,i}^{(R)} - \Psi_c^{(R)}), \quad (66)$$

$$\sigma^2(\tilde{\Psi}_c^{(R)}) = \frac{1}{M-1} \sum_{i=1}^M (\tilde{\Psi}_{c,i}^{(R)} - \Psi_c^{(R)})^2. \quad (67)$$

### 5.2.1. Performance in HK scattering

In the case of HK scattering (equation (22)), we numerically solve the SDE (24) for known clutter parameters  $\mathcal{A}, \mathcal{B}, \alpha$  and target  $\Psi_c = \Psi_c^{(R)}$ . We then try to retrieve  $\Psi_c^{(R)}$  using equations (47) and (64) which correspond respectively to the ML and ergodicity estimators. We explore the dependency of the estimators performances to both the trajectory duration (from 0.1 s to 10 s) and target intensity (from 0.1 to 100). When the duration is explored, the target intensity is set to 10, and when the target intensity is explored, the duration is set to 1 s. The relative squared estimation bias and the estimation variance are plotted in figure 1. They are simply the squared bias and variance defined in equation (66) normalized by the target intensity. We observe that both decrease quickly as the trajectory duration or target intensity increase. This is easily understandable since a longer trajectory carries more information, and a stronger target dominates more strongly the overall

reflectivity. However, the squared bias is orders of magnitude smaller than the variance and is therefore negligible. It is also what explains its apparent noisiness. If we now compare the ergodicity and ML estimators based on the variance, we notice a slight advantage for the ML estimator for low target intensity.

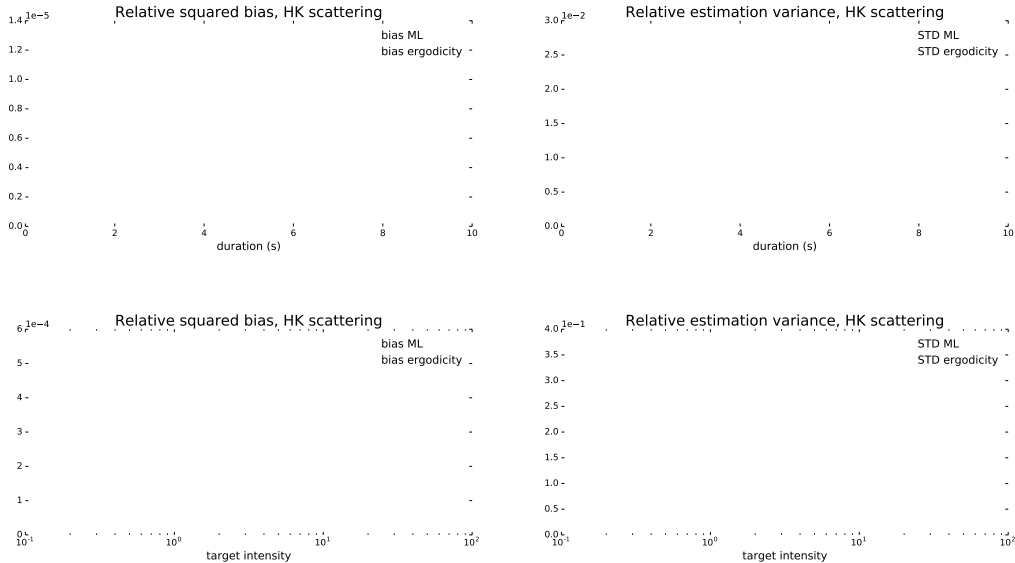


Figure 1. Relative estimation squared bias and variance of the target in HK scattering for both the ergodicity and ML estimators. 1000 trajectories are computed, with  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz and  $\alpha = 1$ . Up: dependence to trajectory duration with  $\Psi_c = \sqrt{10}$ . Down: dependence to target intensity with a duration of 1 s.

We represent in figure 2 a scatter plot of all the estimated  $\Psi_c^{(R)}$  for 1000 trajectories. There is no striking difference between the ergodicity and ML clouds.

### 5.2.2. Performance in GK scattering

In the case of GK scattering (equation (26)), we numerically solve the SDE (29) for known clutter parameters  $\mathcal{A}, \mathcal{B}, \alpha$ , coupling  $\eta$  and target  $\Psi_c = \Psi_c^{(R)}$ . We then try to retrieve  $\Psi_c^{(R)}$  using equations (61) and (64) which correspond respectively to the ML and ergodicity estimators. We explore the dependency of the estimators performances to both the trajectory duration (from 0.1 s to 10 s) and target intensity (from 0.1 to 100). Again, when the duration is explored, the target intensity is set to 10, and when the target intensity is explored, the duration is set to 1 s. The relative squared estimation bias and the estimation variance are plotted in figure 3. We also observe that the variance decreases quickly as the trajectory duration or target intensity increase. The bias is very noiselike but it is again orders of magnitude smaller than the variance so we shall ignore it. Unlike HK scattering, there is this time a large difference between the performance of the ergodicity and ML estimator. The ML estimator outperforms the ergodicity estimator by an order of magnitude on average (less for small durations and target intensity and vice versa).

We represent in figure 4 a scatter plot of all the estimated  $\Psi_c^{(R)}$  for 1000 trajectories. There is a striking difference between the ergodicity and ML clouds. There

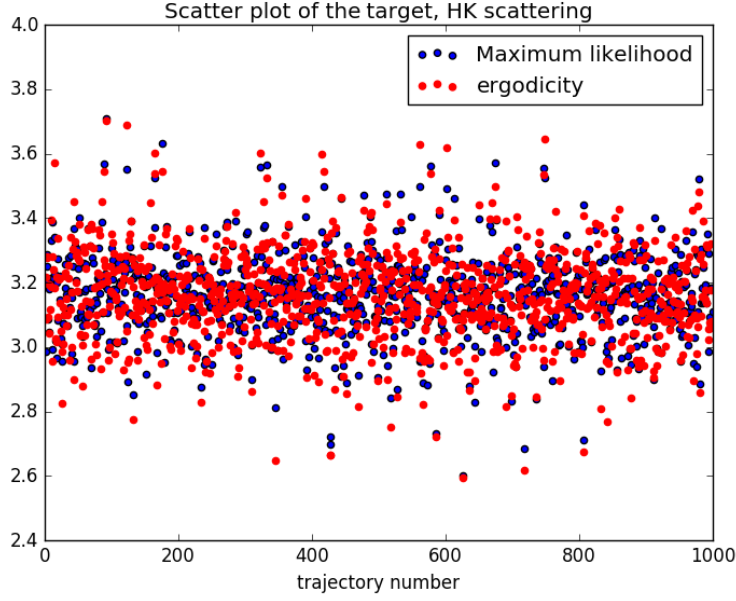


Figure 2. Scatter plot of the estimated target for both the ergodicity and ML estimators. 1000 trajectories of duration 1 s are computed, with  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz,  $\alpha = 1$  and  $\Psi_c^{(R)} = \sqrt{10}$ .

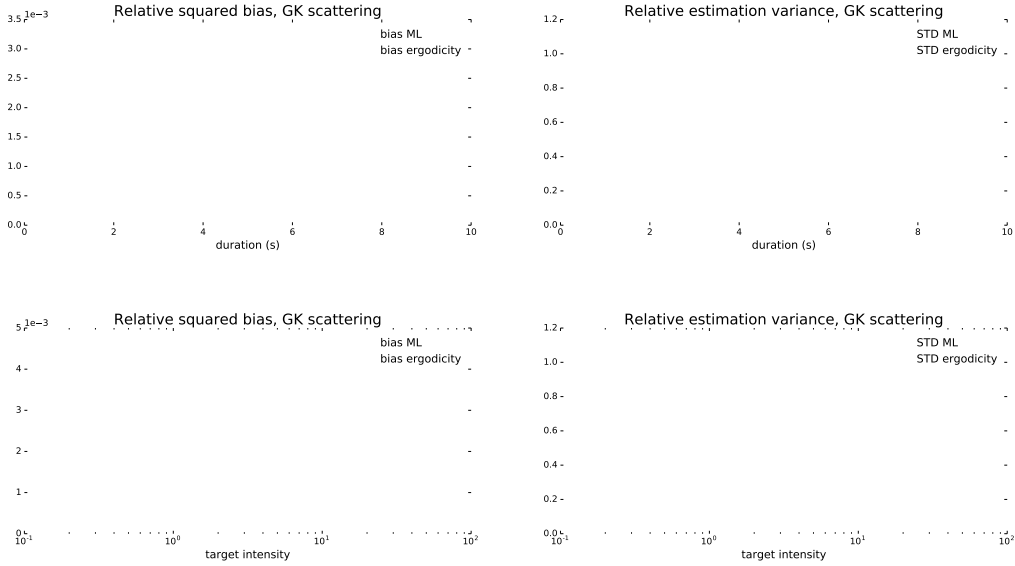


Figure 3. Relative estimation squared bias and variance of the target in GK scattering for both the ergodicity and ML estimators. 1000 trajectories are computed, with  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz,  $\alpha = 1$  and  $\eta = 1$ . Up: dependence to trajectory duration with  $\Psi_c = \sqrt{10}$ . Down: dependence to target intensity with a duration of 1 s..

are both centered correctly, but the ML cloud is much narrower than the ergodicity one, in accordance with the estimation variances. There is a second difference that we should mention: the ML estimator has the drawback of generating some outliers (on the order of 0.5%). These outliers have been filtered out for the variance calculations of figure 3. We noticed that they arise when during the trajectory,  $x_t$  goes very close to 0. Of course, it is not physically possible to have a vanishing SCS. However, we will leave the outliers in the section about decision theory, since

they merely increase the false alarm rate by a small percentage.

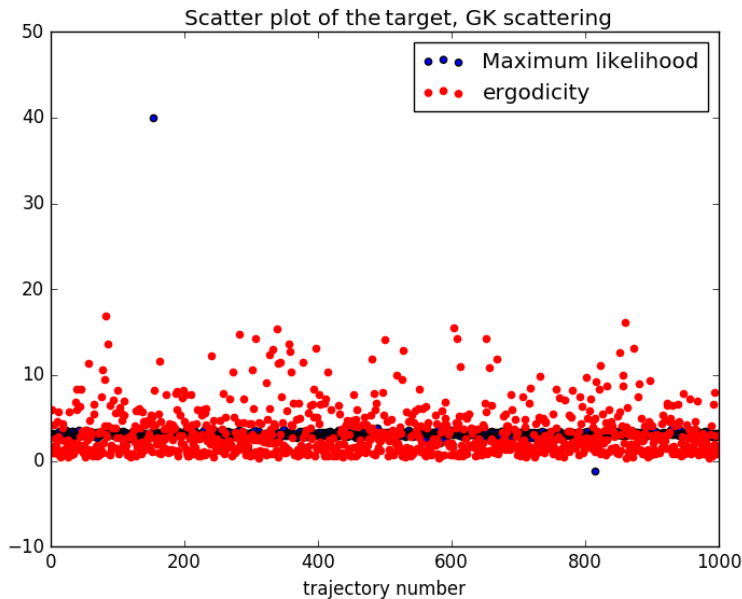


Figure 4. Scatter plot of the estimated target for both the ergodicity and ML estimators. 1000 trajectories of duration 1 s are computed, with  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz,  $\alpha = 1$ ,  $\eta = 1$  and  $\Psi_c^{(R)} = \sqrt{10}$ .

### 5.2.3. Comparison between HK and GK scattering

In HK scattering, the target reflectivity  $\Psi_c$  is simply added to the random medium reflectivity (clutter)  $\Psi_t = x_t^{1/2}\gamma_t$ . There is no interaction between the target and random medium. Retrospectively, it is then natural that the transition probabilities do not 'really' depend on the target and that ML estimation does not present any advantage. More precisely, we mean that the probability of increasing or decreasing by some amount during time interval  $\Delta t$  is independent on the target strength in HK scattering. In that case, ML estimation is not better than the ergodicity estimator, except in the case of a low intensity target.

In GK scattering, the target reflectivity is modulated by the SCS to give  $\Psi_c\eta x_t$  which is added to  $\Psi_t$ . There is an interaction between the target and random medium, and in this case the transition probabilities 'really' depend on the target. As a result, ML estimation is relevant since every single transition depends on  $\Psi_c$ . Comparing figure 1 and figure 3, we also see that the estimation variance is greater for GK than HK scattering, which is also a result of the target being modulated by  $x_t$  in GK scattering.

### 5.2.4. Performance as a function of the sampling frequency in GK scattering

The last parameter we wish to assess the influence of, is the sampling frequency. So far, we fixed the time step to  $\Delta t = 10^{-3}$  s. We saw in section 5.2.2 that the ML estimator significantly outperforms the ergodicity estimation in the case of GK scattering. Since this is due to the information about  $\Psi_c$  contained in every single transition, it is natural to think that the more transitions, the better the performance. To test this idea, we carried numerical simulations and estimated  $\Psi_c$  for a range of sampling timesteps from  $10^{-5}$  s to  $10^{-1}$  s. For each timestep and value

of  $\mathcal{A}$ , 10000 trajectories are computed with  $\Psi_c = \sqrt{10}$  (only for GK scattering), and  $\Psi_c$  is estimated by ML. The estimation variance is then computed. The results are represented in figure 5. We observe as expected that the variance decreases as the sampling time step decreases (increase of the sampling frequency) most likely to an asymptotic lower bound. In GK scattering, there is information about the target in every transition due to the fact that  $x_t$  modulates the target constant  $\Psi_c$  in GK scattering. Therefore, as well as changing the timestep, we also change how fast  $x_t$  evolves, by tuning  $\mathcal{A}$ . Based on our simulation results, it seems that increasing the dynamics of  $x_t$ , *i.e.* increasing  $\mathcal{A}$ , leads to a decrease in the estimation variance.

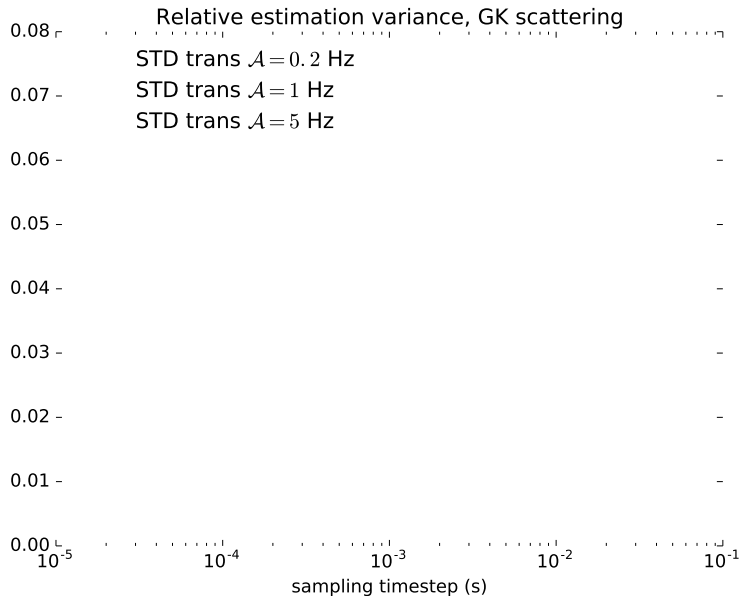


Figure 5. Relative estimation variance of the target in GK scattering for the ML estimators. 10000 trajectories are computed, with  $\mathcal{B} = 100$  Hz,  $\alpha = 1$  and  $\eta = 1$ . The variance is plotted as a function of the sampling time step.

## 6. Discussion

### 6.1. Detection

We propose a first approach to the detection problem (instead of estimation) considering the distribution of the estimated parameter  $\tilde{\Psi}_c$ . For simplicity, let us assume that the target constant is real-valued *i.e.*  $\Psi_c = \Psi_c^{(R)}$ . We denote  $p_{\Psi_c}(\tilde{\Psi}_c^{(R)} = x)$  and  $p_0(\tilde{\Psi}_c^{(R)} = x)$  the distributions of  $\tilde{\Psi}_c^{(R)}$  with and without a target  $\Psi_c$ , respectively.  $\tilde{\Psi}_c^{(R)}$  denotes any estimator of  $\Psi_c^{(R)}$ , for example  $\tilde{\Psi}_{c,ML}^{(R)}$  or  $\tilde{\Psi}_{c,e}^{(R)}$ . For target detection, one can evaluate the likelihood ratio:

$$\Lambda(\tilde{\Psi}_c^{(R)}) = \frac{p_{\Psi_c}(\tilde{\Psi}_c^{(R)} = x)}{p_0(\tilde{\Psi}_c^{(R)} = x)}. \quad (68)$$

Our method for estimating  $\Psi_c$  does not provide the distributions at the numerator and denominator for computing  $\Lambda$ . Instead, one can compute numerical distributions by estimating  $\Psi_c$  on many trajectories. Examples of numerical distributions obtained with 10000 trajectories are represented in figure 6 (left) for two target constants:  $\Psi_c = 1$  and  $\Psi_c = \sqrt{0.1}$ . In addition, we represent the Gaussian distributions with the numerical means and variances. To first order, the curves fit reasonably well. As might be anticipated, we observe that the separation between the no target and target distributions is much smaller for  $\Psi_c = \sqrt{0.1}$ .

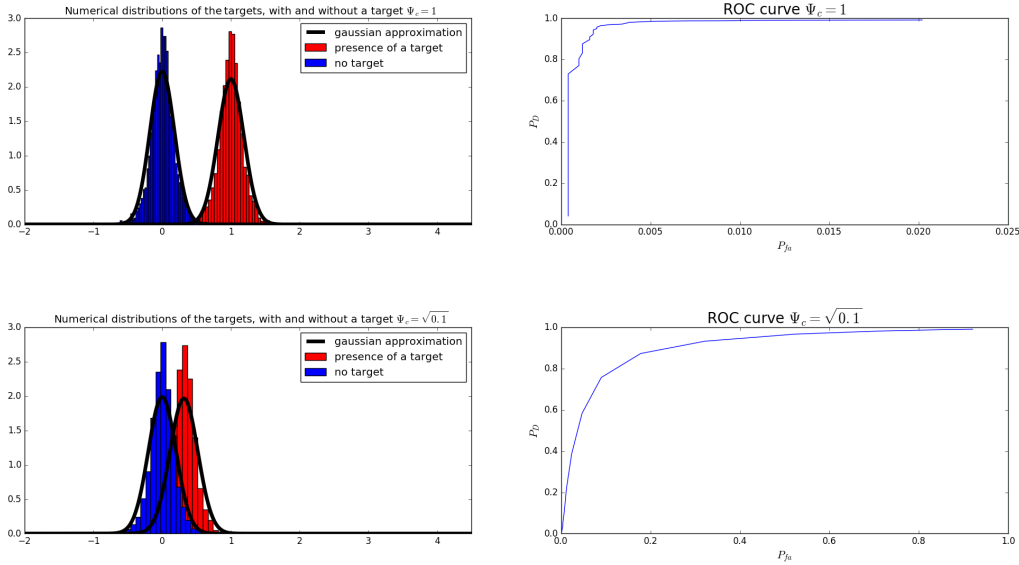


Figure 6. Numerical distributions of  $\tilde{\Psi}_c$  with and without a target, and numerical ROC curves based on the distribution of  $\tilde{\Psi}_c$ . 10000 trajectories of duration 1 s are computed, with  $\mathcal{A} = 1$  Hz,  $\mathcal{B} = 100$  Hz,  $\alpha = 1$ ,  $\eta = 1$ . Up:  $\Psi_c = 1$ ; Down:  $\Psi_c = \sqrt{0.1}$ .

If we denote  $\sigma_{\Psi_c}, \sigma_0$  and  $m_{\Psi_c}, m_0$  the numerical standard deviations and means with and without a target, we have:

$$\Lambda(\tilde{\Psi}_c^{(R)}) \approx \frac{\frac{1}{\sqrt{2\pi}\sigma_{\Psi_c}} e^{-(\tilde{\Psi}_c^{(R)} - m_{\Psi_c})/2\sigma_{\Psi_c}^2}}{\frac{1}{\sqrt{2\pi}\sigma_0} e^{-(\tilde{\Psi}_c^{(R)} - m_0)/2\sigma_0^2}}. \quad (69)$$

By deciding that there is a target if  $\Lambda$  exceeds the threshold  $\lambda_T$  and varying the threshold, we compute a range of probabilities of detection and probabilities of false alarms  $P_D$  and  $P_{FA}$  and get a ROC curve. More precisely, 10000 trajectories of duration 1 s are numerically computed with fixed  $\Psi_c = \Psi_c^{(R)} > 0$  (target). The  $i$ -th trajectory gives an estimated target  $\tilde{\Psi}_{c,i}^{(R)}$ .  $m_{\Psi_c}$  and  $\sigma_{\Psi_c}$  are then computed from the set  $\{\tilde{\Psi}_{c,i}^{(R)}, i = 1, 2, \dots, 10000\}$  of estimated  $\Psi_c^{(R)}$ . A similar procedure is carried to compute numerically  $m_0$  and  $\sigma_0$  (no target). Finally, to compute the probability of detection at fixed threshold, we count how many trajectories generated with a target  $\Psi_c = \Psi_c^{(R)} > 0$  are such that their associated  $\tilde{\Psi}_{c,i}^{(R)}$  verifies the condition  $\Lambda(\tilde{\Psi}_{c,i}^{(R)}) > \lambda_T$ . A similar procedure is done for the probability of false alarm.

The ROC curves with  $\Psi_c = 1$  and  $\Psi_c = \sqrt{0.1}$  are represented in figure 6 (right). Notice that the range covered by the  $x$ -axis depends on the target. It is manifest



that, as expected, the ROC curve is much better with  $\Psi_c = 1$  than with  $\Psi_c = \sqrt{0.1}$ .

The reader may object that the Gaussian distributions do not fit so well the numerical distributions. We wish to emphasize that we chose it just for illustrative purposes. More generally, the objective of this discussion is to show that the estimation of the target  $\Psi_c$  can lead to target detection.

## 6.2. Observability of the SCS

We assumed in section 4 that we observed three discrete time series,  $(\tilde{x}, \tilde{R}, \tilde{I}) = \{(\tilde{x}_k, \tilde{R}_k, \tilde{I}_k), k = 0, \dots, n\}$ , which are our available data for estimating the target constant  $\Psi_c = \Psi_c^{(R)} + i\Psi_c^{(I)}$ . It is visible in the expressions of the estimators, equation (47) for HK scattering and equation (61) for GK scattering, which are in the form:

$$\begin{cases} \tilde{\Psi}_c^{(R)} = F(\tilde{x}, \tilde{R}) \\ \tilde{\Psi}_c^{(I)} = F(\tilde{x}, \tilde{I}) \end{cases} \quad (70)$$

where  $F$  is some function. That we have access in reality to  $x_t$ , *i.e.* the SCS of the clutter, is not obvious. In reality, a sensor measures the total reflectivity  $\Psi_t^{(HK)}$  or  $\Psi_t^{(GK)}$ . The real and imaginary parts are then immediately derived, such that  $\tilde{R}$  and  $\tilde{I}$  are indeed observable. However, the SCS  $x_t$  is ‘encapsulated’ in the reflectivity as evidenced by equations (22) and (26). Fayard and Field provide a formula to optimally infer  $\tilde{x}$  from increments of  $z_t = |\Psi_t|^2$  and of the phase  $\theta_t$  when there is no target, *i.e.* clutter only [17]. However, their solutions requires the knowledge of the parameters  $\mathcal{A}$  and  $\alpha$ , which is also taken for granted in the present work. If we do not know  $\mathcal{A}$  and  $\alpha$ , and still with no target, we can approximate  $\tilde{x}$  by applying a sliding window to  $\tilde{z}$ , where  $\tilde{z}$  is the observed discrete time series of  $z_t$ . Indeed,  $z_t$  is K-distributed, and the K distribution arises from (see [4]):

$$p(z_t = z) = p(z_t = z | x_t = x)p(x_t = x) = \frac{1}{x} e^{-z/x} \times \frac{\alpha(\alpha x)^{\alpha-1} e^{-\alpha x}}{\Gamma(\alpha)}, \quad (71)$$

since

$$\begin{cases} p(z_t = z | x_t = x) = \frac{1}{x} e^{-z/x} \\ p(x_t = x) = \frac{\alpha(\alpha x)^{\alpha-1} e^{-\alpha x}}{\Gamma(\alpha)}. \end{cases} \quad (72)$$

The distribution of  $z_t$  for constant SCS  $x_t = x$  is then a negative exponential of expectation  $x$ . The SCS evolves on a much longer timescales than  $z_t = x_t |\gamma_t|^2$  (equation (13)). Consequently, averaging  $z_t$  over a time interval long compared to  $T_{speckle}$  but short compared to  $T_{SCS}$  gives an estimate of the SCS for this time interval.

It remains that the general case of estimating both  $\Psi_c$  and  $x_t$  has not been solved. As a first approach, one could estimate  $\Psi_c$  using the ergodic estimator, *i.e.* compute  $\hat{\Psi}_c$  from equation (64), deduce the approximate clutter by subtracting  $\hat{\Psi}_c$  to the observed reflectivity, estimate the SCS using Fayard and Field formula or a sliding window as described above, and then do ML estimation using equations (47) or (61). An even more challenging task would be to estimate, based only on the observed reflectivity (not the SCS), all the parameters of the model: the clutter parameters  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\alpha$  and the target  $\Psi_c = \Psi_c^{(R)} + i\Psi_c^{(I)}$ . Those are possible directions

for future work.

## 7. Conclusion

The isotropic (uniformly distributed phase over  $[0, 2\pi[$ ) random walk model can be used to model electromagnetic scattering by a random media, *i.e.* the clutter. Field's model for the clutter is an extension of the isotropic random walk model which accounts for the dynamics of the random medium, and in which the observables (reflectivity, intensity etc) are expressed in terms of stochastic differential equations. Coherent scattering (weak scattering or presence of a coherent scatterer) can be accommodated in this formalism by deriving SDE for the total scattered field. We derived these SDE for the in-phase and quadrature phase components of the reflectivity. Two models of coherent scattering have been considered: HK scattering and GK, which both depends on two parameters referred to as "target parameters". Approximate transition probabilities have been derived for small time steps using Euler-Maruyama scheme. We used these approximate transition probabilities for maximum likelihood estimation of the target parameters. In both HK and GK scattering, the estimators are given by explicit analytical formula.

Using numerical simulations, we have assessed the performance of the estimators. The estimation bias is negligible and the estimation variance is small compared to the parameters. We observed the intuitive result that the estimation variance decreases as the duration, or target intensity, or sampling frequency increases. We showed that mostly for GK scattering does the maximum likelihood estimation provide a significant advantage over the very simple ergodic estimator. Only in this case, do the transition probabilities really (*i.e.* not a simple translation) depend on the target parameters (see section 5.2.3). For HK scattering, the maximum likelihood estimator is better than the ergodic estimator in the case of small parameters. We discussed how our method of estimating the target leads to target detection, ROC curves etc.

We think that it would be valuable to study more precisely the performance of the estimation in terms of the quantity of information about the target parameters contained in each transition and the number of transitions available given the time series duration and the sampling frequency. This quantity of information would also depend on the parameters  $\mathcal{A}$ ,  $\mathcal{B}$  and  $\alpha$ , which control the dynamics of the clutter. Finally, HK and GK scattering are two examples of coherent scattering models. It is possible that based on physical considerations, one would choose a more adapted model, maybe a combination of HK and GK scattering or a model where the target is not modulated by  $x_t$  but by a different process. Our approach could then be extended to such models as long as the dynamics are expressed by stochastic differential equations.

## Acknowledgements

This work was supported by the DGA, Delegate for Armament of French MoD.

## References

- [1] E. Jakeman and K. D. Ridley, *Modeling Fluctuations in Scattered Waves*. CRC Press: Series in Optics and Optoelectronics, 2006.

- [2] E. Jakeman and P. Pusey, "A model for non-Rayleigh sea echo," *IEEE Transactions on Antennas and Propagation*, vol. 24, no. 6, pp. 806–814, 1976.
- [3] E. Jakeman, "On the statistics of K-distributed noise," *Journal of Physics A: Mathematical and General*, vol. 13, no. 1, pp. 31–48, 1980.
- [4] K. D. Ward, R. J. A. Tough, and S. Watts, *Sea Clutter: Scattering, the K distribution and Radar Performance*. 20, IET Radar, Sonar and Navigation, first edition ed., 2006.
- [5] T. R. Field, *Electromagnetic Scattering from Random Media*. Oxford University Press, 2009.
- [6] R. Barakat, "Weak-scatterer generalization of the K-density function with application to laser scattering in atmospheric turbulence," *J. Opt. Soc. Am. A*, vol. 3, pp. 401–409, Apr 1986.
- [7] E. Jakeman and R. J. A. Tough, "Generalized K distribution: a statistical model for weak scattering," *Journal of the Optical Society of America A*, vol. 4, Sept. 1987.
- [8] E. Jakeman and R. J. A. Tough, "Non-gaussian models for the statistics of scattered waves," *Advances in Physics*, vol. 37, no. 5, pp. 471–529, 1988.
- [9] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications, Fifth Edition*. Springer-Verlag, 2000.
- [10] C. J. Roussel, A. Coatanhay, and A. Baussard, "Forward and backward probabilistic inference of the sea clutter," *Waves in Random and Complex Media*, vol. 29, no. 3, pp. 540–568, 2018.
- [11] C. J. Roussel, A. Coatanhay, and A. Baussard, "Estimation of the parameters of stochastic differential equations for sea clutter," *IET Radar, Sonar and Navigation*, vol. 13, no. 4, pp. 497 – 504, 2018.
- [12] D. J. Higham, "An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations," *Society for Industrial and Applied Mathematics*, vol. 43, no. 3, pp. 525–546, 2001.
- [13] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Springer, 1992.
- [14] J. H. Matis and T. R. Kiffe, *Stochastic Population Models: A Compartmental Perspective*. Springer, 2000.
- [15] H. Risken, *The Fokker-Planck Equation: methods of solution and applications*. Springer, 1989.
- [16] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data*. Artech House, 2005.
- [17] P. Fayard and T. R. Field, "Optimal inference of the scattering cross-section through the phase decoherence," *Waves in Random and Complex Media*, vol. 18, no. 4, pp. 571–584, 2008.