



**HAL**  
open science

## Monte Carlo methods for radiative transfer with singular kernels

Christophe Gomez, Olivier Pinaud

► **To cite this version:**

Christophe Gomez, Olivier Pinaud. Monte Carlo methods for radiative transfer with singular kernels. SIAM Journal on Computing, 2018, 40 (3), pp.A1714-A1741. 10.1137/17M1134755 . hal-01528667

**HAL Id: hal-01528667**

**<https://hal.science/hal-01528667>**

Submitted on 29 May 2017

**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.

# Monte Carlo methods for radiative transfer with singular kernels

Christophe Gomez, <sup>\*,1</sup> and Olivier Pinaud <sup>†2</sup>

<sup>1</sup>Aix Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373,  
13453 Marseille, France

<sup>2</sup>Department of Mathematics, Colorado State University, Fort Collins  
CO, 80523

May 15, 2017

## Abstract

This work is devoted to Monte Carlo methods for radiative transfer equations with singular kernels, and is motivated by the study of wave propagation in random media with long-range dependence. As opposed to the short-range case where the collision cross section is integrable and leads to a non-zero mean free time, the cross section is not integrable in the long-range situation and yields a vanishing mean free time. For computational efficiency, a particular care is then required in the construction of the stochastic processes used in the Monte Carlo methods. For this, we adapt a method of Asmussen-Rosiński and Cohen-Rosiński based on a small jumps/large jumps decomposition of the generator. We compare this method with another approach based on alpha-stable processes, and show the superiority of the first one. We consider various algorithms for the simulation of the jump distribution, and underline the efficiency of an appropriate stochastic collocation method. Comparisons between integrable and singular kernel solutions are given.

## 1 Introduction

This work is devoted to the derivation of Monte Carlo methods for the resolution of two-dimensional radiative transfer equations (RTE) of the form

$$\partial_t f + k \cdot \nabla_x f = \mathcal{Q}(f), \quad f(t=0) = f_0, \quad (x, k) \in \mathbb{R}^2 \times \mathbb{S}^1, \quad (1)$$

with collision operator

$$\mathcal{Q}(f)(k) = \int_{\mathbb{S}^1} F(k \cdot p)(f(p) - f(k))d\sigma(p), \quad F \geq 0.$$

---

\*christophe.gomez@univ-amu.fr

†pinaud@math.colostate.edu

Above,  $\sigma(p)$  is the surface measure on  $\mathbb{S}^1$ . Our motivation comes from the study of high frequency wave propagation in random media. In such a context, the function  $f$  solution to the RTE describes the wave energy density in the phase space, see [12, 19, 28, 31] for a few references. The collision term models the interaction between the wave and the fluctuations of the underlying medium in the so-called weak coupling regime. For instance, if  $c(x)$  is the velocity field and reads

$$\frac{1}{c(x)^2} = \frac{1}{c_0^2} (1 + V(x)),$$

where  $c_0$  is a constant background velocity and  $V$  is a mean zero, stationary random field that models fluctuations around the background, then the collision kernel  $F$  is proportional to the Fourier transform of the correlation function

$$R(x - y) = \mathbb{E}\{V(x)V(y)\}.$$

The symbol  $\mathbb{E}$  means expectation over realizations of the random medium, and the kernel  $F$  is nonnegative by application of Bochner theorem.

We are interested here in situations where the random media presents long-range correlations, in the sense that the correlation function  $R$  decays so slow at the infinity that it is not integrable. Such a scenario arises for instance in turbulent atmosphere, in the earth crust, or in biomedical applications [7, 18, 32]. We display in Figure 1 typical realizations of media with short-range (SR) and long-range correlations (LR). The structure is very different, and this can be seen in the RTE itself. When the medium is SR, the kernel  $F$  is integrable (with integral  $\mu$ ), and the collision operator can be split into gain and loss terms. In terms of stochastic processes, which are at the core of Monte Carlo methods,  $\mathcal{Q}$  is the generator of a compound Poisson process with intensity  $\mu$ , whose simulation is now standard material [16, 21, 33]. Besides, the solution  $f$  has essentially the regularity of the initial condition  $f_0$  (we do not take into account here some well known averaging effects of the transport operator for integrals of  $f$  in the variable  $k$  [10]).

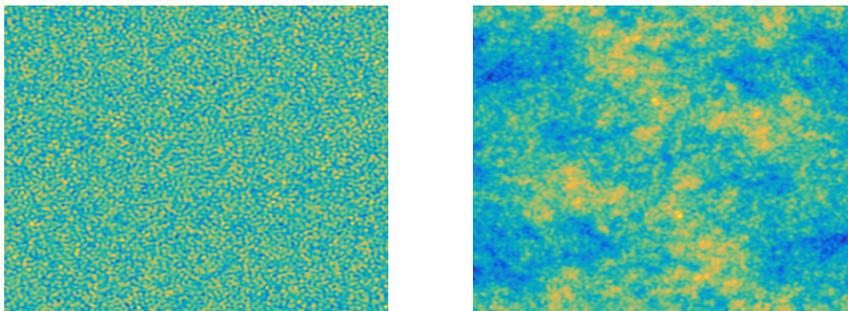


Figure 1: Random media with short-range correlations (left), and long-range correlations (right).

The situation is very different in the LR case. First of all, due to the slow decay of  $R$  at the infinity, the kernel  $F$  has a non integrable singularity at zero. As a consequence, one can only make sense of the term  $\mathcal{Q}$  by exploiting the difference  $f(p) - f(k)$ , which precludes any splitting between gain and losses. The intensity  $\mu$  is infinite, and the

operator  $\mathcal{Q}$  is now the generator of a general Lévy process on  $\mathbb{S}^1$ , and not just of a pure jump Poisson process as in the SR case. Informally,  $\mu^{-1}$  measures the average time between collisions (the “Mean Free Time”), which means in the LR case that collisions are taking place in a continuous basis. The simulation of such Lévy processes in the context of transport equations seems much less studied than its compound Poisson counterpart, and is then the object of this work. We focus here on the two-dimensional case with velocities on  $\mathbb{S}^1$ , and will address the three-dimensional case elsewhere. Note that the choice of  $\mathbb{S}^1$  is arbitrary, any circle of given radius would be dealt with in the same fashion. The singularity of kernel tends to favor grazing collisions, and as a consequence this LR regime bears some strong similarities with the classical peaked-forward regime [20, 23, 25], with the forward effect magnified by the singularity.

Regarding the regularity of  $f$ , it was established in [13, 14] that  $f$  is  $C^\infty$  in all variables for all  $t > 0$ , which is in stark contrast with the SR case. The regularity is reminiscent of the hypoelliptic nature of (1) when  $F$  is not integrable. Our original motivation for studying (1) is the resolution of some inverse problems, for instance the localization of sources or inclusions in clutter. Since inversion techniques based on transport equations are often based on the singularities of  $f$ , see [3], the smoothness of  $f$  indicates that the inversion, if possible at all, should be more difficult than in the SR situation.

Monte Carlo (MC) methods have several advantages compared to classical finite elements or finite volumes methods. They can handle the high-dimensionality of (1) in a simple manner, since no mesh is required and the quantity of interest only need be discretized at a detector where measurements take place. They are very flexible in terms of geometry, provided the boundary conditions on  $f$  can be translated into boundary conditions on the underlying stochastic process, and are generally applicable when the coefficients in the equation depend on the variables  $(x, k)$ . MC methods are also very easy to implement and can be parallelized in a straightforward manner, which is a strong feature with today’s technology. Note that discretization-based methods would have to handle carefully the singularity of the cross-section.

On the down side, while MC methods offer a relatively good accuracy at low cost, getting more accurate results can come at a significant increase in the computational time. Some variance reduction techniques might then be necessary to lower this cost.

Typical collisions kernels in turbulent atmosphere are the celebrated Kolmogorov (K) power spectrum [27], or its Von Karman (VK) variant including the inner and outer scales of the turbulence [26]. With  $F(k, p) = \Phi(|k - p|)$ , they read

$$\Phi(|k|) = C_0 |k|^{-\frac{11}{3}} \quad (\text{K}), \quad \Phi(|k|) = C_1 (1 + L_0^2 |k|^2)^{-\frac{11}{6}} e^{-\ell_0^2 |k|^2}, \quad (\text{VK}),$$

where  $C_0$  and  $C_1$  are constants, and  $L_0, \ell_0$  are the outer and inner scales, respectively. In a turbulent regime, we have  $L_0 \gg \ell_0$ , so that the Von Karman power spectrum behaves like the Kolmogorov power spectrum at  $k = 0$ , and is therefore not integrable in dimensions two and three. In this work, we will mostly focus on kernels of the form  $F(k \cdot p)$ , and explain how our methods can be directly adapted to kernels like  $F(x, k, p)$ . We consider kernels with power singularities such as the Kolmogorov power spectrum, that read

$$F(k \cdot p) = \frac{A(k \cdot p)}{|k - p|^{1+\alpha}}, \quad \alpha \in (0, 2), \quad 0 < \underline{a} \leq A(s) \leq \bar{a} < \infty, \quad (2)$$

for  $s \in [-1, 1]$ .

With the hypotheses above, the operator  $\mathcal{Q}$  can be seen the generator of a Lévy process  $(K^k(t))_{t \geq 0}$  on  $\mathbb{S}^1$ , with initial condition  $K^k(0) = k$ , see [13]. Introducing the position

$$X^x(t) = x + \int_0^t K^k(s) ds,$$

the solution  $f$  to (1) is then

$$f(t, x, k) = \mathbb{E}\{f_0(X^x(t), K^k(t))\}.$$

Generating  $N$  independent copies  $(X_i^x, K_i^k)_{i=1, \dots, N}$  of the random trajectories  $(X^x(t), K^k(t))_{t \geq 0}$ ,  $f$  is approximated by the empirical average

$$f_N(t, x, k) = \frac{1}{N} \sum_{i=1}^N f_0(X_i^x(t), K_i^k(t)),$$

while the central limit theorem shows that the error is controlled by some measure of the standard deviation of  $f_N$  divided by  $\sqrt{N}$ .

The key to MC methods is an efficient simulation of the process  $(K(t))_{t \geq 0}$ . To this end, we adapt a method of the probabilistic literature developed by Asmussen-Rosiński [1], and Cohen-Rosiński [6], that we will refer to as the ACR method. The idea is to introduce a cut-off in the collision operator to remove the singularity. The resulting operator is then the generator of a compound Poisson process with finite (but large) intensity. The issue is that a small value of the cut-off parameter is required for good accuracy, leading to a high intensity and therefore to many collisions increasing the computational time. The issue is fixed by adding a correcting term describing asymptotically the behavior of the operator around the singularity. This term models frequent jumps of small amplitude, while the regularized generator describes more rare jumps with larger amplitudes. In the work of Asmussen-Rosiński, and Cohen-Rosiński, the correction takes the form of a Brownian motion on the real line, while in our case we obtain a Brownian motion on the circle (more generally we would obtain a Brownian motion on  $\mathbb{S}^d$  with generator given by the Laplace-Beltrami operator on  $\mathbb{S}^d$ ).

Without another MC method to compare against, the efficiency of the ACR method for the simulation of transport equations is not obvious at first since it is only an approximation and the cost of the generation of a compound Poisson process with large intensities is to be determined. To this goal, we carefully design an alternative MC method that does not involve a regularization of the generator, and is based on the so-called alpha-stable processes. We will refer to this second method as the AS method. We also derive an efficient way to simulate the jump distribution of the compound process that accounts for the singularities of the kernel, and is based on the stochastic collocation method. We show that the ACR method is very efficient and, perhaps surprisingly, that it is superior to the AS method for both weakly singular kernels (i.e.  $\alpha$  close to 0) and highly singular kernel ( $\alpha$  close to 2).

The paper is organized as follows. In Section 2, we present the ACR method and give an error analysis. We introduce the AS method in Section 3 and Section 4 is devoted to algorithms for the jump distributions. We describe rejection sampling methods and the stochastic collocation method. We perform simulations in section 5, and start

by comparing the latter methods, and show the superiority of the last one. We then address the impact of the Brownian correction in the ACR method, and compare the performances of ACR and AS methods. We also investigate the sensitivity of the ACR method to the cut-off parameter. We conclude the simulations section by comparing the solutions to the RTE for underlying random media with SR and LR correlations, and by investigating the role of the function  $a$ . We finally present generalizations to kernels of the form  $F(x, k, p)$  in Section 6, and a conclusion is offered in Section 7.

**Acknowledgment.** OP is supported by NSF CAREER grant DMS-1452349.

## 2 The ACR method

We adapt in this section the method introduced by Asmussen-Rosiński in [1] and Cohen-Rosiński in [6]. For this, we need first to appropriately parametrize the operator  $\mathcal{Q}$ . The natural choice for  $\mathcal{Q}$  are the polar coordinates and we introduce the surjective map

$$P : \phi \in \mathbb{R} \longmapsto (\cos(\phi), \sin(\phi)) \in \mathbb{S}^1.$$

Setting  $k = (\cos(\theta), \sin(\theta))$  and  $p = (\cos(\theta'), \sin(\theta'))$  so that  $k \cdot p = \cos(\theta' - \theta)$ , and rewriting

$$A(k \cdot p) = a(\theta' - \theta),$$

the collision kernel  $\mathcal{Q}$  can be recast as

$$\begin{aligned} \mathcal{Q}(f)(k) &= \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{-\pi}^{\pi} \frac{a(\theta' - \theta)(\tilde{f}(\theta') - \tilde{f}(\theta))}{(1 - \cos(\theta' - \theta))^{\frac{1+\alpha}{2}}} d\theta' \\ &= \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{-\pi}^{\pi} \frac{a(\theta')(\tilde{f}(\theta' + \theta) - \tilde{f}(\theta))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' \\ &= \tilde{\mathcal{Q}}(\tilde{f})(\theta), \quad \theta \in [-\pi, \pi], \end{aligned}$$

where  $\tilde{f} = f \circ P$  is a  $2\pi$ -periodic function. Note that the  $2\pi$ -periodic function  $a$  is symmetric with respect to 0 and is nonnegative. The operator  $\tilde{\mathcal{Q}}$  can be seen as the generator of a Lévy process on  $[-\pi, \pi]$  starting from 0, that we denote by  $(\Theta(t))_{t \geq 0}$ . According to our parametrization,

$$(K^k(t))_{t \geq 0} = (P(\Theta(t) + \theta))_{t \geq 0} \quad \text{in law.}$$

For  $\varepsilon \in (0, 1)$ , the ACR method consists in splitting  $\tilde{\mathcal{Q}}$  into

$$\begin{aligned} \tilde{\mathcal{Q}}(g)(\theta) &= \mathcal{Q}_{<}^{\varepsilon}(g)(\theta) + \mathcal{Q}_{>}^{\varepsilon}(g)(\theta) \\ &:= \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{I_{\varepsilon}^<} \frac{a(\theta')(g(\theta' + \theta) - g(\theta))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' + \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{I_{\varepsilon}^>} \frac{a(\theta')(g(\theta' + \theta) - g(\theta))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' \end{aligned}$$

where  $I_{\varepsilon}$  is the set

$$I_{\varepsilon} := \{\theta' \in [-\pi, \pi], \quad |\tan(\theta'/4)| > \varepsilon/4\},$$

and  $I_\varepsilon^C$  its complementary in  $[-\pi, \pi]$ . The reason for the latter form of the cut-off will be apparent further. Because of the regularization, the second piece of the generator, namely  $\mathcal{Q}_>^\varepsilon(g)$ , is now the generator of a classical compound Poisson process with jump intensity

$$\mu_\varepsilon = \frac{\Pi_\varepsilon^0}{2^{\frac{1+\alpha}{2}}}, \quad \text{where} \quad \Pi_\varepsilon^0 = \int_{I_\varepsilon} \frac{a(\theta)d\theta}{(1 - \cos(\theta))^{\frac{1+\alpha}{2}}},$$

and jump distribution

$$\Pi_\varepsilon(d\theta) = \frac{\mathbb{1}_{I_\varepsilon}(\theta)a(\theta)}{\Pi_\varepsilon^0(1 - \cos(\theta))^{\frac{1+\alpha}{2}}}d\theta.$$

The expected number of jumps for the compound Poisson process in any time interval  $[0, T]$  is given by  $\mu_\varepsilon T$ . The term  $\mu_\varepsilon$  blows up as  $\varepsilon \rightarrow 0$ , we will give a precise estimate further, and this fact is the reason for the need of an efficient simulation technique for the jump distribution.

The first term of the generator, i.e.  $\mathcal{Q}_<^\varepsilon$ , models frequent jumps of small amplitude and can be asymptotically reduced to the generator of a Brownian motion on the circle as explained in the next section.

## 2.1 The generator $\mathcal{Q}_<^\varepsilon$

We have the following lemma, proved in Appendix A.1, which describes the relation between  $\mathcal{Q}_<^\varepsilon$  and the Laplace operator.

**Lemma 2.1** *Suppose that  $a \in C^2([-\pi, \pi])$  and that  $g$  is a smooth  $2\pi$ -periodic bounded function with bounded derivatives. Then, for any  $\phi \in [-\pi, \pi]$ ,*

$$\mathcal{Q}_<^\varepsilon(g)(\phi) = \frac{D_\varepsilon}{2} \frac{\partial^2 g(\phi)}{\partial \phi^2} + \mathcal{R}_\varepsilon[g](\phi),$$

where

$$D_\varepsilon = \frac{2a(0)\varepsilon^{2-\alpha}}{2-\alpha} \quad \text{and} \quad \|\mathcal{R}_\varepsilon[g]\|_\infty \leq \frac{\varepsilon^{4-\alpha}}{4-\alpha} \|a\|_{C^2([-\pi, \pi])} \|g\|_{C^4([-\pi, \pi])}.$$

The leading term in  $\mathcal{Q}_<^\varepsilon$  is the generator of a Lévy process with density

$$p_B^\varepsilon(t, \phi) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{\ell=1}^{\infty} e^{-tD_\varepsilon \ell^2/2} \cos \ell\phi, \quad t > 0.$$

It can be directly simulated by remarking that, using the Poisson summation formula,

$$p_B^\varepsilon(t, \phi) = \frac{1}{\sqrt{2\pi t D_\varepsilon}} \sum_{\ell \in \mathbb{N}} e^{-\frac{(2\pi\ell + \phi)^2}{2tD_\varepsilon}}, \quad t > 0,$$

which, if  $\phi_t$  is a centered Brownian motion with variance  $tD_\varepsilon$ , is the density of  $(\phi_t \bmod 2\pi)$ . Brownian trajectories starting from 0 can easily be generated using the following relation :

$$\begin{aligned} \phi_{nh}^\varepsilon &= \phi_{nh}^\varepsilon - \phi_{(n-1)h}^\varepsilon + \cdots + \phi_{2h}^\varepsilon - \phi_h^\varepsilon + \phi_h^\varepsilon \\ &:= X_n + \cdots + X_2 + X_1, \end{aligned}$$

where  $(X_i)_{i=1,\dots,n}$  are i.i.d centered Gaussian random variable with variance  $hD_\varepsilon$ , that we denote by  $\mathcal{N}(0, hD_\varepsilon)$ . Note that  $D_\varepsilon$  is proportional to  $\varepsilon^{2-\alpha}/(2-\alpha)$ , which becomes larger as  $\alpha$  increases. We therefore expect the inclusion of the Brownian term in the simulation to be critical in order to obtain good accuracy and computational cost.

## 2.2 The generator $\mathcal{Q}_>^\varepsilon$

As already mentioned, this operator is the generator of a compound Poisson process with Lévy measure  $\mu_\varepsilon \Pi_\varepsilon$ . Denoting by  $(\theta_\varepsilon(t))_{t \geq 0}$  such a process starting at 0, a sample path has the form

$$\theta_\varepsilon(t) = \sum_{i=1}^{N_t} \theta_i, \quad (3)$$

where  $(N_t)_{t \geq 0}$  is a Poisson process with intensity  $\mu_\varepsilon$ , independent of  $(\theta_i)_{i \geq 1}$ , which is a sequence of i.i.d random variables distributed according to  $\Pi_\varepsilon(d\theta)$ . The sojourn times in each state  $i$  are independent and satisfy an exponential distribution with parameter  $\mu_\varepsilon$ .

In order to simulate the process, we use the following parametrization. Since the jump distribution  $\Pi_\varepsilon(d\theta)$  is symmetric around the origin, we can write  $\theta = \text{sign}(\theta)|\theta| := s_\theta \theta'$ , where  $s_\theta$  takes the values  $\pm 1$  with equal probability, and  $\theta' \in I_{\varepsilon,+} := I_\varepsilon \cap [0, \pi]$ . We then rewrite  $\Pi_\varepsilon$  as

$$\Pi_\varepsilon(d\theta) = \pi_\varepsilon(\theta') d\theta' d\mu(s_\theta),$$

with

$$d\mu(s_\theta) = \frac{1}{2} (\delta(s_\theta + 1) + \delta(s_\theta - 1)), \quad \pi_\varepsilon(\theta') = \frac{\mathbb{1}_{I_{\varepsilon,+}}(\theta') a(\theta')}{\Pi_{\varepsilon,+}^0(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}},$$

and

$$\Pi_{\varepsilon,+}^0 = \int_{I_{\varepsilon,+}} \frac{a(\theta) d\theta}{(1 - \cos(\theta))^{\frac{1+\alpha}{2}}}. \quad (4)$$

Therefore, in order to simulate a jump  $\theta$  under  $\Pi_\varepsilon(d\theta)$ , we generate  $\theta'$  according to  $\pi_\varepsilon(\theta') d\theta'$ , and multiply it by  $s_\theta = \pm 1$  with equal probability.

Note that the rate  $\mu_\varepsilon$  satisfies the asymptotics

$$\mu_\varepsilon \underset{\varepsilon \rightarrow 0}{\simeq} \frac{a(0)}{2\alpha\varepsilon^\alpha}, \quad (5)$$

and therefore blows up as expected in the limit  $\varepsilon \rightarrow 0$ . This expresses the fact that the process jumps more and more as  $\varepsilon$  becomes small, which hence increases the simulation cost. Note also that the rate blows up more rapidly when  $\alpha$  is large. Several methods to simulate  $\pi_\varepsilon$  will be presented in Section 4.

## 2.3 Error analysis

The main idea of the ACR method is that the Lévy process  $(K^k(t))_{t \geq 0}$  on  $\mathbb{S}^1$  starting at  $k$  with generator  $\mathcal{Q}$  is approximated by the process

$$(K^{k,\varepsilon}(t))_{t \geq 0} = ((\cos(\phi_\varepsilon(t) + \theta_\varepsilon(t) + \theta), \sin(\phi_\varepsilon(t) + \theta_\varepsilon(t) + \theta))_{t \geq 0},$$

where  $\phi_\varepsilon$  and  $\theta_\varepsilon$ , defined in the last two sections, are independent. Note that we replaced  $\phi_\varepsilon \bmod 2\pi$  by  $\phi_\varepsilon$  by periodicity. The process  $(\phi_\varepsilon(t))_{t \geq 0}$  is generated as described in Section 2.1, while  $(\theta_\varepsilon(t))_{t \geq 0}$  has the form (3).

With

$$X^{x,\varepsilon}(t) = x + \int_0^t K^{k,\varepsilon}(s) ds,$$

the function

$$f^\varepsilon(t, x, k) = \mathbb{E}\{f_0(X^{x,\varepsilon}(t), K^{k,\varepsilon}(t))\}$$

is then the solution to the transport equation below, with the parametrization  $k = (\cos(\theta), \sin(\theta))$ ,

$$\partial_t f^\varepsilon + k \cdot \nabla_x f^\varepsilon = \mathcal{Q}_>^\varepsilon(f^\varepsilon) + \frac{D_\varepsilon}{2} \partial_\theta^2 f^\varepsilon, \quad f^\varepsilon(t=0) = f_0,$$

and it is direct to show that the solution is unique. We have then the following proposition, proved in Appendix A.2.

**Proposition 2.2** *Suppose that the unique solution to the transport equation (1) satisfies  $f \in L^1((0, T), L^2(\mathbb{R}^2, C^4(\mathbb{S}^1)))$ , and that  $a \in C^2([-\pi, \pi])$ . Then, for all  $t \in (0, T)$ ,*

$$\|f^\varepsilon(t, \cdot, \cdot) - f(t, \cdot, \cdot)\|_{L^2(\mathbb{R}^2 \times \mathbb{S}^1)} \leq \frac{\varepsilon^{4-\alpha}}{4-\alpha} \|a\|_{C^2([-\pi, \pi])} \int_0^t \|f(s)\|_{L^2(\mathbb{R}^2, C^4(\mathbb{S}^1))} ds.$$

Note that the hypothesis (2) on  $a$  can be relaxed to hold only in a neighborhood of  $\theta = 0$ . The smoothness assumption on  $f$  is satisfied for instance if it holds for  $f_0$ , as equation (1) propagates regularity. As was mentioned before, solutions to (1) are automatically  $C^\infty$  in all variables for all  $t \geq t_0 > 0$  if  $f_0 \in L^2(\mathbb{R}^2 \times \mathbb{S}^1)$ . Proposition 2.2 could then be restated in a more optimal way by taking into account this regularity, and this would require an additional analysis in an initial layer  $(0, t_0)$  since regularity occurs only for  $t \geq t_0$ .

The next section is devoted to the derivation of an alternative approach to the ACR method. Our main motivation in doing so is to provide an efficient method to compare the ACR method against.

### 3 The AS method

The main idea is to directly simulate the process  $(K^k(t))_{t \geq 0}$  without using the small jumps/larger jumps approximation of the generator. The method is applicable provided we make the additional assumption that  $a(\theta)$  admits a global minimizer at  $\theta = 0$ , or equivalently  $A(s)$  admits a global minimizer at  $s = 1$ . This allows us to decompose  $\mathcal{Q}$  as

$$\begin{aligned} \mathcal{Q}(f)(k) &= \frac{A(1)}{2^{\frac{1+\alpha}{2}}} \int_{\mathbb{S}^1} \frac{f(p) - f(k)}{(1 - \cos(k \cdot p))^{\frac{1+\alpha}{2}}} d\sigma(p) \\ &\quad + \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{\mathbb{S}^1} \frac{A(k \cdot p) - A(1)}{(1 - \cos(k \cdot p))^{\frac{1+\alpha}{2}}} (f(p) - f(k)) d\sigma(p) \\ &:= \mathcal{Q}^1(f)(k) + \mathcal{Q}^2(f)(k) \end{aligned}$$

The second term  $\mathcal{Q}^2$  is the generator of a compound Poisson process on  $\mathbb{S}^1$  with finite intensity, which does not pose any computational issue since now the average length of the jumps is large compared to those generated by  $\mathcal{Q}_{>}^\varepsilon$ . The main difficulty is then in the first term  $\mathcal{Q}^1$ , for which we need to construct an appropriate random Markov process  $(K^1(t))_{t \geq 0}$ . The first step is to compute the probability density function,

$$p(t, k, k_0) = \mathbb{P}(K^1(t) = k \mid K^1(0) = k_0)$$

that satisfies the forward Kolmogorov equation

$$\partial_t p(k) = \mathcal{Q}^1(p)(k), \quad \text{with} \quad p(t = 0, k, k_0) = \delta(k - k_0), \quad k, k_0 \in \mathbb{S}^1,$$

and where  $\delta$  is the Dirac measure. With the parametrization  $k \cdot k_0 = \cos \theta$ , the operator  $\mathcal{Q}$  can be diagonalized using Fourier series, and we obtain the following exact expression for  $p$  (written as a function of  $\theta$ ),

$$p(t, \theta) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{\ell=1}^{\infty} e^{t\lambda_\ell} \cos \ell\theta, \quad (6)$$

where (below  $a_0 = a(0) = A(1)$ )

$$\lambda_\ell = \frac{a_0 \pi^{\frac{1}{2}} \Gamma(-\frac{\alpha}{2})}{2^\alpha \Gamma(\frac{1+\alpha}{2})} \left( \frac{\Gamma(\ell + \frac{1+\alpha}{2})}{\Gamma(\ell + \frac{1-\alpha}{2})} - \frac{\Gamma(\frac{1+\alpha}{2})}{\Gamma(\frac{1-\alpha}{2})} \right), \quad (7)$$

see [29] with a slight adjustment of the constant prefactors, with  $\Gamma$  the usual gamma function. Note that the eigenvalues  $\lambda_\ell$  are negative, and we then recast them as follows,

$$\lambda_\ell = -D_\alpha(\beta_\ell - \beta_0),$$

with

$$D_\alpha = -\frac{a_0 \pi^{\frac{1}{2}} \Gamma(-\frac{\alpha}{2})}{2^\alpha \Gamma(\frac{1+\alpha}{2})}, \quad \beta_\ell = \frac{\Gamma(\ell + \frac{1+\alpha}{2})}{\Gamma(\ell + \frac{1-\alpha}{2})},$$

where  $D_\alpha$  is positive. Before going further, we need to introduce the so-called alpha-stable processes, that are central in the simulation of  $(K^1(t))_{t \geq 0}$ .

### 3.1 Alpha-stable processes

We use here the notation of [30]. A random variable  $X$  has a stable distribution if there are real parameters  $\alpha \in (0, 2]$ ,  $\beta \in [-1, 1]$ ,  $\sigma \geq 0$ , and  $\mu$  such that

$$\mathbb{E}\{\exp itX\} = \begin{cases} \exp \{-\sigma^\alpha |t|^\alpha (1 - i\beta(\text{sign}(t) \tan(\pi\alpha/2))) + i\mu t\} & \text{if } \alpha \neq 1, \\ \exp \{-\sigma |t| (1 + 2i\beta(\text{sign}(t) \log |t|/\pi) + i\mu t)\} & \text{if } \alpha = 1, \end{cases}$$

The law of  $X$  is denoted by  $S_\alpha(\sigma, \beta, \mu)$ . We will only be concerned with symmetric stable processes for which  $\beta = \mu = 0$ , and write the corresponding law as  $S_\alpha(\sigma) = S_\alpha(\sigma, 0, 0)$ . The density of such a symmetric alpha-stable random variable is then

$$\varphi_{\alpha, \sigma}(\xi) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\sigma^\alpha |x|^\alpha} e^{-i\xi x} dx.$$

Stable random variables can be simulated with the algorithm described in [35].

We introduce in the next paragraph an approximated process that is essentially an alpha-stable process on  $\mathbb{S}^1$ .

### 3.2 An approximate process

The process is constructed following the key observation that  $\beta_\ell$  converges very fast to  $\ell^\alpha$  as  $\ell \rightarrow \infty$ , see Figure 2. We then replace  $\beta_\ell$  by  $\ell^\alpha$  in the definition of  $p(t, \theta)$ , as well as  $\beta_0$  by 0 with the rationale that  $|\beta_0| \ll \beta_\ell$  when  $\ell \gg 1$ . We therefore consider the density

$$p_0(t, \theta) := \frac{1}{2\pi} + \frac{1}{\pi} \sum_{\ell=1}^{\infty} e^{-tD_\alpha \ell^\alpha} \cos \ell\theta, \quad \theta \in [-\pi, \pi].$$

By the Poisson summation formula,  $p_0$  can be written as,

$$p_0(t, \theta) = \sum_{\ell \in \mathbb{Z}} \varphi_{\alpha, (tD_\alpha)^{1/\alpha}}(2\pi\ell + \theta), \quad \theta \in [-\pi, \pi],$$

where  $\varphi_{\alpha, (tD_\alpha)^{1/\alpha}}$  is the density of a symmetric alpha-stable random variable introduced before. The above expression shows that if  $X \sim S_\alpha((tD_\alpha)^{1/\alpha})$ , then  $p_0$  is the density of  $X \bmod 2\pi$ . Periodized densities of the form of  $p_0$  are referred to as "wrapped distribution".

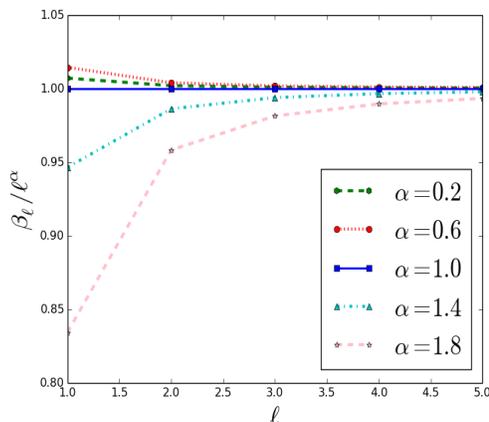


Figure 2: The ratio  $\beta_\ell/\ell^\alpha$  as a function of  $\ell$  for various values of  $\alpha$ .

### 3.3 Correcting the approximate process: stochastic collocation

The density  $p_0$  is an excellent approximation of  $p$ , as is shown on figure 3. For the two cases in the figure, the  $\ell_2$  relative error is about 0.003. In the case  $\alpha = 1.6$ , the error is essentially around the origin, where most of the values of  $\theta$  will be drawn, and is about half a percent of the peak value. The error for larger values of  $\theta$  is smaller by roughly a factor 10.

Even though these small errors translate into small errors in the transport solution, it is not possible to obtain an arbitrary precision by using  $p_0$  instead of  $p$ . For arbitrary accuracy, a random variable drawn according to  $p_0$  needs to be corrected to satisfy  $p$ . A straightforward way to do this would be by a rejection method. While the acceptance rate is expected to be high since  $p_0$  and  $p$  are close, the rejection requires many calculations of the densities that could be costly. A better approach is to use the stochastic

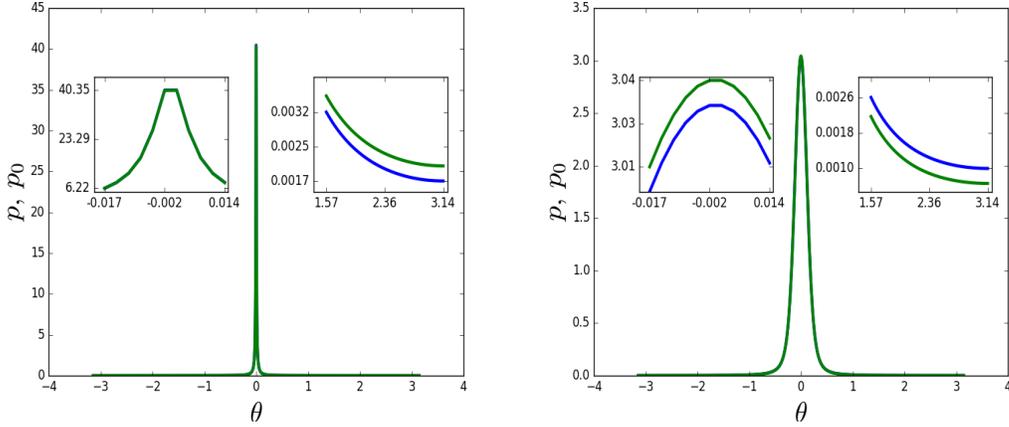


Figure 3: Comparison between  $p$  (blue) and  $p_0$  (green) for  $\alpha = 0.8$  (left) and  $\alpha = 1.6$  (right). Above,  $a(x) = 0.1$  and  $t = h = 0.03$ , which is a typical value for the time stepsize in the resolution of the radiative transfer equation

collocation method, see [15]: if  $X$  has density  $p_0$ , and if its cumulative distribution function (CDF) is denoted by  $G$ , then  $G(X)$  is a uniform random variable in  $[0, 1]$ . Besides, if  $Y$  has density  $p$  with CDF  $F$ , then the following relation holds in law:

$$Y = (F^{-1} \circ G)(X). \quad (8)$$

Equation (8) forms the basis of the stochastic collocation method. In our case, we can expect  $F^{-1} \circ G$  to be close to the identity since  $p_0$  is close to  $p$ . The idea is then to approximate  $F^{-1} \circ G$  by Lagrange interpolation with appropriate nodes, and since  $F^{-1}(G(x))$  is almost linear, only a small numbers of nodes is required. The cost of inverting  $F^{-1}$  is then minor. The function  $F^{-1} \circ G$  is represented in Figure 4 for several value of  $\alpha$ . As in Figure 3, the time  $t$  used for the simulations is the value of a typical stepsize  $h$  (see further) for the resolution of the transport equation, that is  $h = 0.03$ . It is clearly seen that indeed  $(F^{-1} \circ G)(x)$  is close to being linear, the nonlinear behavior being stronger for small  $\alpha$ . The main reason is that we neglected the term  $\beta_0$  that becomes larger as  $\alpha$  decreases.

A natural question is to wonder whether other (simple) distributions for  $X$  in (8) provide as good candidates as the wrapped alpha-stable. We investigate this by considering a wrapped Gaussian distribution:

$$p_G(t, \phi) = \frac{1}{\sqrt{2\pi t D}} \sum_{\ell \in \mathbb{Z}} e^{-\frac{(2\pi\ell + \phi)^2}{2tD}}.$$

Such a random variable can be easily simulated by drawing a Gaussian random variable modulo  $2\pi$ . Denoting by  $G_1$  the associated CDF, plots of  $F^{-1} \circ G_1$  are given in Figure 4 for several values of  $D$ . It is apparent that  $F^{-1} \circ G_1$  is far from being linear, and therefore that the function  $G$  is a much better choice.

Interpolation of  $(F^{-1} \circ G)(x)$  is generally done with Gauss quadrature nodes for good accuracy. We actually use here the Gauss-Lobatto rule that includes the endpoints, which we noticed was giving slightly better results. How these nodes are calculated is

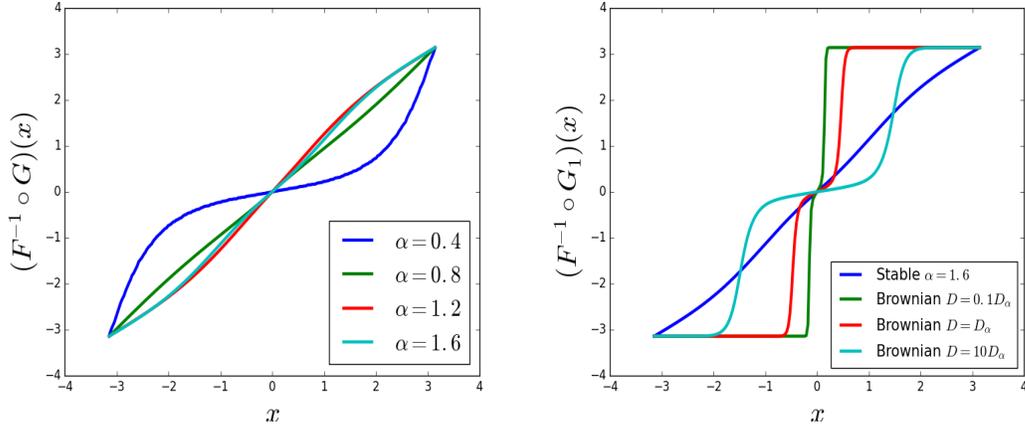


Figure 4: Left: the function  $(F^{-1} \circ G)(x)$  for several values of  $\alpha$ . Note the close to linear behavior. Right: the function  $(F^{-1} \circ G_1)(x)$  for several values of  $D_0$ . The nonlinear behavior is apparent. For both figures,  $t = h = 0.03$ , and  $a_0 = 0.1$ .

explained in Section A.4 of the Appendix. Given a collection  $(x_i)_{i=1, \dots, N_I}$  of collocation points in  $[0, \pi]$ , (8) is replaced by, using the symmetry of  $F^{-1} \circ G$  around the origin,

$$Y = \text{sign}(X) \sum_{i=1}^{N_I} F^{-1}(x_i) L_{N_I}(|X|) := H_{N_I}(X),$$

where  $L_{N_I}$  is a Lagrange interpolation polynomial of degree  $N_I - 1$ . The Gauss-Lobatto rule with  $N_I$  points is exact for polynomials up to order  $2N_I - 1$ . We display in Figure 5 the function  $F^{-1} \circ G$  on  $[0, \pi]$ , its approximation  $H_{N_I}$ , and the Gauss-Lobatto nodes for  $N_I = 5$  and  $N_I = 9$ , and various values of  $\alpha$ . We set as before  $t = h = 0.03$ . We observe a very good fit with an  $\ell_2$  relative error between  $F^{-1} \circ G$  and  $H_{N_I}$  of 0.5%, 0.07%, 0.1%, and 0.3%, for  $\alpha = 0.4$ ,  $\alpha = 0.8$ ,  $\alpha = 1.2$ ,  $\alpha = 1.9$ , respectively. The convergence of the

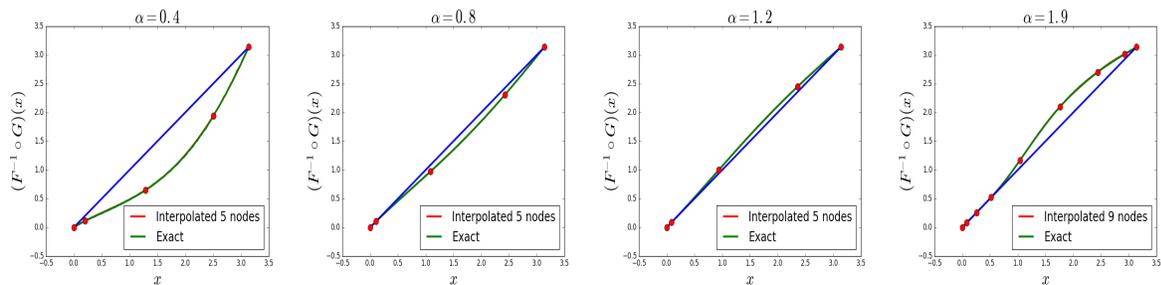


Figure 5: Comparison between  $F^{-1} \circ G$  and  $H_{N_I}$  for several values of  $\alpha$ . The blue line is the identity.

stochastic collocation method as  $N_I \rightarrow \infty$  is addressed in [15].

The construction of the interpolation function  $H_{N_I}$  can be done offline (the cost is negligible compared to that of a whole simulation of the RTE), so that the overall cost of the AS method includes the generation of  $X$ , and the calculation of a low order polynomial (say maximum of order 10).

The next section is devoted to the simulation of the jump distribution  $\pi_\varepsilon$  introduced in Section 2.2.

## 4 Algorithms for the jump distribution

We present here two methods to draw random numbers according to  $\pi_\varepsilon$ : rejection sampling, and the stochastic collocation. Since the intensity  $\mu_\varepsilon$  is small, many drawings according to  $\pi_\varepsilon$  have to be performed, and an efficient way to do this is necessary. The main difficulty is the presence of the singularity at  $\theta = 0$ , that is not integrable when the cut-off is removed, and is driving the general shape of  $\pi_\varepsilon$ , see Figure 6 for an illustration. The key is therefore to handle appropriately this singularity. The qualitative behaviors are similar around  $\theta = 0$  for non constant functions  $a$  bounded below and above.

We present first a carefully designed rejection method, that will be used as a benchmark for the stochastic collocation method introduced in section 4.2.

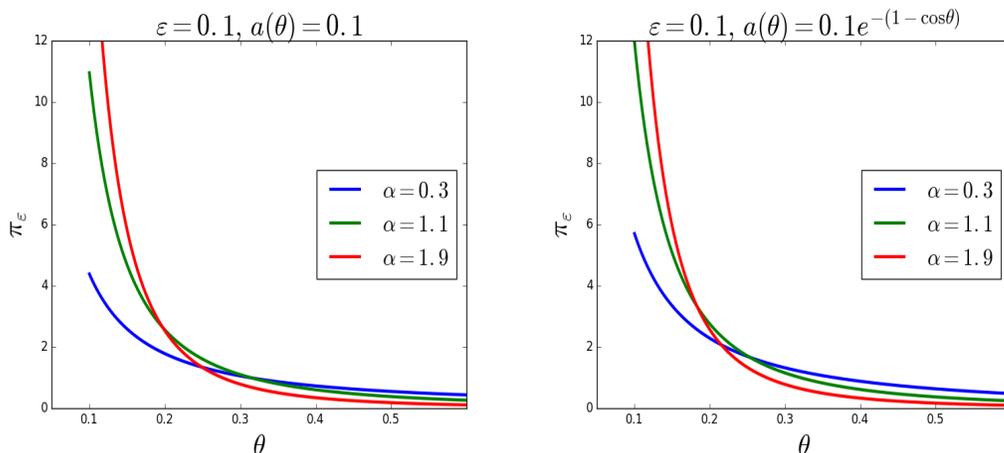


Figure 6: Representations of  $\pi_\varepsilon(\theta)$  for  $a(\theta) = 0.1$  and  $a(\theta) = 0.1e^{-(1-\cos(\theta))}$ ,  $\varepsilon = 0.1$  and several values of  $\alpha$ .

### 4.1 Rejection sampling

The method relies on the condition that the density  $\pi_\varepsilon$  can be bounded by another density (a proposal), which, due to the singularity, limits the options. A good proposal for the rejection sampling has to capture properly this singularity and has to be easily simulated. Such a proposal can be derived using a change of variables as described in the following lemma, proved in Appendix A.3:

**Lemma 4.1** *Let  $V$  be a random variable with density*

$$\pi_{\varepsilon,V}(v) = \frac{1}{2^{(3\alpha-1)/2}\Pi_{\varepsilon,+}^0} \cdot \frac{a(4 \arctan(v))(1+v^2)^\alpha}{v^{1+\alpha}} \mathbb{1}_{(\varepsilon/4,1)}(v),$$

where  $\Pi_{\varepsilon,+}^0$  is defined by (4). Then, the random variable  $\Theta = 4 \arctan(V)$  has  $\pi_\varepsilon$  for density.

The density  $\pi_{\varepsilon,V}$  can then be bounded as follows:

$$\pi_{\varepsilon,V}(v) \leq c_\varepsilon f_P(v), \quad \text{where} \quad c_\varepsilon := \frac{\bar{a}(1 - \varepsilon'^\alpha)}{2^{(\alpha-1)/2} \Pi_{\varepsilon,+}^0 \alpha \varepsilon'^\alpha}, \quad \varepsilon' = \frac{\varepsilon}{4},$$

and

$$f_P(v) := \frac{C_\varepsilon}{v^{1+\alpha}} \mathbb{1}_{(\varepsilon/4,1)}(v), \quad \text{with} \quad C_\varepsilon := \frac{\alpha \varepsilon'^\alpha}{1 - \varepsilon'^\alpha}.$$

Inverting the CDF corresponding to  $f_P$ , one can see that

$$W = \frac{\varepsilon'}{(1 - (1 - \varepsilon'^\alpha)U)^{1/\alpha}},$$

has  $f_P$  for density when  $U$  is uniformly distributed over  $(0,1)$ . This distribution is known as a bounded Pareto distribution with parameters  $(\alpha, \varepsilon, 1)$ . Note that according to (5), the constant  $c_\varepsilon$ , central to the rejection method, is of order 1 in  $\varepsilon$ . One can then expect a rejection rate weakly sensitive to  $\varepsilon$ . The algorithm goes as follows:

Until acceptance:

1. Draw a random variable  $W$  according to  $f_P$ .
2. Draw a random number  $U$  uniformly distributed over  $(0,1)$ .
3. Accept  $W$  if

$$U \leq \frac{\pi_{\varepsilon,V}(W)}{c_\varepsilon f_P(W)} = \frac{a(4 \arctan(W))}{\bar{a}2^\alpha} (1 + W^2)^\alpha,$$

and reject otherwise.

In order to illustrate how  $f_P$  approximates  $\pi_{\varepsilon,V}$  and how the rejection method performs, we provide Quantile-Quantile plots (QQ-plots from now on) of  $\pi_{\varepsilon,V}$  versus  $f_P$  in Figure 7, and the acceptance rate  $1/c_\varepsilon$  are summarized in Table 1 for several values of  $\varepsilon$  and  $\alpha$ .

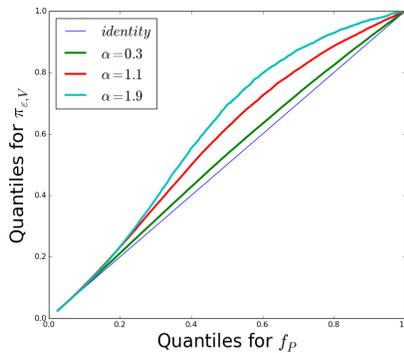


Figure 7: QQ-plot for  $\pi_{\varepsilon,V}$  versus  $f_P$ , from  $10^6$  realizations of the distributions, in the case  $a(\theta) = 0.1$ , and  $\varepsilon = 0.1$  and several values of  $\alpha$ .

From this results, one can see, as expected, that the acceptance rate is essentially constant with respect to  $\varepsilon$ . We observe as well that the rejection method performs worse as the singularity becomes stronger, that is as  $\alpha$  becomes closer to 2. We fix this issue in the next paragraph.

$1/c_\varepsilon$	$\alpha = 0.3$	1.1	1.9
$\varepsilon=0.5$	0.85	0.52	0.30
0.1	0.83	0.47	0.27
0.05	0.82	0.47	0.27
0.01	0.82	0.47	0.27

Table 1: Acceptance rate for the rejection method with proposal  $f_P$ , in the case  $a(\theta) = 0.1$ , for several values of  $\varepsilon$  and  $\alpha$ .

**Improved proposal.** Since most of the mass of the density is located around the singularity at 0, instead of bounding the term  $(1 + v^2)^\alpha$  by  $2^\alpha$ , we consider its Taylor expansion at first order when  $\alpha \in (0, 1]$  and second order when  $\alpha \in (1, 2)$ . Then, we have

$$\pi_{\varepsilon, V}(v) \leq \tilde{c}_\varepsilon \tilde{f}_P(v),$$

with

$$\tilde{f}_P(v) = \frac{1}{D_\varepsilon} \mathbb{1}_{(\varepsilon', 1)}(v) \times \begin{cases} \frac{1}{v^{1+\alpha}} + \alpha v^{1-\alpha} & \text{if } \alpha \in (0, 1] \\ \frac{1}{v^{1+\alpha}} + \frac{\alpha}{v^{\alpha-1}} + \frac{\alpha(\alpha-1)}{2} v^{3-\alpha} & \text{if } \alpha \in (1, 2), \end{cases} \quad (9)$$

where

$$D_\varepsilon = \begin{cases} \left( \frac{1 - \varepsilon'^\alpha}{\alpha \varepsilon'^\alpha} + \frac{\alpha(1 - \varepsilon'^{2-\alpha})}{2 - \alpha} \right) & \text{if } \alpha \in (0, 1] \\ \left( \frac{1 - \varepsilon'^\alpha}{\alpha \varepsilon'^\alpha} + \frac{\alpha(1 - \varepsilon'^{2-\alpha})}{2 - \alpha} + \frac{\alpha(\alpha-1)(1 - \varepsilon'^{4-\alpha})}{2(4 - \alpha)} \right) & \text{if } \alpha \in (1, 2), \end{cases}$$

and

$$\tilde{c}_\varepsilon = \frac{\bar{a}}{2^{(3\alpha-1)/2} \Pi_{\varepsilon, +}^0} D_\varepsilon.$$

Drawing random numbers from  $\tilde{f}_P$  is not difficult by using the decomposition

$$\tilde{f}_P(v) = \begin{cases} \mu_1 f_P^1(v) + \mu_2 f_P^2(v) & \text{if } \alpha \in (0, 1] \\ \mu_1 f_P^1(v) + \mu_2 f_P^2(v) + \mu_3 f_P^3(v) & \text{if } \alpha \in (1, 2), \end{cases}$$

where

$$\mu_1 = \frac{1 - \varepsilon'^\alpha}{D_\varepsilon \alpha \varepsilon'^\alpha}, \quad \mu_2 = \frac{\alpha(1 - \varepsilon'^{2-\alpha})}{D_\varepsilon(2 - \alpha)}, \quad \text{and} \quad \mu_3 = \frac{\alpha(\alpha-1)(1 - \varepsilon'^{4-\alpha})}{2D_\varepsilon(4 - \alpha)},$$

so that the sum of the  $\mu_j$  is one, and where the densities are

$$f_P^1 = f_P, \quad f_P^2(v) = \frac{2 - \alpha}{(1 - \varepsilon'^{2-\alpha})} v^{1-\alpha} \mathbb{1}_{(\varepsilon', 1)}(v) \quad \text{and} \quad f_P^3(v) = \frac{4 - \alpha}{(1 - \varepsilon'^{4-\alpha})} v^{3-\alpha} \mathbb{1}_{(\varepsilon', 1)}(v).$$

As we can see, compared to  $f_P$ ,  $\tilde{f}_P$  has corrective terms. We already know how to simulate  $f_P^1$ , and it is not difficult to see that

$$W_2 = (1 - (1 - \varepsilon'^{2-\alpha})U)^{1/(2-\alpha)} \quad \text{and} \quad W_3 = (1 - (1 - \varepsilon'^{4-\alpha})U)^{1/(4-\alpha)},$$

where  $U$  is uniformly distributed over  $(0, 1)$ , have respectively  $f_P^2$  and  $f_P^3$  for densities. Then, to draw random number according to  $\tilde{f}_P$  we proceed as follow:

1. Draw an integer  $j \in \{1, 2, 3\}$  according to the distribution  $(\mu_1, \mu_2, \mu_3)$ .
2. Draw a random number  $W$  according to  $f_P^j$ .

One can see from Figure 8, compared to Figure 7, that  $\tilde{f}_P$  is a better proposal than  $f_P$ , for which we then expect a better acceptance rate. This is confirmed with the results summarized in Table 2.

$1/\tilde{c}_\varepsilon$	$\alpha = 0.3$	1.1	1.9	$\ \pi_{\varepsilon,V} - \tilde{f}_P\ _\infty$	$\alpha = 0.3$	1.1	1.9
$\varepsilon=0.5$	0.993	0.999	0.999	$\varepsilon=0.5$	0.034	$2.5 \cdot 10^{-3}$	$2.7 \cdot 10^{-3}$
0.1	0.997	0.999	0.999	0.1	0.054	$2.1 \cdot 10^{-3}$	$7.6 \cdot 10^{-4}$
0.05	0.998	0.999	0.999	0.05	0.074	$2.0 \cdot 10^{-3}$	$4.1 \cdot 10^{-4}$
0.01	0.999	0.999	0.999	0.01	0.179	$1.7 \cdot 10^{-3}$	$9.8 \cdot 10^{-5}$

Table 2: Acceptance rate for the rejection method with proposal  $\tilde{f}_P$  and  $\|\pi_{\varepsilon,V} - \tilde{f}_P\|_\infty$ , in the case  $a(\theta) = 0.1$ , for several values of  $\varepsilon$  and  $\alpha$ .

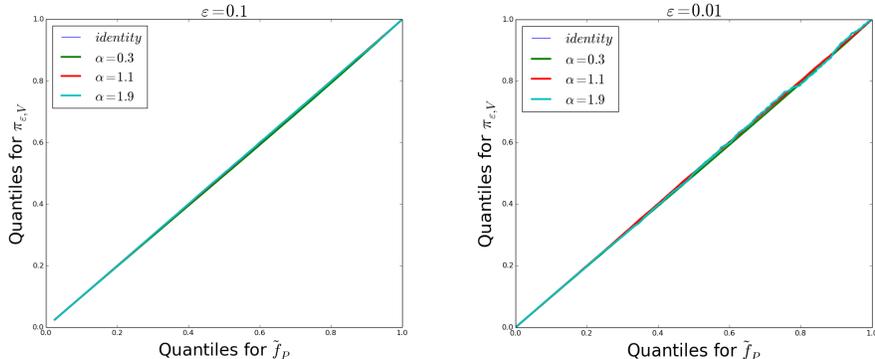


Figure 8: QQ-plot for  $\pi_{\varepsilon,V}$  versus  $\tilde{f}_P$ , from  $10^7$  realizations of the distributions, in the case  $a(\theta) = 0.1$ , for  $\varepsilon = 0.1$  and  $\varepsilon = 0.01$ , and several values of  $\alpha$ .

**Nonconstant function  $a(\theta)$ .** We now test the rejection method for a nonconstant function  $a$ , and choose as an example  $a(\theta) = 0.1e^{-(1-\cos(\theta))}$ . The performances are illustrated in Figure 9 and in Table 3. The method still works well, but now the results

are more sensitive to the parameter  $\alpha$  and  $\varepsilon$ . Note that the choice of function  $a$  here is particularly favorable since it presents a maximum at  $\theta = 0$  and decays fast away from zero. We will see an example of function  $a$  in the next section for which the rejection method fails. This is actually our motivation for designing an alternative approach. While the rejection method is well-suited to handle the singularities, it has a few flaws: the performance of strongly depends on the function  $a$ , and we can expect poor results if  $a$  has a minimum at  $\theta = 0$  for instance; also,  $\tilde{f}_P$  depends on the variable  $V = \tan(\theta/4)$ , and it would more convenient to work directly with the variable  $\cos(\theta)$  without using trigonometric formulas since  $\cos \theta$  appears directly in the kernel. We addresses these issues in the next section.

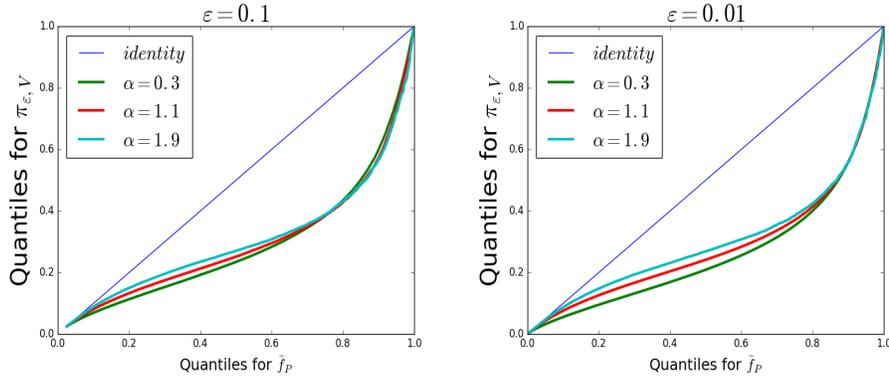


Figure 9: QQ-plot for  $\pi_{\varepsilon, V}$  versus  $\tilde{f}_P$ , from  $10^7$  realizations of the distributions, in the case  $a(\theta) = 0.1e^{-(1-\cos(\theta))}$ , for  $\varepsilon = 0.1$  and  $\varepsilon = 0.01$ , and several values of  $\alpha$ .

$1/\tilde{c}_\varepsilon$	$\alpha = 0.3$	1.1	1.9
$\varepsilon=0.1$	0.767	0.916	0.972
0.01	0.903	0.992	0.999

Table 3: Acceptance rate for the rejection method with proposal  $\tilde{f}_P$  and  $\|\pi_{\varepsilon, V} - \tilde{f}_P\|_\infty$ , in the case  $a(\theta) = 0.1e^{-(1-\cos(\theta))}$ , and for several values of  $\varepsilon$  and  $\alpha$ .

## 4.2 Stochastic collocation

We design here a method that is essentially independent of the choice of the function  $a$  (provided the latter is bounded below and above): we use the stochastic collocation method of section 3.3 with a simple proposal fast to compute, in particular simpler than the  $\tilde{f}_P$  of the last section. For this, we reparametrize  $\pi_\varepsilon(\theta')$  as follows: with  $\tau = (\cos \theta' + 1)/2 \in [0, 1]$ , we rewrite  $\pi_\varepsilon(\theta')$  as

$$\pi_\varepsilon(\theta')d\theta' \equiv \pi_\varepsilon(\tau)d\tau = \frac{\mathbb{1}_{[0, 1-\frac{\varepsilon^2}{4}]}(\tau)a(2\tau-1)}{2^{\frac{1+\alpha}{2}}C_\varepsilon(1-\tau)^{1+\frac{\alpha}{2}}\tau^{\frac{1}{2}}}d\tau,$$

where  $C_\varepsilon$  is an appropriate normalization constant. Above, we made the abuse of notation  $a(\theta) \equiv a(\cos \theta) \equiv a(2\tau - 1)$  and we slightly modified the definition of the interval  $I_\varepsilon$  and replaced it for convenience by

$$I_\varepsilon = \left\{ \theta \in [-\pi, \pi], 1 - \cos \theta > \frac{\varepsilon^2}{2} \right\}.$$

Consider then the density

$$q_1(x) = N_\varepsilon \mathbb{1}_{[0, 1 - \frac{\varepsilon^2}{4}]}(x) (1 - x)^{-1 - \frac{\alpha}{2}} \quad \text{with} \quad N_\varepsilon^{-1} = \frac{2}{\alpha} \left( \left( \frac{2}{\varepsilon} \right)^\alpha - 1 \right).$$

Random variables with density  $q_1$  can trivially be simulated since the cumulative distribution is invertible analytically. Hence, if  $U$  satisfies a uniform distribution on  $[0, 1]$ , then

$$X = 1 - \frac{1}{(1 + ((\frac{2}{\varepsilon})^\alpha - 1)U)^{\frac{2}{\alpha}}} \quad (10)$$

has density  $q_1$ . Note that the function  $q_1$  captures exactly the singularity of  $\pi_\varepsilon$  at  $\tau = 1$ .

Recall that the stochastic collocation method is based on the relation  $Y = (F^{-1} \circ G)(X)$ , where  $F$  is the CDF of our target distribution  $\pi_\varepsilon$  and  $G$  is a proposal distribution, that we choose with density  $q_1$ . We use here Gauss interpolation as, contrary to the AS method, it offered slightly better accuracy than Gauss-Lobatto.

In figure 10, we display the function  $F^{-1} \circ G$  for  $G = G_1$  and its interpolant for  $N_I = 5$  and  $N_I = 15$  collocation points. For  $N_I = 5$ , the relative  $\ell_2$  error between  $F^{-1} \circ G$  and its interpolant is 2%, 2.7%, and 1.3% for  $\alpha = 0.3$ ,  $\alpha = 1.1$ , and  $\alpha = 1.9$ , respectively. When  $N_I = 15$ , the error becomes 0.1%, 0.09%, and 0.07%, and the fit is excellent. In figure 11, we zoom around the singularities of  $\pi_\varepsilon$  at  $x = 0$  and  $x = 1$  where the density is larger. We observe as well an almost perfect fit. The generation of the interpolation polynomial has a negligible cost and is done offline, before propagating the particles. Also, the additional cost of using 15 collocations points instead 5 points is very minor, and as a consequence we will set  $N_I = 15$  in the simulations. In terms of overall cost,  $\pi_\varepsilon$  is simulated by generating a random number  $U$  with a uniform distribution, and by computing a polynomial of degree 15 at the point  $X$  drawn from  $q_1$ .

In figure 12, we plot  $F^{-1} \circ G$  for nonconstant functions  $a$ . The important observation is that the behavior  $F^{-1} \circ G$  is qualitatively the same as  $a$  varies, and therefore that the stochastic collocation method will be just as efficient as in the case where  $a$  is a constant. This is expected since the singularities of  $\pi_\varepsilon$  are not affected by  $a$  when the latter is bounded below and above.

## 5 Simulations

We need first to define a typical time scale. When the scattering kernel is integrable, one is given by the mean free time introduced earlier, and corresponding to  $\mu_\varepsilon^{-1}$ . In our non-integrable case, the mean free time is zero, and we need a different quantity. A natural one is the inverse of  $-\lambda_1$ , that is the inverse of the first non zero eigenvalue of  $\mathcal{Q}$ . When  $t \gg t_S := (-\lambda_1)^{-1}$ , the distribution of the angle is almost uniform and particles are in a diffusive regime. At such large times, the solution  $f(t, x, k)$  to the

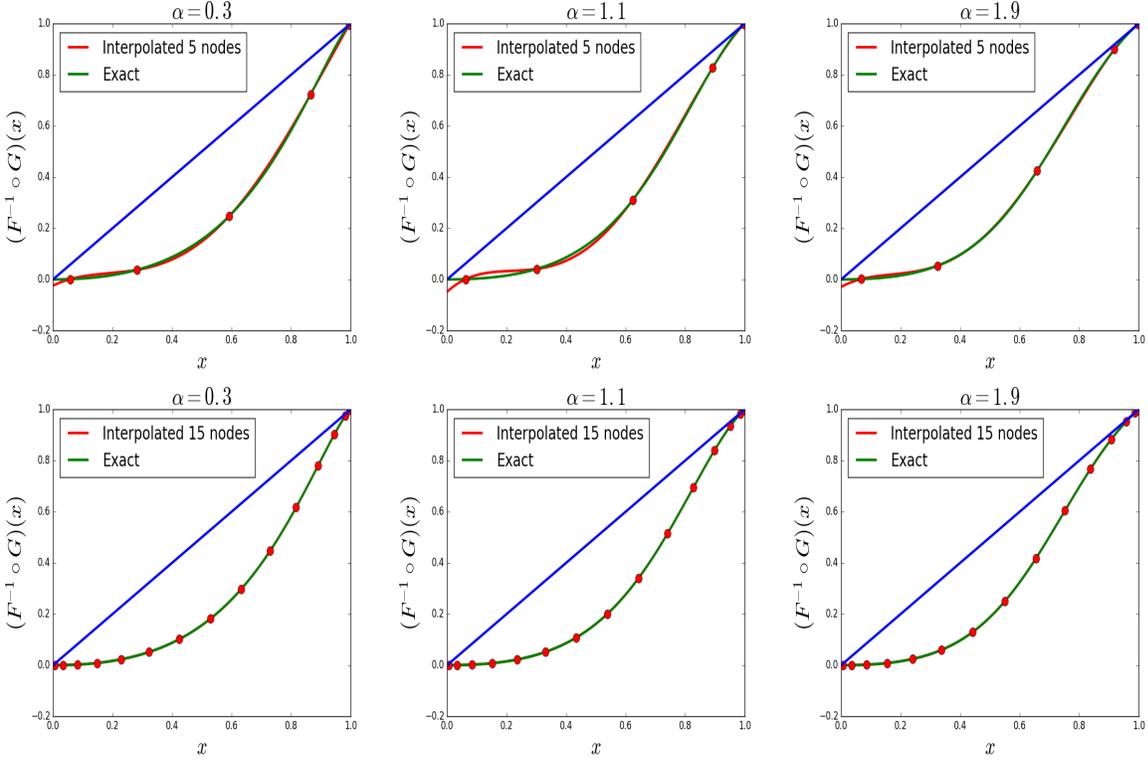


Figure 10: Interpolation of the function  $F^{-1} \circ G$  with Gauss collocation points for  $\varepsilon = a(\theta) = 0.1$ . In the case  $N_I = 15$  (bottom figures),  $F^{-1} \circ G$  and its interpolant cannot be distinguished, with relative  $\ell_2$  errors of 0.1%, 0.09%, and 0.07%, for  $\alpha = 0.3$ ,  $\alpha = 1.1$ , and  $\alpha = 1.9$ , respectively.

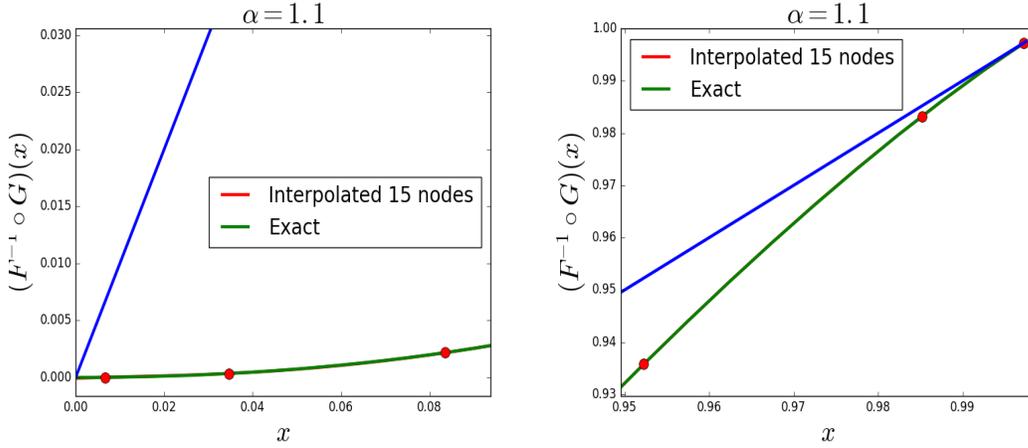


Figure 11: Zoom of figure 10 for  $\alpha = 1.1$  and  $N_I = 15$  around the singularities of  $\pi_\varepsilon$  at  $x = 0$  and  $x = 1$ . Again, the fit is excellent and  $F^{-1} \circ G$  and its interpolant cannot be distinguished.

RTE (1) does not depend on  $k$  anymore and satisfies a diffusion equation [22] (see also [13] in the context of singular kernels). For these time scales, it is preferable to solve a diffusion equation instead of the RTE, and we therefore consider times not significantly

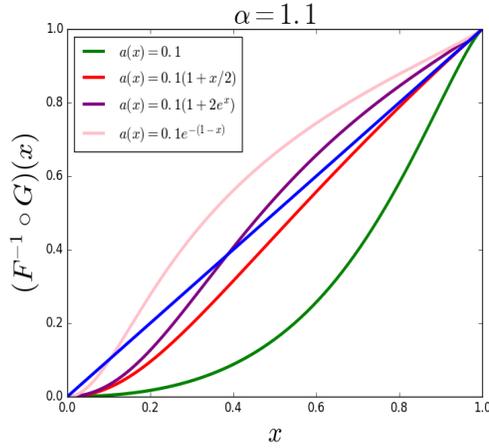


Figure 12: Representation of  $F^{-1} \circ G$  for  $\varepsilon = 0.1$ ,  $\alpha = 1.1$ , and several functions  $a(x)$ . The qualitative behavior stays the same.

larger than  $t_S$ .

We represent in Figure 13 the characteristic time  $t_S$  as a function of  $\alpha$  for  $a(x) = a_0 = 0.1$ . It does not vary much for  $\alpha \in (0, 1.5)$ , and then drops to zero when  $\alpha \in (1.5, 2)$ . We plot in Figure 14 the mean free time  $\mu_\varepsilon^{-1}$  of the ACR method, as a function of  $\alpha$ ,

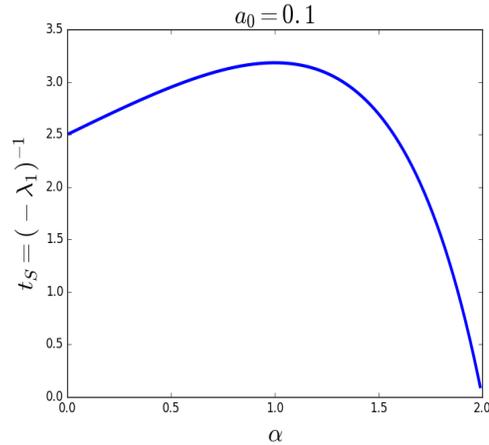


Figure 13: Characteristic time  $t_S$  as a function of  $\alpha$  for  $a(x) = a_0 = 0.1$ .

for  $\varepsilon = 0.03, 0.1, 0.3$  and various functions  $a$ . When  $a(x) = a_0 = 0.1$  and  $\varepsilon = 0.1$  for instance, the mean free time can be as low as 0.1 for  $\alpha$  close to 2, or about 0.5 when  $\alpha = 1$ , which is significantly smaller than the characteristic time  $t_S \simeq 3$  (see fig. 13). Results are roughly quantitatively the same for the non constant functions  $a$  of the figure. Figures 6 and 14 together illustrate well the typical behavior of the jump process  $(\theta_\varepsilon(t))_{t \geq 0}$  of Section 2.2: when  $\alpha$  is small, the mean free time is of order of  $t_S$ , and few collisions take place. Figure 6 shows that the jumps can be of large amplitude since the probability of moving away from  $\theta = 0$  where the density is largest is not negligible. When  $\alpha$  is larger, the collision rate increases, and the amplitude of the jumps becomes smaller, as we see on Figure 6 that the density is very small away from  $\theta = 0$ .

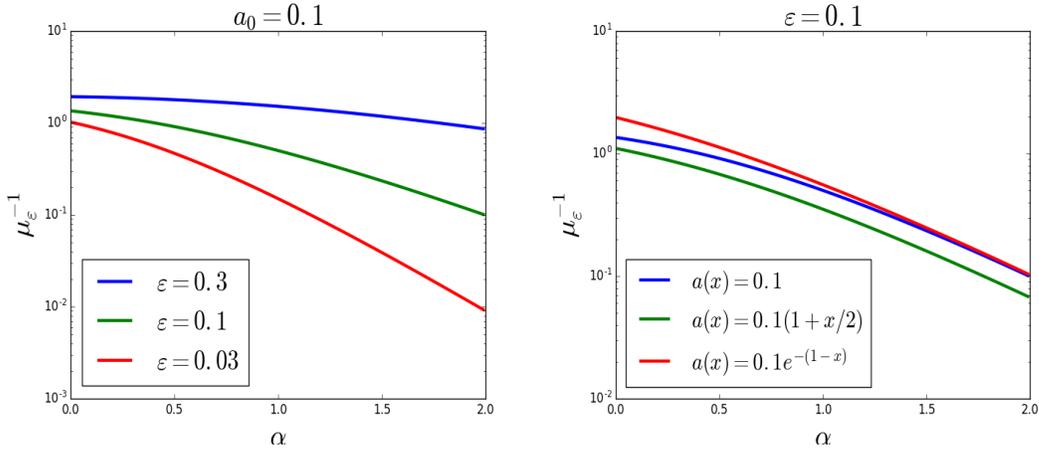


Figure 14: Mean free time  $\mu_\varepsilon^{-1}$  of the ACR method as a function of  $\alpha$ , for  $\varepsilon = 0.3, 0.1, 0.03$  and various functions  $a$ .

In the next section, we compare the rejection sampling and the collocation method for the simulation of the jump distribution.

## 5.1 Comparison of the algorithms for the jump distribution

We proceed as follows: we fix  $\varepsilon = 0.1$ , choose  $N_I = 15$  collocation points, and two functions  $a(\theta)$ ,  $a(\theta) = 0.1e^{-(1-\cos\theta)}$  and  $a(\theta) = 0.1(1 - 3\cos(\theta)/4)$ . We then generate  $10^7$  samples drawn from  $\pi_\varepsilon$  and compare the computational time of the two methods. The results are summarized in table 4. The different costs are expressed in terms of the cost of the collocation method. Simulations are performed on one core (Intel Xeon E5-2697A at 2.60GHz), and the unit is 1.46s for a code written in the Julia language. Note that these results are only indicative of trends and may vary on different platforms.

	$a(\theta) = 0.1e^{-(1-\cos\theta)}$		$a(\theta) = 0.1(1 - 3\cos(\theta)/4)$	
	Collocation	Rejection	Collocation	Rejection
$\alpha = 0.3$	1	2.2	1	6.3
$\alpha = 1.1$	1	2.2	1	9.4
$\alpha = 1.9$	1	2.05	1	11.2

Table 4: Comparison of the methods for the jump distribution for  $\varepsilon = 0.1$  and two functions  $a(\theta)$ .  $10^7$  samples are drawn from  $\pi_\varepsilon$ , with units of 1.46s on one core (Intel Xeon E5-2697A at 2.60GHz).

It is apparent that the collocation method is more efficient, which is expected since it does not depend on the function  $a$ . Its cost is also independent of  $\alpha$ , which is clear according to the construction of the method. The difference in the computational time between the two functions  $a$  for the rejection method is explained as follows: the function

$(1 - 3 \cos(\theta)/4)$  has a minimum at  $\theta = 0$ , around which most of the drawings of the rejection are done, while  $e^{-(1-\cos\theta)}$  has its maximum at  $\theta = 0$  which limits the impact of a nonconstant function  $a$ . The results are very similar for other values of  $\varepsilon$  like  $\varepsilon = 0.03$  or  $\varepsilon = 0.3$ .

We compared the stochastic collocation with the Metropolis-Hastings algorithm [17, 24] and the stochastic step function method [34], with again better performances for the collocation.

Following the results of this section, we will use from now on the collocation method for the simulation of the jump distribution.

## 5.2 Impact of the Brownian correction of the ACR method

We investigate in this section the impact of the Brownian correction in the ACR method on the accuracy of the results. As mentioned in section 2.1, we expect this correction of size  $\varepsilon^{2-\alpha}$  to be crucial for large  $\alpha$ . We need a reference solution to quantify the error. In our case with a singular collision kernel, we were not able to find a simple analytical solution to the transport equation. We therefore derive a semi-analytical expression in Appendix A.5. The initial condition has the form, with  $x = (x_1, x_2)$  and  $k = (\cos \theta, \sin \theta)$ ,

$$f_0(x, \theta) = \frac{1}{\sqrt{2\pi c^2}} e^{-\frac{|x|^2}{2c^2}} \frac{1}{2\pi} (1 + \cos \theta), \quad (11)$$

where  $c = 0.1$  in all calculations. Denoting by  $f$  the solution to the RTE for  $a(\theta) = a_0$  constant, we have an expression for  $f$  integrated over the spatial direction  $x_2$  and the slab  $x_1 \in [z_1, z_2]$ , with the angle integrated over several bins  $[\theta_i, \theta_{i+1}]$ ,  $i = 1, \dots, N_B$ ,  $\theta_{i+1} = \theta_i + \Delta\theta$ . Our reference solution is hence the following function of  $t$  and  $i = 1, \dots, N_B$ ,

$$J_i(t) = \int_{z_1}^{z_2} \int_{\mathbb{R}} \int_{\theta_i}^{\theta_{i+1}} f(t, x_1, x_2, \theta) dx_1 dx_2 d\theta. \quad (12)$$

Set  $\varepsilon = a_0 = 0.1$ , and  $t = 3$ . For  $\alpha = 0.3$  and  $\alpha = 1.1$ ,  $t = 3$  is roughly the characteristic time  $t_S$  depicted in Figure 13, while when  $\alpha = 1.9$ , this is about  $3t_S$ . The direction  $K^\varepsilon(t)$  is generated following the ACR method, and the position  $X^\varepsilon(t)$  is calculated as explained below.

**Integrating  $K^\varepsilon(t)$ .** As seen in Section 2.3, there are two contributions to the angle defining the direction, the Brownian part  $\phi_\varepsilon$ , and the pure jump part  $\theta_\varepsilon$ , with

$$K^\varepsilon(t) = (\cos(\phi_\varepsilon(t) + \theta_\varepsilon(t)), \sin(\phi_\varepsilon(t) + \theta_\varepsilon(t))) := K^\varepsilon(\phi_\varepsilon(t), \theta_\varepsilon(t)).$$

Since  $\theta_\varepsilon$  is constant between two consecutive jumps (say at  $t_1$  and  $t_2$ ), the jump part can be integrated exactly. The integral needs to be discretized for the Brownian part though, and since the latter is continuous but not smooth, we just use a rectangle method of low order. As an example, let  $t_1$  and  $t_2$  ( $t_1 < t_2$ ) two successive jumps of  $\theta_\varepsilon$ , fix a discretization parameter  $h$  and write  $t_2 - t_1 = N_h h + r_h$ , where  $N_h$  is an integer and  $r_h \in [0, h)$ . For  $N_h \geq 1$ , we have then the following expression for the approximation of

$X(t)$ :

$$X^{\varepsilon,h}(t_2) = X^{\varepsilon,h}(t_1) + h \sum_{i=0}^{N_h-1} K^\varepsilon(\phi_\varepsilon(t_1 + ih), \theta_\varepsilon(t_1^+)) + r_h K^\varepsilon(\phi_\varepsilon(t_1 + N_h h), \theta_\varepsilon(t_1^+)),$$

and  $\phi_\varepsilon(t_2)$  for the next iteration is obtained by

$$\phi_\varepsilon(t_2) - \phi_\varepsilon(t_1 + N_h h) \sim \mathcal{N}(0, r_h D_\varepsilon),$$

while  $\theta(t_2^+)$  is the new angle after the jump. When  $N_h = 0$ , we simply have

$$X^{\varepsilon,h}(t_2) = X^{\varepsilon,h}(t_1) + (t_2 - t_1)K^\varepsilon(\phi_\varepsilon(t_1), \theta_\varepsilon(t_1^+)),$$

with update  $\phi_\varepsilon(t_2) - \phi_\varepsilon(t_1) \sim \mathcal{N}(0, (t_2 - t_1)D_\varepsilon)$ .

We plot trajectories  $((X^{\varepsilon,h}(t), \phi_\varepsilon(t) + \theta_\varepsilon(t) \bmod 2\pi))_{t \geq 0}$  in Figure 15 for several values of  $\alpha$  and initial condition  $((0, 0), 0)$ . We recall that the final time is  $t = 3$  and that  $\varepsilon = a_0 = 0.1$ . The parameter  $h$  is set to 0.03. When  $\alpha = 0.3$ , the mean free time  $\mu_\varepsilon^{-1}$  is 1.09, and the Brownian correction is weak (of order  $\varepsilon^{2-\alpha}$ ), so we can mostly see the pure jump part. When  $\alpha = 1.1$ , the Brownian correction becomes apparent, and is fully entangled with the jump part when  $\alpha = 1.9$ . We have  $\mu_\varepsilon^{-1} = 0.43$  for  $\alpha = 1.1$  and  $\mu_\varepsilon^{-1} = 0.12$  for  $\alpha = 1.9$ , note the decrease in the amplitude of the jumps from the case  $\alpha = 0.3$ .

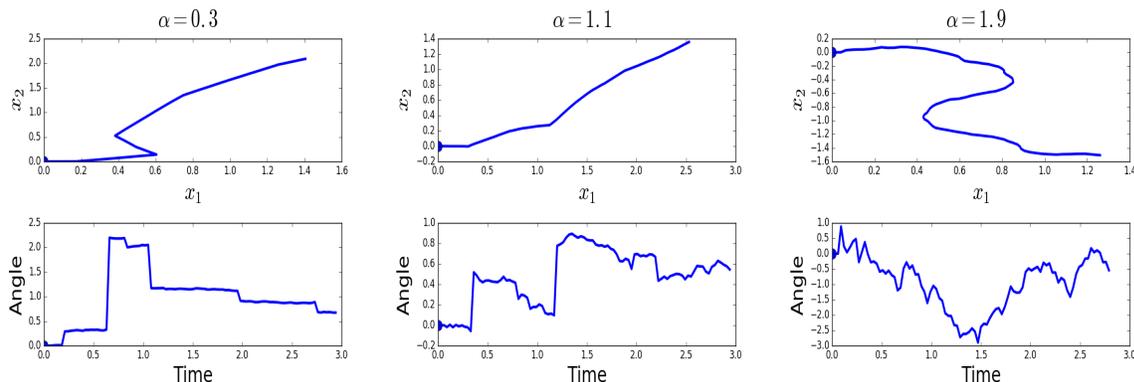


Figure 15: Trajectories  $(X^{\varepsilon,h}(t), \phi_\varepsilon(t) + \theta_\varepsilon(t) \bmod 2\pi)$  for several values of  $\alpha$  and initial condition  $((0, 0), 0)$ . The Brownian correction becomes more apparent as  $\alpha$  increases.

In Figure 16, we represent the reference solution  $(J_i(t))_{i=1, \dots, N_B}$  and the numerical solution to the RTE  $f_{\varepsilon,h}$ , including or not the Brownian correction. We use  $10^7$  trajectories,  $N_B = 80$  bins and set  $z_1 = 2.7$ ,  $z_2 = 3$ . When  $\alpha = 1.9$ , the solution is completely off when the correction is not included, while the relative error is about 3% when it is included (the error is calculated for only one simulation, and is therefore just indicative). When  $\alpha = 1.1$ , the relative errors are 6% without the Brownian, and 0.3% with, and 1.4% and 0.4% respectively for  $\alpha = 0.3$ . In all cases, this shows that the Brownian correction is crucial in order to obtain a good accuracy in the simulations.

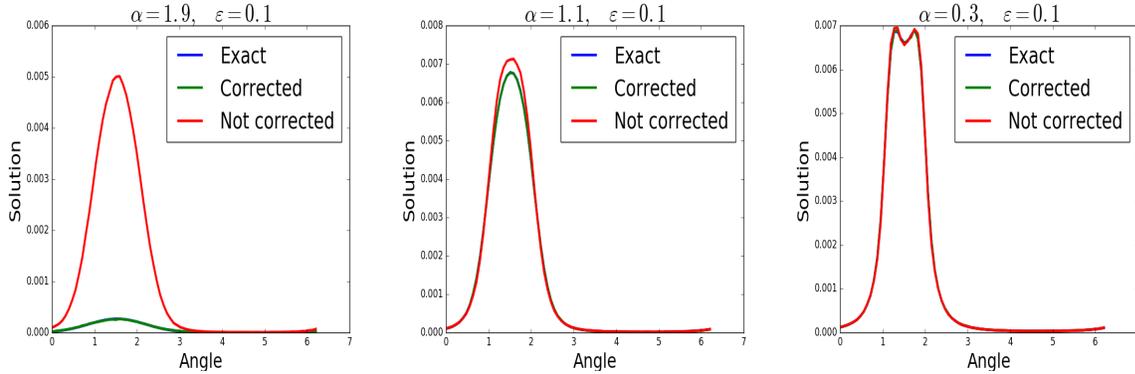


Figure 16: Comparison between the reference solution and the numerical solutions including or not the Brownian correction. The non corrected solution is completely off when  $\alpha = 1.9$ , while there is a good accuracy when the Brownian part is included.

### 5.3 Comparison ACR/AS methods

We compare in this section the performances of the ACR and AS methods. We do so in terms of accuracy versus computational cost for  $\alpha = 0.7, 1.3, 1.8$ . We do not address smaller values of  $\alpha$  since we will see that the ACR method is already much more competitive than AS for  $\alpha = 0.7$ . We fix  $a(x) = a_0 = 0.1$ , and use our reference solution  $(J_i(t))_{i=1, \dots, N_B}$  with  $N_B = 20$  to quantify the error. As in the previous section, we set  $t = 3$ , which is about the characteristic time  $t_S$  for  $\alpha = 0.7$  and  $\alpha = 1.3$ , and about  $2t_S$  for  $\alpha = 1.8$ . Since the error is itself a random variable, we average it over 20 simulations to obtain a more stable estimate. We then vary the number of particles, the time stepsize  $h$  (see the previous section), and the parameter  $\varepsilon$  for the ACR method, and record the minimal computational cost required to achieve a given accuracy. The results are depicted in Figure 17. The cost is expressed in units of 0.17s, which is the computational time to achieve an error of about 10% for a code written in the Julia language, and ran in parallel on a 32 cores Intel Xeon E5-2697A at 2.60GHz. We integrate the position  $X(t)$  in the ACR method as explained in the previous section, while  $X(t)$  in the AS case is calculated by a low order quadrature for the integral of  $K(t)$ .

In the case  $\alpha = 0.7$ , the cost of achieving an error of 0.2% is about 10 times higher for the AS method. The gap is reduced for lower accuracies, about twice the cost of AS when the error is 10% for instance. The fact that the ACR method is more efficient is confirmed with  $\alpha = 1.3$  and  $\alpha = 1.8$ , with the gap shrinking as  $\alpha$  increases. This is explained on the one hand by the fact that the ACR method is more sensitive to the parameter  $\alpha$  than AS since there are more and more collisions as  $\alpha$  grows, and on the other that the jumps are smaller for large  $\alpha$  which reduces the error in the calculation of  $X(t)$  in the AS method. Still, the cost of a 0.5% error when  $\alpha = 1.8$  is more than twice higher for the AS method. Note that the cost significantly increases when going from an accuracy of about 10% to about 0.1%.

The better performances of the ACR method can be explained as follows: while the process  $(K(t))_{t \geq 0}$  is simulated exactly in the AS method (in the sense that there is no approximation apart from the collocation sampling), the position  $X(t)$  is obtained by

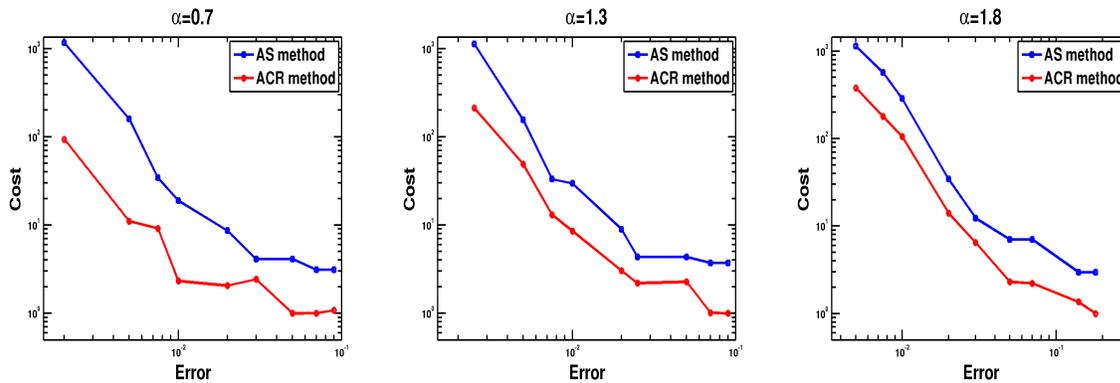


Figure 17: Comparison of the AS and ACR methods in terms of cost versus accuracy for  $\alpha = 0.7, 1.3, 1.8$ . The ACR method is consistently more efficient.

numerical integration of  $K(t)$ . Since the latter presents rough variations, a sufficiently fine discretization is needed. A better way to handle the integration of  $K(t)$  would be to adapt the stepsize to the random variations of  $K$ , which is essentially what is provided by the ACR method: the decomposition of the generator into the pure jump and the Brownian parts allows us to separate the large jumps from the smaller variations. As a consequence, the former part can be integrated exactly, leading to a reduced error compared to the AS method.

## 5.4 Sensitivity to the parameter $\varepsilon$

The parameter  $\varepsilon$  is crucial in the ACR method, and it is therefore important to investigate how it is affecting the efficiency. We already know that the mean free time is of order  $\varepsilon^\alpha$ , and that the error is of order  $\varepsilon^{4-\alpha}$ , which means that there is a trade-off between accuracy and cost in the choice of  $\varepsilon$  (see Section 2.3). We would like to investigate here this fact numerically. In Figure 18, left panel, we represent the error as a function of the number of particles, for  $\alpha = 0.7, 1.8$ , and  $\varepsilon = 0.03, 0.1, 0.3$  the higher group of curves corresponds to  $\alpha = 1.3$ . We are in the same setting as the last section, with  $t = 3$ ,  $a(x) = 0.1$ , and a stepsize  $h = 0.03$ . We observe as expected that the smaller the  $\varepsilon$ , the smaller the error with a significant margin for  $\alpha = 0.7$ , and a much smaller one when  $\alpha = 1.8$ . This is to be compared with the cost of decreasing  $\varepsilon$ : in Figure 18, right panel, we represent the cost of running a simulation for  $\varepsilon = 0.03, 0.1, 0.3$ . The reference is the cost of a simulation with  $10^6$  particles,  $h = 0.03$ ,  $\varepsilon = 0.3$  and  $\alpha = 0.7$  (the results are qualitatively similar for other choices). When  $\alpha = 0.7$  and  $\alpha = 1.3$ , the cost is not very different when decreasing  $\varepsilon$  from 0.3 to 0.03, while it increases significantly in the case  $\alpha = 1.8$  when going from 0.1 and 0.03 (this is explained by the mean free time of order  $\varepsilon^\alpha$ ). Since the errors between the cases  $\varepsilon = 0.1$  and  $\varepsilon = 0.03$  are quite similar as depicted in the left figure, this suggests that  $\varepsilon = 0.1$  is a good compromise and a better choice than  $\varepsilon = 0.3$  or  $\varepsilon = 0.03$ .

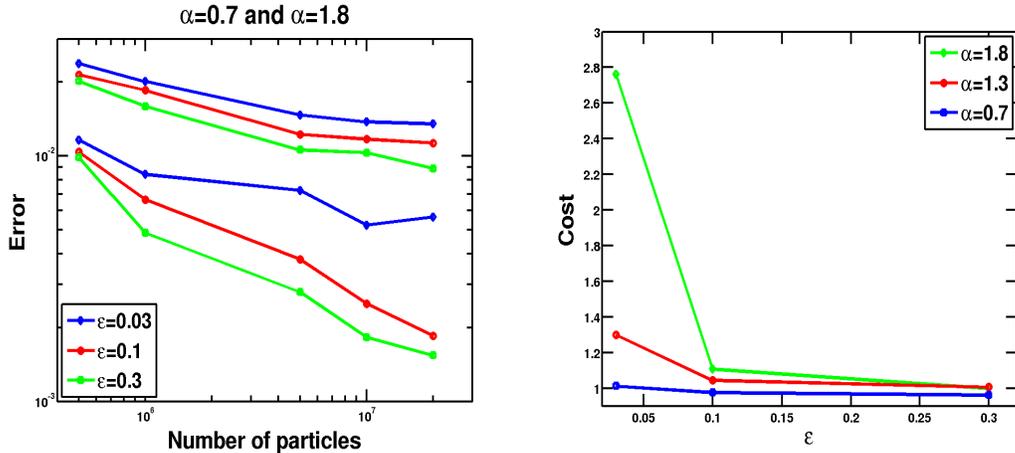


Figure 18: Left: error as a function of the number of particles for various values of  $\alpha$ . The upper group of curves corresponds to  $\alpha = 1.8$ , and the lower group to  $\alpha = 0.7$ . Right: computational cost as a function of  $\epsilon$ . Note the significant increases in the case  $\alpha = 1.8$ .

## 5.5 Comparison Short-Range versus Long-Range underlying random medium

Our goal in this section is to compare the solutions to the RTE (1) with singular kernels to solutions with integrable kernels. We choose a simple kernel given by a constant function, that is

$$\mathcal{Q}_I(f)(k) = a_1 \int_{\mathbb{S}^1} (f(p) - f(k)) d\sigma(p).$$

The process  $(K(t))_{t \geq 0}$  with generator  $\mathcal{Q}_I$  is straightforwardly simulated as it suffices to draw the angle uniformly over  $[-\pi, \pi]$ . For a meaningful comparison with singular kernels of the form (2), we choose  $a_1$  such that the mean free time associated with  $\mathcal{Q}_I$  (namely  $(2\pi a_1)^{-1}$ ) is equal to the characteristic time  $t_S$  associated with the singular kernel. We set as usual  $a(x) = 0.1$ , for  $\alpha = 0.5$  and  $\alpha = 1.8$ , we display in Figure 19 the solutions to the RTE on a slab of width 0.2 centered at  $x_2 = 0$ , as functions of  $x_1$ , for  $t = t_S, 2t_S, 3t_S$  and integrated in angle. We use the initial condition (11). For  $t = t_S$ , the three solutions for  $\alpha = 0.5$ ,  $\alpha = 1.8$ , and for the constant kernel case (referred as “Uniform” in the figure), are quite similar with the distinction that the scattered part (i.e. the part that lags behind the peak) is weaker in singular cases since scattering is mostly peaked forward. When  $t = 2t_S$ , significant differences appear, with in particular the fronts in the singular case being widened by the regularization of the RTE. The widening is stronger when  $\alpha$  is larger, as expected. On the other hand, the ballistic front in the uniform case does not change, also as can be expected. When  $t = 3t_S$ , the front in the case  $\alpha = 0.5$  merges with the scattered part, while it can still be observed when  $\alpha = 1.8$  since forward scattering is stronger.

These results substantiate the claim made in the introduction that inverse problems based on transport equations with singular kernels should be more ill-posed than those with integrable kernels: inversion strategies are often based on the singularities of the solution, that here are smoothed out by the RTE.

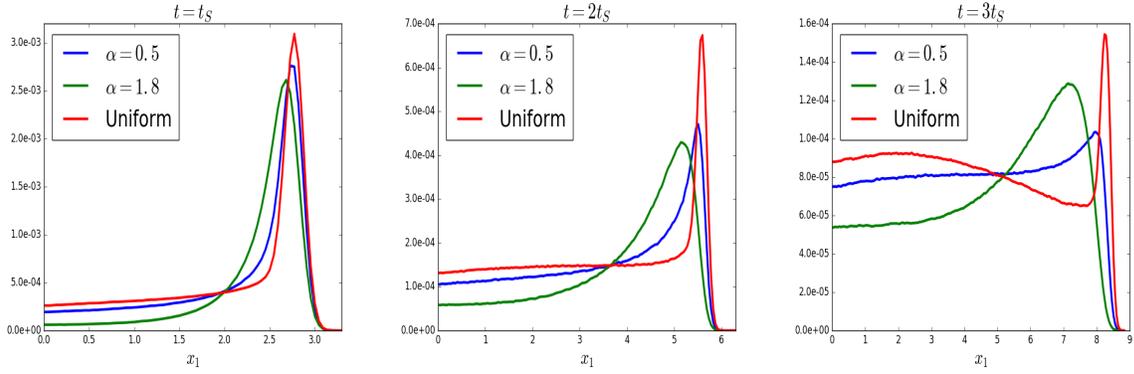


Figure 19: Comparison of solutions in the short-range (red curves) and long-range cases (blue and green curves) at multiples of the characteristic time. Note the regularization of the front in the long-range case.

We investigate in the next section the effect of the function  $a$  of the kernel on the form of the solution.

## 5.6 Impact of the function $a$

We use the same setting as in the last section, namely we represent the solutions on a slab around  $x_2 = 0$  at multiples of the characteristic time associated with a constant function  $a$ . We compare the results for  $a(u) = a_0 = 0.1$  (with  $u \equiv k \cdot p$ ),  $a(u) = a_1(u) = (2 + u)/30$ , and  $a(u) = a_2(u) = 0.1e^{-(1-u)}$ . The coefficients were chosen such that  $a(1) = a_1(1) = a_2(1)$ , that is such that the “amplitude” at the singularity is the same in the three cases. Results are displayed in figure 20. Since the function  $a_2$  is more peaked around  $u = 1$  than  $a_1$ , which is itself more peaked than  $a_0$ , its associated scattered part is weaker than in the two other cases.

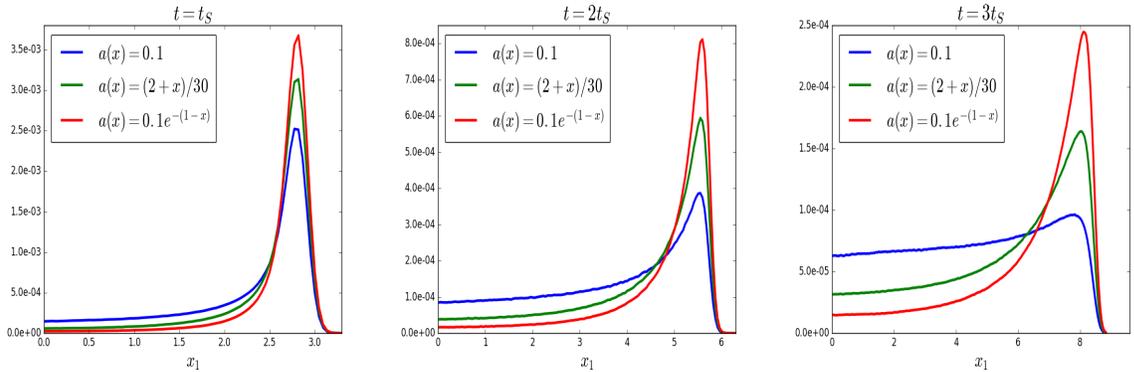


Figure 20: Comparison of the solutions for constant and non constant functions  $a$ . Observe the largest front for the most peaked function.

## 6 Generalizations

We explain in this section how the ACR method can be generalized to collision operators of the form

$$\mathcal{Q}(f)(x, k) = \int_{\mathbb{S}^1} \frac{f(p) - f(k)}{|k - p|^{1+\alpha(x)}} a(x, k, p) d\sigma(p),$$

which arise for instance when the statistical properties of the underlying random media vary with the position.

**The pure jump part.** We start with the generator  $\mathcal{Q}_>^\varepsilon$ . A standard method to handle compound Poisson processes with parameter-dependent Lévy measures is the so-called “fictitious shock method”, see [21], also referred to as “thinning”. The intensity  $\mu_\varepsilon$  depends now on  $(x, k)$ , and let

$$\bar{\mu}_\varepsilon := \sup_{x \in \mathbb{R}^2, k \in \mathbb{S}^1} \mu_\varepsilon(x, k).$$

We then generate a compound Poisson process with intensity  $\bar{\mu}_\varepsilon$  (note that the global supremum in the above definition can be localized to yield better efficiency). Suppose the process has a discontinuity at some time  $t$ . Then:

- With probability  $p = 1 - \mu_\varepsilon(X(t), K(t_-))/\bar{\mu}_\varepsilon$ , we set  $K(t_+) = K(t_-)$ , that is  $K$  is not modified.
- With probability  $1 - p$ ,  $K(t_-)$  is modified according to  $\pi_\varepsilon(X(t), K(t_-), p)$ , where

$$\pi_\varepsilon(x, k, p) d\sigma(p) = \frac{a(x, k, p)}{\pi_\varepsilon^0(x, k) |k - p|^{1+\alpha(x)}} d\sigma(p),$$

and

$$\pi_\varepsilon^0(x, k) = \int_{\mathbb{S}^1} \frac{a(x, k, p)}{|k - p|^{1+\alpha(x)}} d\sigma(p).$$

The method is exact, and the question left is how to simulate  $\pi_\varepsilon$ . The rejection method can readily be adapted, the constant  $\bar{a}$  being now the supremum of  $a$  over  $(x, k, p)$ . The stochastic collocation method can also be used in this framework as follows: rewrite  $a$  as  $a(x, k, p) = a(x, k \cdot p, p_\perp)$ , where  $p_\perp = p - (k \cdot p)k$ , and let  $A(x, k \cdot p) := \sup_{p_\perp} a(x, k \cdot p, p_\perp)$ . Defining

$$\nu_\varepsilon(x, p) d\sigma(p) = \frac{A(x, k \cdot p)}{\nu_\varepsilon^0 |k - p|^{1+\alpha(x)}} d\sigma(p),$$

where  $\nu_\varepsilon^0$  is an appropriate normalization,  $\pi_\varepsilon$  is simulated by rejection sampling with proposal  $\nu_\varepsilon$ . The distribution  $\nu_\varepsilon$  can then be simulated with the stochastic collocation method: discretize the range of the function  $\alpha$  into  $N_1$  intervals, and that of  $A$  into  $N_2$  intervals. Consider the resulting  $N_1 N_2$  densities  $\mu_{\varepsilon, i}$ ,  $i = 1, \dots, N_1 N_2$ . For each  $i$ , one can generate offline the interpolation polynomial as described in Section 4.2, and store the coefficients. The accuracy depends on  $N_1 N_2$ , but with a polynomial of order 15, we would just need to store  $15 N_1 N_2$  coefficients, which is not expensive.

**The Brownian part.** Following the calculations of section A.1, the coefficient  $a(0)$  in the definition of  $D_\varepsilon \equiv D_\varepsilon(x, k)$  has to be replaced by  $a(x, k, k)$ . Assuming there are no jumps between  $t$  and  $t + h$ , the angle  $\phi_\varepsilon$  can be approximately generated with the relation

$$\phi_\varepsilon(t + h) - \phi_\varepsilon(t) \sim \mathcal{N}(0, hD_\varepsilon(X(t), K(t))).$$

If there are jumps in  $[t, t + h]$ , then the interval is decomposed appropriately, and we apply a formula as above in each subintervals.

## 7 Conclusion

This work is devoted to Monte Carlo methods for radiative transfer equations with singular collision kernels. We compared the ACR method based on an approximation of the generator with the AS method based on the true generator. We showed the superiority of the ACR method, due in part to the separation of the large jumps contribution allowing for a more accurate integration of the momentum.

We also compared various methods for the simulation of the jump distribution, and obtained significantly better performances with the stochastic collocation method. We made clear of the importance of including the Brownian correction for very singular kernels, and compared solutions to the RTE in the singular case with solutions in the integrable case. We in particular confirmed the regularization effect of singular kernels.

This work is a first step in the simulation of high frequency wave energy transport in random media with long-range correlations, our ultimate goal being the resolution of imaging problems with transport-based techniques as in [4, 5]. Various relevant computational issues were not addressed here, in particular how to reduce the variance in the simulations, and how to adjust the parameter  $\varepsilon$  to it. A comparison with deterministic methods is also of importance. Note that these latter methods would need to take a particular care of the singularity of the kernel. These questions will be the object of future works, as well as the extension of the Monte Carlo methods to three-dimensional settings.

## A Appendix

### A.1 Proof of Lemma 2.1

Let us start by rewriting the generator as follows

$$\begin{aligned} \mathcal{Q}_<^\varepsilon(g)(\theta) &= \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{I_\varepsilon^C} \frac{a(\theta')(g(\theta' + \theta) - g(\theta))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' \\ &= \frac{a(0)}{2^{\frac{1+\alpha}{2}}} \int_{I_\varepsilon^C} \frac{g(\theta' + \theta) - g(\theta)}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' + \frac{1}{2^{\frac{1+\alpha}{2}}} \int_{I_\varepsilon^C} \frac{(a(\theta') - a(0))(g(\theta' + \theta) - g(\theta))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' \\ &:= \mathcal{R}_1^\varepsilon(\theta) + \mathcal{R}_2^\varepsilon(\theta), \end{aligned}$$

where  $I_\varepsilon^C = \{\theta' \in [-\pi, \pi], |\tan(\theta'/4)| \leq \varepsilon/4\}$ .

Regarding  $\mathcal{R}_1^\varepsilon$ , Taylor expanding  $g$  up to fourth order, and using symmetries to cancel the first and third order terms, we find, for some  $\xi_\theta \in (\theta, \theta + \theta')$ ,

$$\begin{aligned}\mathcal{R}_1^\varepsilon(\theta) &= \frac{a(0)g''(\theta)}{2^{\frac{1+\alpha}{2}+1}} \int_{I_\varepsilon^C} \frac{\theta'^2}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' + \frac{a(0)}{2^{\frac{1+\alpha}{2}} 4!} \int_{I_\varepsilon^C} \frac{\theta'^4 g^{(4)}(\xi_\theta)}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' \\ &:= \mathcal{L}_\varepsilon g(\theta) + \mathcal{R}_{12}^\varepsilon(\theta),\end{aligned}$$

where  $\mathcal{L}_\varepsilon g$  is the leading term. Regarding  $\mathcal{R}_2^\varepsilon$ , we perform again a Taylor expansion on  $g$  up to the fourth order and using again symmetries, we obtain

$$\mathcal{R}_2^\varepsilon(\theta) = \frac{g''(\theta)}{2^{\frac{1+\alpha}{2}+1}} \int_{I_\varepsilon^C} \frac{\theta'^2 (a(\theta') - a(0))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta' + \frac{1}{2^{\frac{1+\alpha}{2}} 4!} \int_{I_\varepsilon^C} \frac{\theta'^4 g^{(4)}(\xi_\theta) (a(\theta') - a(0))}{(1 - \cos(\theta'))^{\frac{1+\alpha}{2}}} d\theta'.$$

We use the following lemma in order to treat the terms above.

**Lemma A.1** *We have*

$$\left| \int_{I_\varepsilon^C} \frac{\phi^2 d\phi}{(1 - \cos(\phi))^{\frac{1+\alpha}{2}}} - \frac{2^{\frac{3+\alpha}{2}}}{2 - \alpha} \varepsilon^{2-\alpha} \right| \leq \frac{3}{2^{\frac{3-\alpha}{2}} (4 - \alpha)} \varepsilon^{4-\alpha},$$

and

$$\int_{I_\varepsilon^C} \frac{\phi^4}{(1 - \cos(\phi))^{\frac{1+\alpha}{2}}} d\phi \leq \frac{2^{\frac{5+\alpha}{2}}}{4 - \alpha} \varepsilon^{4-\alpha}.$$

As a result, Taylor expanding  $a$  up to the second order and using symmetries to cancel the first order term, we have for any  $\theta \in [\pi, \pi]$ ,

$$\left| \mathcal{Q}_<^\varepsilon(g)(\theta) - \frac{a(0)\varepsilon^{2-\alpha}}{2 - \alpha} g''(\theta) \right| \leq \frac{\varepsilon^{4-\alpha}}{4 - \alpha} \left( \|a\|_\infty + \|a''\|_\infty \right) \left( \|g''\|_\infty + \|g^{(4)}\|_\infty \right).$$

*Proof.* [of Lemma A.1] Using that

$$1 - \cos(\theta') = 2 \sin^2(\theta'/2) = \frac{4 \tan^2(\theta'/4)}{(1 + \tan^2(\theta'/4))^2},$$

we have, with the change of variable  $u = \tan(\phi/4)$ ,

$$\begin{aligned}\int_{I_\varepsilon^C} \frac{\phi^2 d\phi}{(1 - \cos(\phi))^{\frac{1+\alpha}{2}}} &= \frac{1}{2^{3\frac{1+\alpha}{2}}} \int_{I_\varepsilon^C} \frac{\phi^2 (1 + \tan^2(\phi/4))^{1+\alpha} d\phi}{|\tan(\phi/4)|^{1+\alpha}} \\ &= \frac{2^6}{2^{3\frac{1+\alpha}{2}}} \int_{-\varepsilon/4}^{\varepsilon/4} \frac{\arctan^2(u) (1 + u^2)^\alpha du}{|u|^{1+\alpha}}.\end{aligned}$$

Now, using standard analysis, we find

$$\left| \int_{-\varepsilon/4}^{\varepsilon/4} \frac{\arctan^2(u) (1 + u^2)^\alpha du}{|u|^{1+\alpha}} - \int_{-\varepsilon/4}^{\varepsilon/4} \frac{|u|^2 du}{|u|^{1+\alpha}} \right| \leq 6 \int_{-\varepsilon/4}^{\varepsilon/4} \frac{|u|^4 du}{|u|^{1+\alpha}},$$

with

$$\int_0^{\varepsilon/4} u^{3-\alpha} du = \frac{\varepsilon^{4-\alpha}}{2^{8-2\alpha}(4 - \alpha)} \quad \text{and} \quad \int_0^{\varepsilon/4} u^{1-\alpha} du = \frac{\varepsilon^{2-\alpha}}{2^{4-2\alpha}(2 - \alpha)}.$$

Gathering all previous estimates, we obtain

$$\left| \int_{I_\varepsilon^c} \frac{\phi^2 d\phi}{(1 - \cos(\phi))^{\frac{1+\alpha}{2}}} - \frac{2^{\frac{3+\alpha}{2}}}{2-\alpha} \varepsilon^{2-\alpha} \right| \leq \frac{3}{2^{\frac{3-\alpha}{2}}(4-\alpha)} \varepsilon^{4-\alpha},$$

concluding the proof of the first point. The second point of the lemma follows from the same lines as above, and we have

$$\begin{aligned} \int_{I_\varepsilon^c} \frac{\phi^4 d\phi}{(1 - \cos(\phi))^{\frac{1+\alpha}{2}}} &= \frac{2^{10}}{2^{3\frac{1+\alpha}{2}}} \int_{-\varepsilon}^{\varepsilon} \frac{\arctan^4(u)(1+u^2)^\alpha du}{u^{1+\alpha}} \\ &\leq \frac{2^{\frac{5+\alpha}{2}}}{4-\alpha} \varepsilon^{4-\alpha}. \end{aligned}$$

□

## A.2 Proof of Proposition 2.2

Let  $u^\varepsilon = f^\varepsilon - f$ , which we assume is smooth in order to justify the formal calculations. With the notation of Lemma 2.1,  $u^\varepsilon$  satisfies

$$\partial_t u^\varepsilon + k \cdot \nabla_x u^\varepsilon = \mathcal{Q}_>^\varepsilon u^\varepsilon + \frac{D_\varepsilon}{2} \partial_\theta^2 u^\varepsilon - \mathcal{R}_\varepsilon[f], \quad u^\varepsilon(t=0) = 0. \quad (13)$$

For  $(\cdot, \cdot)$  the scalar product on  $\mathbb{R}^2 \times \mathbb{S}^1$ , we have

$$(\mathcal{Q}_>^\varepsilon u^\varepsilon, u^\varepsilon) \leq 0, \quad (\partial_\theta^2 u^\varepsilon, u^\varepsilon) \leq 0,$$

and therefore, we obtain from (13),

$$\frac{1}{2} \frac{d}{dt} \|u^\varepsilon(t)\|_{L^2(\mathbb{R}_+ \times \mathbb{S}^1)}^2 \leq (\mathcal{R}_\varepsilon[f(t)], u^\varepsilon(t)).$$

This yields from the Cauchy-Schwarz inequality,

$$\frac{d}{dt} \|u^\varepsilon(t)\|_{L^2(\mathbb{R}_+ \times \mathbb{S}^1)} \leq \|\mathcal{R}_\varepsilon[f(t)]\|_{L^2(\mathbb{R}^2 \times \mathbb{S}^1)},$$

and the conclusion follows from Lemma 2.1.

## A.3 Proof of Lemma 4.1

Let us assume that  $\Theta$  is a random variable with density  $\pi_\varepsilon$ , and let  $g$  be a bounded continuous function. With the notation of section 2.2, we have

$$\begin{aligned} \mathbb{E}[g(\Theta)] &= \int_{I_{\varepsilon,+}} \frac{g(\theta)a(\theta)}{\Pi_{\varepsilon,+}^0(1 - \cos(\theta))^{(1+\alpha)/2}} d\theta \\ &= \frac{1}{2^{(1+\alpha)/2}\Pi_{\varepsilon,+}^0} \int_{I_{\varepsilon,+}} \frac{g(\theta)a(\theta)}{\sin^{1+\alpha}(\theta/2)} d\theta \end{aligned}$$

and making the change of variable  $\theta \rightarrow 2\theta$ ,

$$\mathbb{E}[g(\Theta)] = \frac{2}{2^{(1+\alpha)/2}\Pi_{\varepsilon,+}^0} \int_{I'_{\varepsilon,+}} \frac{g(2\theta)a(2\theta)}{\sin^{1+\alpha}(\theta)} d\theta,$$

with

$$I'_{\varepsilon,+} := \{\theta \in (0, \pi/2) \text{ s.t. } \tan(\theta/2) > \varepsilon/4\}.$$

Now, using that  $\sin(\theta) = 2 \tan(\theta/2)/(1 + \tan^2(\theta/2))$ , and making again the change of variable  $\theta \rightarrow 2\theta$ , we obtain

$$\begin{aligned} \mathbb{E}[g(\Theta)] &= \frac{2}{2^{3(1+\alpha)/2}\Pi_{\varepsilon,+}^0} \int_{I'_{\varepsilon,+}} \frac{g(2\theta)a(2\theta)}{\tan^{1+\alpha}(\theta/2)} (1 + \tan^2(\theta/2))^{1+\alpha} d\theta \\ &= \frac{2^2}{2^{3(1+\alpha)/2}\Pi_{\varepsilon,+}^0} \int_{I''_{\varepsilon,+}} \frac{g(4\theta)a(4\theta)}{\tan^{1+\alpha}(\theta)} (1 + \tan^2(\theta))^{1+\alpha} d\theta, \end{aligned}$$

with

$$I''_{\varepsilon,+} := \{\theta \in (0, \pi/4) \text{ s.t. } \tan(\theta) > \varepsilon/4\}.$$

Now, with  $v = \tan(\theta)$ , we finally obtain

$$\begin{aligned} \mathbb{E}[g(\Theta)] &= \frac{1}{2^{(3\alpha-1)/2}\Pi_{\varepsilon,+}^0} \int g(4 \arctan(v)) \frac{a(4 \arctan(v))(1 + v^2)^\alpha}{v^{1+\alpha}} \mathbb{1}_{(\varepsilon/4,1)}(v) dv \\ &= \mathbb{E}[g(4 \arctan(V))], \end{aligned}$$

which concludes the proof.

## A.4 Gauss and Gauss-Lobatto quadratures

**Gauss quadrature.** This is classical material, and the standard method is the Golub-Welsch algorithm [11]. For a measure  $\mu(x)$ , and given a three-term recurrence relation between orthogonal polynomials  $P_k$  for the measure  $\mu$  of the form

$$P_{k+1}(x) = (x - \alpha_k)P_k(x) - \beta_k P_{k-1}(x), \quad k \geq 0, \quad P_{-1}(x) = 0, \quad P_0(x) = 1,$$

one forms the  $N \times N$  symmetric tridiagonal matrix  $J$  with diagonal  $(\alpha_0, \dots, \alpha_{N-1})$  and upper diagonal  $(\sqrt{\beta_1}, \dots, \sqrt{\beta_{N-2}})$ . The collocation nodes of the  $N$  nodes Gauss quadrature are then the eigenvalues of  $J$ . The important point is therefore to obtain the coefficient  $\alpha_k$  and  $\beta_k$ . There is simple method based on the moments of  $\mu$ , which unfortunately becomes numerically unstable when about more than 10 collocations points are needed. We then use a different approach based on a discretization method explicated in [9]. The idea is to discretize the measure  $\mu$  as

$$\mu(x) = \sum_{p=1}^P w_p \delta(x - x_p),$$

for some weights  $w_p$  and points  $x_p$  to be determined. We then obtain a three-term recurrence relation with coefficients  $(\alpha_k^P, \beta_k^P)$  for this discrete measure using the Stieljes

algorithm (see again [9]), where the limit of  $(\alpha_k^P, \beta_k^P)$  as  $P \rightarrow \infty$  yields the coefficients of the original measure  $\mu$ . The choice of  $w_p$  and  $x_p$  is important for the efficiency of the method, and is done as follows: for some weight function  $\omega(x)$  for which Gauss quadrature nodes  $y_p$  and weights  $\gamma_p$  can be easily calculated, we write (with an abuse of notation, we suppose  $\mu$  has density  $\mu(x)$ ), for some interval  $I$ ,

$$\int_I f(x)\mu(x)dx = \int_I \frac{f(x)\mu(x)}{\omega(x)}\omega(x)dx \simeq \sum_{p=1}^P \frac{f(y_p)\mu(y_p)}{\omega(y_p)}\gamma_p.$$

We then set  $x_p = y_p$ , and  $w_p = \mu(x_p)\gamma_p/\omega(x_p)$ . As an example, consider the collocation method of Section 4.2. We have  $\mu(x) = q_1(x)$  for  $x \in [0, 1 - \frac{\varepsilon^2}{4}]$ , and we choose  $\omega(x) = (1-x)^{\delta-1}$ , for  $\delta = 0.1$ , which is a Gauss-Jacobi type weight. This choice is made since, as the function  $p_1$ ,  $\omega$  has a singularity at  $x = 1$ . The associated nodes and weights are known and tabulated, generally over the interval  $[-1, 1]$ . We then do a linear transformation to relocate them to the interval  $[0, 1 - \frac{\varepsilon^2}{4}]$ . The overall cost of the calculation of the collocation points is negligible compared to complete simulation of the RTE.

**Gauss-Lobatto quadrature.** The procedure is direct using what is above. If  $I = [a, b]$ , then define the measure  $\nu(x) = (x-a)(b-x)\mu(x)$ . The interior nodes of the Gauss-Lobatto rule for  $\mu$  are then the Gauss nodes for  $\nu$ , see [8].

## A.5 Semi-analytical solution

We derive here a representation formula for a particular solution to (1). For  $f$  the solution to (1) with an initial condition of the form (11) and  $x = (x_1, x_2)$ , let

$$g(t, x_1, k) = \int_{\mathbb{R}} f(t, (x_1, x_2), k)dx_2, \quad k \in \mathbb{S}^1.$$

Writing  $k = (\cos \theta, \sin \theta)$  and  $z = x_1$  for convenience,  $g$  solves (assuming  $f$  decays sufficiently fast at the infinity),

$$\partial_t g + \cos \theta \partial_z g = \mathcal{Q}(g) = a_0 \int_{\mathbb{S}^1} \frac{g(p) - g(k)}{|k - p|^{1+\alpha}} d\sigma(p).$$

We decompose then  $g$  into Fourier modes as

$$g(t, z, \theta) = \sum_{\ell \geq 0} g_\ell(t, z) e_\ell(\theta), \quad \text{with} \quad e_0 = \frac{1}{\sqrt{2\pi}}, \quad e_\ell = \frac{\cos \ell \theta}{\sqrt{\pi}}, \quad \ell \geq 1.$$

Denoting by  $\lambda_\ell$  the eigenvalues of the operator  $\mathcal{Q}$  explicitly defined in (7), we find the system of equations

$$\begin{cases} \partial_t g_0 + \frac{1}{\sqrt{2}} \partial_z g_1 = \lambda_0 g_0 \\ \partial_t g_1 + \frac{1}{\sqrt{2}} \partial_z g_0 + \frac{1}{2} \partial_z g_2 = \lambda_1 g_1 \\ \partial_t g_\ell + \frac{1}{2} \partial_z g_{\ell-1} + \frac{1}{2} \partial_z g_{\ell+1} = \lambda_\ell g_\ell, \quad \ell \geq 2. \end{cases}$$

After taking the Fourier transform in  $z$ , we find

$$\partial_t \hat{g}_\ell(t, \xi) = \lambda_\ell \hat{g}_\ell - i\xi(c_\ell \hat{g}_{\ell+1} + d_\ell \hat{g}_{\ell-1}) \quad (14)$$

where

$$c_0 = \frac{1}{\sqrt{2}}, \quad d_0 = 0, \quad d_1 = \frac{1}{\sqrt{2}}, \quad c_\ell = \frac{1}{2}, \quad \ell \geq 1, \quad d_\ell = \frac{1}{2}, \quad \ell \geq 2.$$

Storing the coefficient  $\hat{g}_\ell$  up to some order  $\ell = L$  into a vector  $\hat{g}_L$ , (14) can be written into the matrix form

$$\partial_t \hat{g}_L(t, \xi) = A_L(\xi) \hat{g}_L(t, \xi).$$

This gives us an approximation  $g_L$  of  $g$  represented by

$$g_L(t, z, \theta) = \frac{1}{2\pi} \sum_{\ell=0}^L \left( \int_{\mathbb{R}} e^{iz\xi} e^{tA_L(\xi)} \hat{g}_L(0, \xi) d\xi \right)_\ell e_\ell(\theta),$$

where  $(f)_\ell$  is the component of  $f$  along  $e_\ell$ . If  $f_0$  in (11) is the initial condition, then  $g(t=0)$  reads

$$g(0, z, \theta) = \frac{1}{\sqrt{2\pi c^2}} e^{-\frac{z^2}{2c^2}} \frac{1}{2\pi} (1 + \cos \theta) = \frac{1}{\sqrt{2\pi c^2}} e^{-\frac{z^2}{2c^2}} \left( \frac{e_0}{\sqrt{2\pi}} + \frac{e_1}{\sqrt{4\pi}} \right),$$

which admits as Fourier transform

$$\hat{g}(0, \xi, \theta) = e^{-c^2 \xi^2 / 2} \left( \frac{e_0}{\sqrt{2\pi}} + \frac{e_1}{\sqrt{4\pi}} \right).$$

The function  $J_i$  defined by (12) is then finally

$$J_i(t) = \int_{z_1}^{z_2} \int_{\theta_i}^{\theta_{i+1}} g(t, z, \theta) dz d\theta = \frac{1}{2\pi} \sum_{\ell=0}^L \left( \int_{\mathbb{R}} \frac{e^{iz_2 \xi} - e^{iz_1 \xi}}{i\xi} e^{tA_L(\xi)} \hat{g}(0, \xi) d\xi \right)_\ell \int_{\theta_i}^{\theta_{i+1}} e_\ell(\theta) d\theta,$$

which we evaluate numerically.

## References

- [1] S. ASMUSSEN AND J. ROSISKI, *Approximations of small jumps of lvy processes with a view towards simulation*, J. Appl. Probab., 38 (2001), pp. 482–493.
- [2] K. B. ATHREYA, D. MCDONALD, AND P. NEY, *Limit theorems for semi-markov processes and renewal theory for markov chains*, Ann. Probab., 6 (1978), pp. 788–797.
- [3] G. BAL, *Inverse transport theory and applications*, Inverse Problems, 25 (2009).
- [4] G. BAL AND O. PINAUD, *Imaging using transport models for wave-wave correlations*, M3AS, 21(5) (2011), pp. 1071–1093.
- [5] G. BAL AND K. REN, *Transport-based imaging in random media*, SIAM Applied Math., 68(6) (2008), pp. 1738–1762.
- [6] S. COHEN AND J. ROSINSKI, *Gaussian approximation of multivariate lvy processes with applications to simulation of tempered stable processes*, Bernoulli, 13 (2007), pp. 195–210.

- [7] S. DOLAN, C. BEAN, AND R. B., *The broad-band fractal nature of heterogeneity in the upper crust from petrophysical logs*, Geophys. J. Int., 132 (1998), pp. 489–507.
- [8] W. GAUTSCHI, *Orthogonal polynomials: computation and approximation*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2004. Oxford Science Publications.
- [9] W. GAUTSCHI, *Orthogonal polynomials (in matlab)*, Journal of Computational and Applied Mathematics, 178 (2005), pp. 215 – 234. Proceedings of the Seventh International Symposium on Orthogonal Polynomials, Special Functions and Applications.
- [10] F. GOLSE, P.-L. LIONS, B. PERTHAME, AND R. SENTIS, *Regularity of the moments of the solution of a transport equation*, Journal of Functional Analysis 76, (1988), pp. 110–125.
- [11] G. H. GOLUB AND J. H. WELSCH, *Calculation of gauss quadrature rules*, Mathematics of Computation, 23 (1969), pp. 221–s10.
- [12] C. GOMEZ, *Radiative transport limit for the random Schrödinger equation with long-range correlations*, J. Math. Pures Appl. (9), 98 (2012), pp. 295–327.
- [13] C. GOMEZ, O. PINAUD, AND L. RYZHIK, *Radiative transfer with long-range interactions: regularity and asymptotics*, to appear in SIAM MMS.
- [14] ———, *Hypoelliptic estimates in radiative transfer*, CPDE, 1 (2016), pp. 150–184.
- [15] L. A. GRZELAK, J. WITTEVEEN, M. SUAREZ-TABOADA, AND C. W. OOSTERLEE, *The stochastic collocation monte carlo sampler: Highly efficient sampling from 'expensive' distributions*, Available at SSRN: <https://ssrn.com/abstract=2529691> or <http://dx.doi.org/10.2139/ssrn.2529691>, (2015).
- [16] A. HAGHIGHAT, *Monte Carlo Methods for Particle Transport*, CRC Press, 2014.
- [17] W. K. HASTINGS, *Monte carlo sampling methods using markov chains and their applications*, Biometrika, 57 (1970), pp. 97–109.
- [18] S. HOLM AND R. SINKUS, *A unifying fractional wave equation for compressional and shear wave*, J. Acoust. Soc. Am., 1 (2010), pp. 542–548.
- [19] A. ISHIMARU, *Wave Propagation and Scattering in Random Media*, New York: Academic, 1978.
- [20] A. D. KIM AND J. B. KELLER, *Light propagation in biological tissue*, J. Opt. Soc. Am. A, 20 (2003), pp. 92–98.
- [21] B. LAPEYRE, E. PARDOUX, AND R. SENTIS, *Introduction to Monte Carlo Methods for Transport and Diffusion Equations*, Oxford, 1998.
- [22] E. W. LARSEN AND J. B. KELLER, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Math. Phys., 15 (1974), pp. 75–81.
- [23] C. LEAKEAS AND E. LARSEN, *Generalized Fokker-Planck approximations of particle transport with highly forward-peaked scattering*, Nuclear science and engineering, 137 (2001), pp. 236–250.
- [24] N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, AND E. TELLER, *Equation of state calculations by fast computing machines*, The Journal of Chemical Physics, 21 (1953), pp. 1087–1092.
- [25] G. POMRANING, *The fokker-planck operator as an asymptotic limit*, Mathematical Models and Methods in Applied Sciences, 02 (1992), pp. 21–36.
- [26] C. RAO, W. JIANG, AND N. LING, *Spatial and temporal characterization of phase fluctuations in non-kolmogorov atmospheric turbulence*, Journal of Modern Optics, 47 (2000), pp. 1111–1126.
- [27] F. RODDIER, *The effects of atmospheric turbulence in optical astronomy*, Progress in optics. Volume 19. Amsterdam, North-Holland Publishing Co., 19 (1981), pp. 281–376.
- [28] L. RYZHIK, G. PAPANICOLAOU, AND J. B. KELLER, *Transport equations for elastic and other waves in random media*, Wave Motion, 24 (1996), pp. 327–370.

- [29] S. SAMKO, *On inversion of fractional spherical potentials by spherical hypersingular operators*, in Singular integral operators, factorization and applications, vol. 142 of Oper. Theory Adv. Appl., Birkhäuser, Basel, 2003, pp. 357–368.
- [30] G. SAMORODNITSKY AND T. M.S., *Stable Non-Gaussian Processes*, CRC Press, 1994.
- [31] P. SHENG, *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*, Academic Press, New York, 1995.
- [32] C. SIDI AND F. DALAUDIER, *Turbulence in the stratified atmosphere: Recent theoretical developments and experimental results*, Adv. in Space Res., 10 (1990), pp. 25–36.
- [33] J. SPANIER AND E. M. GELBARD, *Monte Carlo principles and neutron transport problems*, Addison-Wesley, Reading, Mass., 1969.
- [34] T. M. SRENSSEN AND F. E. BENTH, *Levy process simulation by stochastic step functions*, SIAM Journal on Scientific Computing, 35 (2013), pp. A2207–A2224.
- [35] R. WERON, *On the chambers-mallows-stuck method for simulating skewed stable random variables*, Statistics and Probability Letters, 28 (1996), pp. 165 – 171.