Hastings-Metropolis Algorithm on Markov Chains for Small-Probability Estimation

François Bachoc, Achref Bachouch, Lionel Lenôtre

ESAIM: PROCEEDINGS, Vol. ?, 2013, 1-10
Editors: Will be set by the publisher

HASTINGS-METROPOLIS ALGORITHM ON MARKOV CHAINS FOR SMALL-PROBABILITY ESTIMATION *, **

F. Bachoc[1, 2], A. Bachouch[3, 4] and L. Lenôtre[5, 6]

Abstract. Shielding studies in neutron transport, with Monte-Carlo codes, yield challenging problems of small-probability estimation. The particularity of these studies is that the small probability to estimate is formulated in terms of the distribution of a Markov chain, instead of that of a random vector in more classical cases. Thus, it is not straightforward to adapt classical statistical methods, for estimating small probabilities involving random vectors, to these neutron-transport problems. A recent interacting-particle method for small-probability estimation, relying on the Hastings-Metropolis algorithm, is presented. It is shown how to adapt the Hastings-Metropolis algorithm when dealing with Markov chains. A convergence result is also shown. Then, the practical implementation of the resulting method for small-probability estimation is treated in details, for an academic one-dimensional problem, and for a two-dimensional shielding study. Finally, it is shown, for these two cases, that the proposed interacting-particle method considerably outperforms a simple-Monte Carlo method, when the probability to estimate is small.

Résumé. Dans les études de protection en neutronique, celles fondées sur des codes Monte-Carlo posent d'importants problèmes d'estimation de faibles probabilités. La particularité de ces études est que les faibles probabilités sont exprimés en termes de lois sur des chaines de Markov, contrairement à des lois sur des vecteurs aléatoires dans les cas les plus classiques. Ainsi, les méthodes classiques d'estimation de faibles probabilités, portant sur des vecteurs aléatoires, ne peuvent s'utiliser telles qu'elles, pour ces problèmes neutroniques. Un méthode récente d'estimation de faibles probabilités, par système de particules en intéraction, reposant sur l'algorithme de Hastings-Metropolis, est présentée. Il est alors montré comment adapter l'algorithme de Hastings-Metropolis au cas des chaines de Markov. Un résultat de convergence est ainsi prouvé. Ensuite, il est expliqué en détail comment appliquer la méthode obtenue à un cas académique à une dimension, et à une étude de protection en neutronique à deux dimensions. Finalement, pour ces deux problèmes, il est montré que la méthode par système de particules en intéraction est considérablement plus efficace qu'une méthode par Monte Carlo classique, lorsque la probabilité à estimer est faible.

---

* *The work presented in this manuscript was carried out in the framework of the REDVAR project of the CEMRACS 2013*

** *This work was financed by the Commissariat à l'energie atomique et aux énergies alternatives*

[1] Department of Statistics and Operations Research, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna

[2] When the work presented in this manuscript was carried out, the author was affiliated to CEA-Saclay, DEN, DM2S, STMF, LGLS, F-91191 Gif-Sur-Yvette, France and to the Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VII

[3] Laboratoire Manceau de Mathématiques (LMM), Université du Maine. Avenue Olivier Messiaen 72085 Le Mans CEDEX 9 . France

[4] Laboratoire de Modélisation Mathématique et numérique dans les sciences de l'ingénieur (Lamsin). ENIT. BP 37, 1002 Tunis. Tunisie

[5] Inria, Research Centre Rennes-Bretagne Atlantique, Campus de Beaulieu 35042 Rennes Cedex France

[6] Université de Rennes 1, Campus de Beaulieu 35042 Rennes Cedex France

1. Introduction

The study of neutronics began in the 40s, when nuclear energy was on the verge of being used both for setting up nuclear devices like bombs and for civil purposes like the production of energy. Neutronics is the study of neutron population in fissile media that can be modeled using the linear Boltzmann equation, also known as the transport equation. More precisely, it can be subdivided in two different sub-domains. On the one hand, criticality studies aim at understanding the neutron population dynamics due to the branching process that mimics fission reaction (see for instance [20] for a recent survey on branching processes in neutronics). On the other hand, when neutrons are propagated through media where fission reactions do not occur, or can safely be neglected, their transport can be modeled by simple exponential flights [21]: indeed, between each collisions, neutrons travel along straight path distributed exponentially.

Among this last category, shielding studies allow to size shielding structures so as to protect humans from ionizing particles, and imply, by definition, the attenuation of initial neutron flux typically by several decades. For instance, the vessel structure of a nuclear reactor core attenuates the thermal neutron flux inside the code by a factor roughly equal to 10^{13}. Many different national nuclear authorities require shielding studies of nuclear systems before giving their agreement for the design of these systems. Examples are reactor cores, but also devices for nuclear medicine (proton-therapy, gamma-therapy, etc). The study of those nuclear systems is complicated by 3-dimensional effects due to the geometry and by non-trivial energetic spectrum that can hardly be modeled.

Since Monte-Carlo transport codes (like MCNP [15], Geant4 [1], Tripoli-4 [8]) require very few hypotheses, they are often used for shielding studies. Nevertheless, those studies represent a long-standing numerical challenge for Monte-Carlo codes in the sense that they schematically require to evaluate the proportion of neutrons that "pass" through the shielding system. This proportion is, by construction, very small. Hence a shielding study by Monte-Carlo code requires to evaluate a small probability, which is the motivation of the present paper.

There is a fair amount of literature on classical techniques for reducing the variance in these small-probability estimation problems for Monte-Carlo codes. Those techniques often rely on a zero-variance scheme [2, 13, 14] adapted to the Boltzmann equation, allied with weight-watching techniques [3]. The particular forms that this scheme takes when concretely developed in various transport codes range from the use of weight windows [5, 13–15], like in MCNP, to the use of the exponential transform [4, 8] like in Tripoli-4. Nowadays, all those techniques have proven to be often limited in view of full-filling the requirements made by national nuclear authorities for the precise measurements of radiation, which standards are progressively strengthen. Thus, new variance reduction techniques have been recently proposed in the literature (see for instance [9] for the use of neural networks for evaluating the importance function).

This paper deals with the application of the recent interacting-particle method developed in [10], for small probability estimation, to a neutronic shielding's Monte Carlo code. The method proposed in [10] have interesting theoretical properties and is particularly efficient in practical cases. Nevertheless, its application to shielding studies with Monte-Carlo codes is not straightforward. Monte-Carlo codes consist in sampling the trajectory of a neutron which can, depending on the complexity of the physical modeling, be the realization of a discrete-time branching process or stochastic process. Indeed, since a neutron travels along straight paths between collisions, there is no loss of information in considering only the characteristics of the collisions (dates, positions, energies, subparticle creations) as random.

In view of simplifying the matter, the subparticle-creation phenomena are not taken into account, neither is the energy dependence. We consider here the simplified but realistic case of a monokinetic particle (constant speed, no offsprings). Then a particle trajectory is just a set of successive collisions and constitutes a Markov chain. Furthermore, with probability one, the particle is absorbed after a finite number of collisions. The small probability we are interested in is the probability that a particle "pass" through a shielding system and reach a domain of interest before absorption.

The method proposed in [10] relies on the Hastings-Metropolis algorithm [12,17] for practical implementation. This algorithm is clearly a textbook method when applied to probability distributions on the Euclidean space.

Nevertheless, we have discussed below that small probability estimation problems in neutronic codes involve Markov Chains instead of random vectors. Thus, it is not automatic to apply the method of [10] to these kind of problems. Our contribution is two-fold. First, we show how the Hastings-Metropolis algorithm can be extended to the case of Markov chains that are absorbed after finite time. Second, we apply the resulting method to an academic one-dimensional case and to a two-dimensional case, representing a monokinetic-particle model in a simplified but realistic shielding system. The smaller the probability to estimate is, the more the interacting-particle method clearly outperforms a simple-Monte Carlo method.

The manuscript is organized as follows. In section 2 we give a reminder of the interacting-particle method [10], and highlight the need of the Hastings-Metropolis algorithm. In section 3, we prove the validity and the convergence of the Hastings-Metropolis algorithm applied to Markov Chains absorbed after a finite time. Then in section 4, we present the one and two-dimensional cases. We give the actual equations for the small probability estimation method[1]. At last in section 5, we present numerical results in the one and two-dimensional cases.

2. Reminder on the interacting-particle method for small probability estimation

We present the interacting-particle method [10] and highlight its need of the Hastings-Metropolis (HM) algorithm for practical application.

We consider a probability space $(\Omega, \mathcal{F}, P)$, and a measurable space $(S, \mathcal{S}, Q)$. We consider a random variable X from $(\Omega, \mathcal{F}, P)$ to $(S, \mathcal{S}, Q)$. We assume that we are able to sample realizations of X.

We consider an objective function $\Phi : S \to \mathbb{R}$, for which we only assume a continuous cumulative distribution function F. The interacting-particle method aims at estimating the probability of the event $\Phi(X) \geq l$, for a given level $l \in \mathbb{R}$. We denote this probability p.

The method can be presented in two steps. First, we assume that an ideal, or a theoretical, method can be implemented exactly. In this case, the finite-sample distribution of the corresponding estimator of the probability p is known exactly, so that exact finite-sample confidence intervals are available. Furthermore, the limit, for large number of sampling from X, of the probability estimation error, has attractive properties as shown in [10]. The ideal method is presented in section 2.1.

Nevertheless, this ideal method can not be implemented exactly for a large range of practical problems. Thus, it is proposed in [10] to approximate the ideal method by using a HM algorithm. This is presented in subsection 2.2.

2.1. The theoretical version of the interacting-particle method

In this section 2.1, we assume that we are able to sample realizations of X, conditionally to the event $\Phi(X) \geq t$, for any $t \in \mathbb{R}$. This is a strong assumption, which is why the corresponding method that we present is called the ideal method.

The ideal algorithm for estimating p is then parameterized by a number of particle N and is as follows.

Algorithm 2.1.

- *Generate an iid sample* $(X_1, .., X_N)$*, from the distribution of* X*, and initialize* $m = 1$*,* $L_1 = \min(\Phi(X_1), .., \Phi(X_N))$ *and* $X_1^1 = X_1, ..., X_N^1 = X_N$*.*
- *While* $L_m \leq l$ *do*
  - *For* $i = 1, ..., N$
    - *Set* $X_i^{m+1} = X_i^m$ *if* $\Phi(X_i^m) > L_m$*, and else* $X_i^{m+1} = X^*$*, where* X^* *follows the distribution of* X *conditionally to* $\Phi(X) \geq L_m$*, and is independent of any other random variables involved in the algorithm.*
  - *Set* $m = m + 1$*.*
  - *Set* $L_m = \min(\Phi(X_1^m), .., \Phi(X_N^m))$*.*

[1]The reader specially interested in the neutronic Monte Carlo application can go directly from section 2 to section 4.

- *The estimate of the probability p is* $\hat{p}_{ipm} = (1 - \frac{1}{N})^{m-1}$.

For each finite N, the ideal estimator $\hat{p}_{ipm}$ obtained from algorithm 2.1 has an explicit distribution that is detailed in [10]. In this paper, we just consider two properties of $\hat{p}_{ipm}$. First, the estimator is unbiased: $\mathbb{E}(\hat{p}_{ipm}) = p$. Second, asymptotic 95% confidence intervals, for N large, are of the form

$$I_{\hat{p}_{ipm}} = \left[\hat{p}_{imc} \exp\left(-1.96 \sqrt{\frac{-\log \hat{p}_{ipm}}{N}} \right), \hat{p}_{imc} \exp\left(1.96 \sqrt{\frac{-\log \hat{p}_{ipm}}{N}} \right) \right]. \tag{1}$$

Finally, we notice that the event $p \in I_{\hat{p}_{ipm}}$ is asymptotically equivalent (for N large) to the event $\hat{p}_{ipm} \in I_p$, with I_P as in (1), with $\hat{p}_{ipm}$ replaced by p. We will use this property in section 5.

2.2. Practical implementation of the interacting-particle method with the Hastings-Metropolis algorithm

For practical implementation of algorithm 2.1, the only problem we have to solve is the conditional sampling, with the distribution of X, conditionally to $\Phi(X) \geq t$, for any $t \in \mathbb{R}$.

An application of the HM algorithm is proposed in [10]. For this, the following is assumed

- The distribution of X has a probability distribution function (pdf) f with respect to $(S, \mathcal{S}, Q)$. For any $x \in S$ we can compute $f(x)$.
- We dispose of a transition kernel on $(S, \mathcal{S}, Q)$ with conditional pdf $\kappa(x, y)$ (pdf of y conditionally to x). We are able to sample from $\kappa(x, .)$ for any $x \in S$ and we can compute $\kappa(x, y)$ for any $x, y \in S$.

Let $t \in \mathbb{R}$ and $x \in S$ so that $\Phi(x) \geq t$. Then, the following algorithm enables to, starting from x, sample approximately with the distribution of X, conditionally to $\Phi(X) \geq t$. The algorithm is parameterized by a number of iterations $T \in \mathbb{N}^*$.

Algorithm 2.2.

- *Let $X = x$.*
- *For $i = 1, ..., T$*
  - *Independently from any other random variable, generate X^* following the $\kappa(X, .)$ distribution.*
  - *If $\Phi(X^*) \geq t$*
    - *Let $r = \frac{f(X^*)\kappa(X^*, X)}{f(X)\kappa(X, X^*)}$.*
    - *With probability $min(r, 1)$, let $X = X^*$.*
- *Return X.*

The random variable returned by algorithm 2.2 is denoted $X_T(x)$.

For consistency, we now give the actual interacting-particle method, involving algorithm 2.2. This method is parameterized by the number of particles N and the number of HM iterations T.

Algorithm 2.3.

- *Generate an iid sample $(X_1, .., X_N)$ from the distribution of X and initialize $m = 1$, $L_1 = \min(\Phi(X_1), .., \Phi(X_N))$ and $X_1^1 = X_1, ..., X_N^1 = X_N$.*
- *While $L_m \leq l$ do*
  - *For $i = 1, ..., N$*
    - *Set $X_i^{m+1} = X_i^m$ if $\Phi(X_i^m) > L_m$, and else pick at random an integer J among the integers $1 \leq j \leq N$ so that $\Phi(X_j^m) > L_m$. Then, let $X_i^{m+1} = X_T(X_J^m)$, with the notation of algorithm 2.2.*
  - *Set $m = m + 1$.*
  - *Set $L_m = \min(\Phi(X_1^m), .., \Phi(X_N^m))$.*
- *The estimate of the probability p is* $\hat{p}_{ipm} = (1 - \frac{1}{N})^{m-1}$.

The estimator $\hat{p}_{ipm}$ of algorithm 2.3 is the practical estimator that we will study in the numerical results of section 5.

In [10], it is shown that, when the space S is a subset of $\mathbb{R}^d$, under mild assumptions, the distribution of the estimator of algorithm 2.3 converges, as $T \to +\infty$, to the distribution of the ideal estimator of algorithm 2.1. For this reason, we call the estimator of algorithm 2.1 the estimator corresponding to the case $T = +\infty$. We also call the confidence intervals (1) the confidence intervals of the case $T = +\infty$.

Nevertheless, as we discussed in section 1, the space S we are interested in is a space of sequences that are killed after a finite time. Thus, it is not straightforward that the convergence, as $T \to +\infty$, discussed above, hold in our case. Furthermore, even the notion of pdf on this space of sequences has to be defined.

This is the object of the next section 3, that defines the notion of pdf, on a space of sequences that are killed after a finite time, and that gives a convergence result for the HM algorithm. The definition of the pdf is also restated in section 4, so that sections 2 and 4 are self-sufficient for the implementation of the small-probability estimation method in the one and two-dimensional cases.

3. An extension of Hastings-Metropolis algorithm to Markov chain sampling

3.1. Introducing Markov Chains Killed Out of a Domain

As discussed in section 1, the dynamic of the collisions is described by a Markov chain $(X_n)_{n\geq 0}$ with values in $\mathbb{R}^d$ and a probability transition function q. The study of the detection probability for the collisions occurs only in a restricted area. We decide to censor the model and redefine it to obtain accurate theoretical results.

Let D be an open bounded subset of $\mathbb{R}^d$ and ∂D be its $\mathcal{C}^\infty$ boundary. D constitutes the domain of interest. We modify the transition function $(X_n)_{n\geq 0}$ as follows

$$k(x, dy) = (q(x, y)1_D(y)\, dy + p(x, D^C)\delta_\Delta(dy))1_D(x) + \delta_\Delta(dy)1_\Delta(x)$$

where Δ is a cemetery point and

$$p(x, D^C) = \int_{D^C} q(x, y)\, dy.$$

This representation describes the following dynamic:

- while $(X_n)_{n\geq 0}$ is inside D, it behaves with a transition kernel that could push it outside D. We use the already define q which reflects the collision dynamic.
- when $(X_n)_{n\geq 0}$ hits D^C, it is killed and send to the cemetery point Δ where it stays. This way we censor the neutrons and keep only the collisions inside D.

We call this model Markov Chain Killed Out of a Domain, MCKOD for short.

3.2. Formulation of the Hastings-Metropolis algorithm

For self-sufficiency of section 3, we start by giving a formulation of the Hastings-Metropolis Algorithm [12,17] for sampling a distribution γ that admits a density versus a measure Π. The main idea is to define a Markov chain $(Y_n)_{n\geq 0}$ that converges in distribution to γ. To construct $(Y_n)_{n\geq 0}$, we use an instrumental Markov chain $(Z_n)_{n\geq 0}$ and an acceptation-rejection function r. We denote by κ the probability transition function of $(Z_n)_{n\geq 0}$ and Γ the transition kernel of $(Y_n)_{n\geq 0}$. We assume that κ and Γ admit a density versus the same measure Π. Step by step, the algorithm is:

- Introduce a starting point x and use it to sample a transition y of $(Z_n)_{n\geq 0}$.
- Accept or reject this transition using r.
- If the sample is accepted redo the procedure with y , else redo with the starting point.

With enough repetitions of this procedure, the distribution of y is approximately γ. Following [12,19], a formula of Γ is

$$\Gamma(u, dv) = \overline{\kappa}(u, v)\Pi(dv) + \overline{r}(u)\delta_u(dv)$$

where

$$\overline{\kappa}(u,v) = \begin{cases} \kappa(u,v)r(u,v), & \text{if } x \neq y, \\ 0, & \text{if } x = y, \end{cases}$$

and

$$\overline{r}(u) = 1 - \int p(u,v)\Pi(dv).$$

For the acceptation-rejection function, following [19], we choose

$$r(u,v) = \begin{cases} \min\left\{\dfrac{\gamma(v)\kappa(v,u)}{\gamma(u)\kappa(u,v)}, 1\right\}, & \text{if } \gamma(u)\kappa(u,v) > 0 \\ 1, & \text{if } \gamma(u)\kappa(u,v) = 0 \end{cases}.$$

This is the case presented in algorithm 2.2. It ensures the reversibility condition required in [19] to prove that Π is invariant for Γ. In those books, a state space endowed with a countably generated σ-algebra and a good topology is a basic condition for most of the definitions and results. A priori, our MCKOD do not stick with that. In addition, we need to introduce a measure Π. Thus, section 3.3 is devoted to solve these two points. In addition, according to [19], the kernel Γ is irreducible only if the instrumental kernel κ is irreducible. Then, section 3.4 defines a family of instrumental and prove their irreducibility. Finally section 3.5 gives the proof using the results in [16], [18] and [19].

3.3. **State space, distribution and density for MCKOD**

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $q(x,y)$ a probability transition kernel on $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d)$ with density w.r.t. the Lebesgue measure. Let $D \subset \mathbb{R}^d$ be an open bounded subset with $\mathcal{C}^\infty$ boundary ∂D. We suppose that $0 \notin D$ and is the cemetery point. We set $D_0 = D \cup \{0\}$. We introduce the space

$$c_0 = \{(u_n)_{n\geq 0} \in D_0^{\mathbb{N}} : \exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_n = 0\},$$

equipped with the distance

$$d_\infty(u,v) = \max_{n\geq 0} d(u_n - v_n)_{\mathbb{R}^d}$$

and denote by $\mathcal{B}(c_0)$ the Borelian σ-algebra. We want to see the Markov chain $(X_n)_{n\geq 0}$ as a random variable:

$$\begin{array}{rcl} X : (\Omega, \mathcal{F}, \mathbb{P}) & \mapsto & (c_0, \mathcal{B}(c_0)) \\ \omega & \mapsto & (X_n(\omega))_{n\geq 0}. \end{array}$$

A tricky point is the measurability with $\mathcal{B}(c_0)$. But, we have the measurability with

$$\overline{\mathcal{F}} = \bigvee_{i=0}^{+\infty} \mathcal{B}(D_0).$$

Thus, the following result give it:

Proposition 3.1. *Let $\overline{\mathcal{F}}_{c_0}$ be the restriction of $\overline{\mathcal{F}}$ to the subspace c_0. Then,*

$$\overline{\mathcal{F}}_{c_o} = \mathcal{B}(c_0)$$

Proof. Let p_n be the projection from c_0 in D_0 who associates u_n to u. This application is Lipschitz. In fact, let u and v be in c_0, we have $d(u_n - v_n) \leq d(u - v)$. Consequently, every projection is measurable and $\overline{\mathcal{F}}_{c_o} \subset \mathcal{B}(c_0)$.

Let show the inclusion $\mathcal{B}(c_0) \subset \overline{\mathcal{F}}_{c_0}$. We know that (c_0, d_∞) is separable. We denote by S a dense subset. $\mathcal{B}(c_0)$ is generated by the ball of radius $\rho \in \mathbb{Q} \cap D_0$ and center point $u \in S$. Thus, it is enough to show that the ball $B(\rho, u)$ is in $\overline{\mathcal{F}}_{c_0}$. So, we write

$$B(\rho, u) = \bigcap_{n=0}^{+\infty} \{v \in c_0, d(u_n - v_n) \leq \rho\}$$

and because each member of this intersection is in $\overline{\mathcal{F}}_{c_0}$, we have the desired result. □

Let define

$$A_n = \{u \in c_0 : u_k \in D, \forall k < n \text{ and } u_k = 0, \forall k \geq n\}.$$

The family $(A_n)_{n\geq 0}$ is a partition of the space c_0. We introduce the family of projections $(\pi_n)_{n\geq 0}$, with:

$$\begin{aligned} \pi_n : A_n &\mapsto D^n \\ u &\mapsto (u_1, \cdots, u_n), \end{aligned}$$

and the measure on c_0

$$\Pi(dx) = \sum_{n=1}^{+\infty} \lambda_n(\pi_n^{-1}(dx)) 1_{A_n}(x),$$

where λ_n is the Lebesgue measure on D^n. We have the following result:

Proposition 3.2. *The distribution γ of a MCKOD $(X_n)_{n\geq 0}$ starting from α is absolutely continuous versus Π. The density is given by:*

$$\sum_{n=1}^{+\infty} \int_{D^C} q_D^n(x, x_n) q(x_n, x_{n+1})\, dx_{n+1}\, 1_{A_n}(x)$$

where

$$q_D^n(x, y) = q(\gamma, x_1) \cdots q(x_{n-1}, x_n)\, 1_D(x_1) \cdots 1_D(x_n).$$

Proof. We fix $n > 0$ and we restrain Π to A_n. Suppose that $A \in \mathcal{B}(c_0)$ with $\Pi(A) = 0$, we have

$$\begin{aligned} \gamma(A \cap A_n) &= \mathbb{P}((X_1, \cdots, X_n) \in p_n(A \cap A_n), X_{n+1} \in D^C) \\ &= \mathbb{P}((X_1, \cdots, X_n) \in p_n(A \cap A_n)) \times p(X_n, 0) \\ &= \int_{\pi_n(A \cap A_n)} q^n(\gamma, x_n)\, dx_n \times p(X_n, 0) \end{aligned}$$

$q^n(x, y)$ is absolutely continuous versus λ_n and $\lambda_n(A \cap A_n) = \Pi_{|A_n}(A) = 0$. Using the classical formula:

$$\mathbb{P}_X(A) = \sum_{n=0}^{+\infty} \mathbb{P}_X(A \cap A_n) = 0,$$

we conclude that $\mathbb{P}_X(A) = 0$. For the density, the result comes from the expression

$$\begin{aligned} \mathbb{P}(T = n+1) &= \int_{\pi_n(A \cap A_n)} q^n(x, y)\, dy \times p(X_n, 0) \\ &= \int_{D^C} q_D^n(\gamma, x_n) q(x_n, x_{n+1})\, dx_n dx_{n+1} \\ &= \int_{D^C} \int q(\gamma, x_1) \cdots q(x_{n-1}, x_n)\, 1_D(x_1) \\ &\quad \cdots 1_D(x_n)\, q(x_n, x_{n+1})\, dx_1 \cdots dx_n\, dx_{n+1}, \end{aligned}$$

where T is the first time X hits D^C. Thus, the mass is one. □

3.4. Some Π-irreducible instrumental kernels on c_0

We introduce and study the Markov chain on c_0 starting from w with kernel:

$$\kappa(u, dv) = \sum_{k=1}^{+\infty} \Theta(u, A_k)\nu_k(u, dv)1_{A_k}(v)$$

where we suppose

- For each $u \in c_0$, the sum of the $(\Theta(u, A_k))_{k\geq 0}$ is one.
- $(\nu_k(u, dv))_{k\geq 1}$ is a family of probability transition kernels on D^n.

This statements insures that κ is a probability transition kernel on c_0. We describe the behavior of the chain:

- from a sequence u, we change the number of non-null points using the family $(\Theta(u, A_k))_{k\geq 0}$.
- for each non-null point, we change its position using a probability transition kernel from $(\nu_k(u, dv))_{k\geq 1}$.

This kind of Markov chains will serve as the so-called instrumental chains for the Hastings-Metropolis algorithm. We give the following definition for the irreducibility of a chain:

Definition 3.3. Let G be a topological space, $\mathcal{G}$ a σ-algebra on G, m a probability measure and μ a probability transition kernel. We say that $A \in \mathcal{G}$ is attainable from $x \in G$ if:

$$\mu^n(u, A) > 0 \text{ for some } n \geq 1.$$

We say that the set $B \in \mathcal{G}$ is m-communicating if:

$$\forall x \in B, \forall A \in \mathcal{G} \text{ such that } A \subset B,\ m(A) > 0, A \text{ is attainable from } x.$$

In addition, if G is m-communicating, the chain is m-irreducible

The γ-irreducibility of the instrumental kernel is required for the irreducibility of the Metropolis chain and the convergence of the algorithm. Thus, the following results are crucial and permit us to have usable kernels.

Proposition 3.4. *If κ is a probability transition kernel satisfying*

- *For every $u \in c_0$ and $k \geq 0$, $\Theta(u, A_k) > 0$.*
- *for each $k \geq 1$, $\nu_k(u, dv)$ is absolutely continuous w.r.t. the Lebesgue measure in D^n.*
- *for each $k \geq 1$, $\nu_k(u, dv)$ is irreducible for the Lebesgue measure in D^n.*

then

- *κ is absolutely continuous versus Π.*
- *κ is Π-irreducible.*

Proof. The absolute continuity is induced by the definition and the third hypothesis. The proof is totally analog to 3.2. Thus, we only show the Π-irreducibility. Let $A \in \mathcal{B}(c_0)$ be a Π-positive subset and $u \in c_0$ a sequence. We want to prove that:

$$\kappa^n(u, A) > 0, \text{ for some } n \geq 1.$$

This result naturally holds if we show for all $k \geq 0$ and for all $A \subset A_k$ that A is attainable from $u \in c_0$. We fix $k \geq 0$ and choose $A \subset A_k$. From the definition of κ, we have

$$\kappa(u, A) = \int_A \Theta(u, A_k)\nu_k(u, dw)\, dw.$$

Thus, we only have to prove that

$$\int_A \nu_k(u, dw)\, dw > 0.$$

The absolute continuity and the irreducibility of the $(\nu_k(u,dv))_{k\geq 0}$ induce that

$$\text{for every } k\geq 0 \text{ and } \lambda_n\text{-positive set A, } \nu_k^m(u,A)>0 \text{ for } m=1.$$

Indeed, suppose the opposite, for each $m\geq 1$ we have

$$\nu_k^m(u,A)=\int_D\cdots\int_D\int_A \nu_k(u,v_1)\cdots\nu_k(u,v_m)\,dv_1\cdots dv_m=0$$

and we have a conflict with the irreducibility. Finally,

$$\int_A \nu_k(u,y)dy>0$$

and the result is proved. □

Corollary 3.5. *With the same hypothesis as in 3.4, κ is γ-irreducible.*

Proof. A result of [18] says that: if a kernel κ is Π-irreducible and there is a measure γ which is absolutely continuous versus Π, then κ is γ-irreducible. □

3.5. **Convergence of the extended Hastings-Metropolis Algorithm**

Before proving the convergence of the algorithm, we give an example of $(\Theta(u,A_k))_{k\geq 0}$ and $(\nu_k(u,dv))_{k\geq 1}$ for which the result hold. G denotes the geometric distribution adapted to $\mathbb{N}$. For each $u\in c_0$, we choose it for $\Theta(u,A_k)=\mathbb{P}(G=k)$. Let g denote the density of the uniform distribution inside D. We take

$$\nu_k(u,v)=\prod_{i=1}^{k} g(v)$$

and one could see that it satisfies the hypothesis of the result below.

Proposition 3.6. *If κ is probability transition kernel satisfying satisfying the same hypothesis as in 3.4 and*

- *for every $k\geq 0$,*
$$\inf_{u\in c_0}\Theta(u,A_k)>0.$$
- *for each $k\geq 1$,*
$$\inf_{u\in D}\nu_k(u,dv)>0.$$

Then, the Hastings-Metropolis kernel Γ converge to γ.

Proof. In order to prove the convergence, we follow [19]. Consequently, we have to show that Γ is γ-irreducible and $\gamma\{\overline{r}(u)>0\}>0$. We start with the γ-irreducibility. Let $A\in\mathcal{B}(c_0)$ be a γ-positive subset and $u\in c_0$ a sequence. We want

$$\Gamma^n(u,A)>0, \text{ for some } n\geq 1.$$

We use the same approach as in the proof of proposition 3.4. For fixed $k\geq 0$ and $A\subset A_k$, we study

$$\overline{\kappa}(u,A)=\int_A \kappa(u,w)r(u,w)\Pi(dw).$$

Since the second term in the expression of Γ is positive, the γ-irreducibility of $\overline{\kappa}$ is fairly enough. We recall that

$$r(u,v)=\begin{cases}\min\left\{\dfrac{\gamma(w)\kappa(w,u)}{\gamma(u)\kappa(u,w)},1\right\}, & \text{if } \gamma(u)\kappa(u,w)>0\\ 1, & \text{if } \gamma(u)\kappa(u,w)=0\end{cases}.$$

If $\gamma(u)\Theta(u,A_k)\nu_k(u,w)=0$ on A γ-almost-surely, then the proof is trivial with the γ-irreducibility of κ. This is the same if

$$\frac{\gamma(w)\kappa(w,u)}{\gamma(u)\kappa(u,w)}>1$$

on A γ-almost-surely. Thus, we have to check that

$$\overline{\kappa}(u,A)=\int_A \kappa(u,w)\cdot\frac{\gamma(w)\kappa(w,u)}{\gamma(u)\kappa(u,w)}\Pi(dw)>0.$$

Suppose that $u\in A_l$, then

$$\begin{aligned}\overline{\kappa}(u,A)&=\int_A \frac{\gamma_{A_k}(w)\Theta(w,A_l)\nu_l(w,u)}{\gamma(u)}\,dw_1\cdots dw_k\\&=\int_A \frac{\gamma_{A_k}(w)\Theta(w,A_l)\nu_l(w,u)}{\gamma_{A_l}(u)}\,dw_1\cdots dw_k.\end{aligned}$$

Because $\gamma_{A_k}(u)$ is strictly positive on A γ-almost-surely, otherwise we are back to the previous case. With the two lower-bound hypotheses, the problem is reduced to

$$\int_A \gamma_{A_k}(w)\,dw_1\cdots dw_k>0,$$

which is trivially true. For the aperiodicity, the probability to stay inside a A_k is positive. If we take $A\subset A_k$ such that $\gamma_{A_k}(A)$, the probability to reach it is positive, since v_k is γ_{A_k}-irreducible (see hypotheses in 3.4 and corollary 3.5). □

4. Practical implementation of the one and two-dimensional cases

In section 4 we present the one and two-dimensional cases, for which the results of the interacting-particle method of section 2 are presented in section 5.

The interacting-particle method 2.3 necessitates, as we have seen, to evaluate pdf on Markov-chain trajectories with finite number of non-absorbed points. These pdf have been defined in section 3. They are redefined in definition 4.8 and proposition 4.9 so that section 4 is self-sufficient.

The actual values of these pdf, for the one and two-dimensional cases, are also given in section 4.

4.1. **Some general vocabulary and notation**

Throughout section 4, we consider a monokinetic particle (a particle with constant speed and yielding no subparticle birth) evolving in $\mathbb{R}^d$, with $d=1,2$. This particle is created at the source point $s\in\mathbb{R}^d$, that is the birth of the particle takes place at s.

Similarly to the description in section 1, the trajectory of the monokinetic particle is characterized by its collision points, which constitute a Markov chain. The sequence of collision points is written $(X_n)_{n\in\mathbb{N}^*}$. The birth of the particle takes place at s, that is $X_0=s$. After its birth, the monokinetic particle travels along straight lines, with random distances and directions, between its collision points, until it is absorbed. The absorption happens almost surely after a finite number of collisions.

The distribution of the Markov chain of the birth and collision points $(X_n)_{n\in\mathbb{N}}$ is characterized by, first, a function $P_a(t):\mathbb{R}^d\to[0,1]$, so that $P_a(t)$ is the probability of absorption for a collision taking place at t. Second, the distribution is characterized by the pdf $q(t_1,t_2):\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^+$, so that $q(t1,.):\mathbb{R}^d\to\mathbb{R}^+$ is a pdf.

The behavior of the monokinetic particle is then as follows. At a collision point X_n, for which the monokinetic particle has not been absorbed yet, one and only one of the two following events can randomly happen. First,

the particle can be absorbed during the collision. In this case, we define this absorption by the equality $X_m = \Delta$ for any $m > n$. The point $\Delta \in \mathbb{R}^d$ symbolizes that the monokinetic particle has been absorbed. It is called the cemetery point, similarly to section 3. The choice of the cemetery point Δ is of course arbitrary, as long as it is distinguished from the source point, that is $\Delta \neq s$. If the monokinetic particle is not absorbed during the collision, it is scattered. In this case, the position of the next collision point X_{n+1} has the $q(X_n, .)$ pdf.

Thus, eventually, the sequence $(X_n)_{n\in\mathbb{N}}$ of collisions points of the monokinetic particle is a Markov chain that has the property that, almost surely, there exists $m \in \mathbb{N}$ so that $X_n = \Delta$ for $n \geq m$. This is the type of Markov Chains that are covered in section 3. We say that the monokinetic particle is active at time n, or at X_n, or before collision n, if $X_n \neq \Delta$.

Finally, when the Markov Chain of a monokinetic particle $(X_i)_{i\in\mathbb{N}}$ has been sampled, and is equal to $(x_i)_{i\in\mathbb{N}}$ we call the sequence $(x_i)_{i\in\mathbb{N}^*}$ of its collision points a trajectory of the monokinetic-particle. (There is no loss of information in that a trajectory does not store the deterministic source point $x_0 = s$.)

Remark 4.1. Strictly speaking, it is possible that a collision n takes place at position Δ, that is $X_n = \Delta$, though the monokinetic particle has not been absorbed before X_n. Nevertheless, in this section 4, we only consider transition kernels $q(., .)$ that are absolutely continuous with respect to the Lebesgue measure. Thus, almost-surely, the cemetery point Δ identifies without ambiguity that the monokinetic-particle has been absorbed.

4.2. Description of the one-dimensional case and expression of the probability density functions

In this section 4.2, we consider an academic Monte-Carlo problem similar to the shielding studies described in section 1, but for which a monokinetic particle evolves in a one-dimensional space.

4.2.1. *A one-dimensional random walk*

We consider that the monokinetic particle evolves in $\mathbb{R}$. As described in section 4.1, the birth of the particle takes place at 0, that is $X_0 = s = 0$. In the one-dimensional model we define, when the monokinetic particle has not been absorbed, the signed-distance traveled between two collisions is a Gaussian variable. That is, if $X_n \neq \Delta$, we have $X_{n+1} = X_n + \epsilon_{n+1}$, where the $(\epsilon_i)_{i\in\mathbb{N}^*}$ are *iid* and follow a $\mathcal{N}(0, \sigma^2_{mk})$ distribution.

Finally, at each collision point X_n, the probability of absorption is one if $X_n \leq L_{inf}$ or $X_n \geq L_{sup}$, where $L_{inf} < 0 < L_{sup}$. If a collision point is $L_{inf} < X_n < L_{sup}$, the monokinetic particle is absorbed with probability $0 \leq P_a < 0$, and is scattered with probability $1 - P_a$.

The event of interest is here that the monokinetic particle reaches the domain $(-\infty, L_{inf}]$. When using the interacting-particle method of section 2, this event is traduced by the event $\Phi(x) \geq 0$, with $\Phi(x) = L_{inf} - \inf_{i\in\mathbb{N}^*; x_i\neq\Delta} x_i$. Notice that, almost-surely, the infimum is over a finite number of points.

The one-dimensional case presented here reproduces some key features of the shielding studies by Monte Carlo code described in sections 1 and 4.1. Indeed, the monokinetic particle travels a random distance, toward a random direction (positive or negative), between two collision points, and random absorption is considered. The particle is absorbed after a number of collisions that is random and almost-surely finite. By setting P_a sufficiently large, and L_{inf} sufficiently away from 0, we will see that we can tackle problems with arbitrary-small probabilities. Thus, in section 5, we will consider a probability small enough so that the interacting-particle method of section 2 outperforms a simple-Monte Carlo method.

Finally, notice that an important feature of the two-dimensional case of section 4.3, that is not reproduced by the one-dimensional case, is the presence of different media, and the medium-crossing phenomena.

4.2.2. *Expression of the probability density function of a trajectory*

We now give the expression of the pdf (with respect to the setting of definition 4.8 and proposition 4.9) of a trajectory obtained from the one-dimensional model above. We let $(x_i)_{i\in\mathbb{N}^*}$ be the sequence of collision points (the trajectory) of a monokinetic particle. We let $\mathcal{D} = (L_{inf}, L_{sup})$. We denote $\phi_{m,\sigma^2}(t)$ the pdf at t of the one-dimensional Gaussian distribution with mean m and variance σ^2.

Similarly to section 3, we define $A_n = \{(x_i)_{i\in\mathbb{N}^*}, x_j \neq \Delta, \forall 1 \leq j \leq n-1, x_k = \Delta, \forall k \geq n\}$, that is the set of trajectories that are absorbed at collision point $n-1$ (so that they are in the absorbed state from collision point n and onward).

Proposition 4.2. *The pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 4.8 and proposition 4.9, of a trajectory $(x_n)_{n\in\mathbb{N}^*}$, sampled from the procedure of section 4.2.1, is $f(x) = \sum_{n\in\mathbb{N}^*} \mathbf{1}_{A_{n+1}}(x) f_n(x)$, with*

$$f_n(x) = \prod_{i=1}^{n-1} \left(\mathbf{1}_{x_i\in\mathcal{D}} \phi_{x_{i-1},\sigma^2_{mk}}(x_i)(1-P_a) \right) \phi_{x_{n-1},\sigma^2_{mk}}(x_n) \left(\mathbf{1}_{x_n\notin\mathcal{D}} + P_a \mathbf{1}_{x_n\in\mathcal{D}} \right),$$

where $x_0 = 0$ by convention.

Several comments can be made on proposition 4.2.

The pdf of proposition 4.2 has to be evaluated for each trajectory, either sampled from its initial distribution, or from a perturbation method in the HM algorithm 2.2. The perturbation methods are presented below in sections 4.2.3, 4.2.4, 4.2.5 and 4.2.6.

There is a mild difference between proposition 3.2 and proposition 4.2. In proposition 3.2, when the monokinetic particle goes outside the domain of interest D, and consequently is absorbed, the collision point in the exterior of D is not stored in the trajectory of the monokinetic particle. Indeed, the exact value of this point is not needed to assess if the monokinetic has reached the domain $(-\infty, L_{inf}]$. However, by not storing this point, the pdf of proposition 3.2 necessitates to know the probability that, starting from a birth or scattering point in the domain D, the next collision point lies outside D. This probability is not known explicitly in the framework of section 4.3, and *a fortiori* in shielding studies involving more complex Monte Carlo codes. To avoid evaluating this probability numerically each time a pdf of a trajectory is computed, we store the collision points outside the domain D. Notice that this, inevitably, add some variance in the HM method, because we use a source of randomness (the exact collision point at which the monokinetic particle leaves D) that does not impact the event of interest.

The evaluation of a pdf like that of proposition 4.2 is an intrusive operation on a Monte Carlo code. Indeed, it necessitates to know all the random-quantity sampling that are done when this code samples a monokinetic-particle trajectory. Thus, the Monte Carlo code is not used as a black box.

Nevertheless, the computational cost of the pdf evaluation is of the same order as the computational cost of a trajectory sampling, and the same kind of operations are involved. Namely, both tasks require a loop which length is the number of collisions made by the monokinetic-particle before its absorption. Furthermore, for each random quantity that is sampled for a trajectory sampling, the pdf evaluation requires to compute the corresponding pdf. For example, in the case of proposition 4.2, when a trajectory sampling requires to sample n Gaussian variables and n or $n-1$ Bernoulli variables, the trajectory-pdf evaluation requires to compute the corresponding Gaussian pdf and Bernoulli probabilities.

The discussion above holds similarly for the two-dimensional case of section 4.3.

4.2.3. *Description of the trajectory perturbation method when $P_a = 0$*

For clarity of exposition, we present first the perturbation method when $P_a = 0$. In this case, the monokinetic particle is a random walk on $\mathbb{R}$, that is absorbed once it goes outside $\mathcal{D}$.

The perturbation method is parameterized by $\sigma^2_{hm} > 0$. Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. Then, the set of birth and collision points of the perturbed monokinetic-particle is an inhomogeneous Markov chain $(Y_i)_{i\in\mathbb{N}}$. This inhomogeneous Markov chain is so that $Y_0 = 0$. Then, if $i \leq n-1$, and if the perturbed monokinetic particle is still in $\mathcal{D}$ at collision point i, we have $Y_{i+1} = Y_i + \epsilon_{i+1}$, where the $(\epsilon_i)_{1\leq i\leq n}$ are independent and where ϵ_i follows a $\mathcal{N}(x_i - x_{i-1}, \sigma^2_{hm})$ distribution.

Similarly to the initial sampling, the perturbed monokinetic particle is absorbed at the first collision point outside $\mathcal{D}$. If the collision point Y_n of the perturbed monokinetic particle is in $\mathcal{D}$ (contrary to x_n for the initial trajectory), the sequel of the trajectory of the perturbed monokinetic particle is sampled as the initial monokinetic particle would be sampled if its collision point n was Y_n.

This conditional sampling method for perturbed trajectories is intrusive: it necessitates to change the stochastic dynamic of the monokinetic particle. Nevertheless, the new dynamic is here chosen as to have the same cost as the unconditional sampling, and to require the same type of computations. This is similar to the discussion following proposition 4.2.

4.2.4. *Expression of the probability density function of a perturbed trajectory when* $P_a = 0$

We now give the expression of the conditional pdf (with respect to the setting of proposition 3.2) of a trajectory obtained from the one-dimensional perturbation method above.

Proposition 4.3. *Let us consider an historical trajectory* $(x_i)_{i\in\mathbb{N}^*}$*, absorbed at collision* n*. The conditional pdf, with respect to* $(c_0, \mathcal{S}, \Pi)$ *of definition 4.8 and proposition 4.9, of a trajectory* $(y_n)_{n\in\mathbb{N}^*}$ *sampled from the procedure of section 4.2.3, is* $\kappa(x,y) = \sum_{m\in\mathbb{N}^*} \mathbf{1}_{A_{m+1}}(y) f_{n,m}(x,y)$ *where, if* $m \leq n$

$$f_{n,m}(x,y) = \prod_{i=1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}}\phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)\right) \mathbf{1}_{y_m\notin\mathcal{D}}\phi_{y_{m-1}+(x_m-x_{m-1}),\sigma_{hm}^2}(y_m),$$

and if $m > n$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n} \left(\mathbf{1}_{y_i\in\mathcal{D}}\phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)\right) \\ &\quad \times \prod_{i=n+1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}}\phi_{y_{i-1},\sigma_{mk}^2}(y_i)\right) \\ &\quad \mathbf{1}_{y_m\notin\mathcal{D}}\phi_{y_{m-1},\sigma_{mk}^2}(y_m), \end{aligned}$$

where $y_0 = 0$ *by convention.*

Similarly to the discussion following 4.2, the computation of the conditional pdf of a perturbed trajectory has the same computational cost as the sampling of this perturbed trajectory.

4.2.5. *Description of the trajectory perturbation method when* $P_a > 0$

Let us now consider the general case where $P_a > 0$.

The perturbation method is parameterized by $\sigma_{hm}^2 > 0$ and $0 < P_c < 1$. Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n.

As when $P_a = 0$, the set of birth and collision points of the perturbed monokinetic-particle is an inhomogeneous Markov chain $(Y_i)_{i\in\mathbb{N}}$, so that $Y_0 = 0$. As when $P_a = 0$, we modify the increments of the initial trajectory, and, if the perturbed trajectory outsurvives the initial one, we generate the sequel with the initial distribution. Specifically to this case $P_a > 0$, we perturb the absorption/non-absorption sampling by changing the initial values with probability P_c.

More precisely, for $i \leq n-1$, and if the perturbed monokinetic particle has not been absorbed before collision point i, it is absorbed with probability $max(P_c, \mathbf{1}_{Y_i\notin D})$. If it is scattered instead, we have $Y_{i+1} = Y_i + \epsilon_{i+1}$, where the $(\epsilon_i)_{1\leq i\leq n}$ are independent and where ϵ_i follows a $\mathcal{N}(x_i - x_{i-1}, \sigma_{hm}^2)$ distribution. If the perturbed monokinetic particle has not been absorbed before collision point n, then it is absorbed if $Y_n \notin \mathcal{D}$. If $Y_n \in \mathcal{D}$, the perturbed monokinetic particle is absorbed with probability p, where $p = 1 - P_c$ if $x_n \in \mathcal{D}$ and $p = P_a$ if $x_n \notin \mathcal{D}$.

As when $P_a = 0$, if the perturbed monokinetic particle has not been absorbed before collision point Y_n, the sequel of the trajectory of the perturbed monokinetic particle is sampled as the initial particle would be sampled if its collision point n was Y_n.

The idea is that, by selecting the difference between P_c and $\min(P_a, 1 - P_a)$, the closeness between the perturbed and initial trajectories can be specified, from the point of view of the absorption/non-absorption events.

4.2.6. *Expression of the probability density function of a perturbed trajectory when $P_a > 0$*

Proposition 4.4. *Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. Let $y_0 = x_0 = 0$ by convention. The conditional pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 4.8 and proposition 4.9, of a trajectory $(y_n)_{n\in\mathbb{N}^*}$ sampled from the procedure of section 4.2.5, is $\kappa(x,y) = \sum_{m\in\mathbb{N}^*} \mathbf{1}_{A_{m+1}}(y) f_{n,m}(x,y)$ where, if $m \leq n-1$,*

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)(1-P_c)\right) \\ &\quad \phi_{y_{m-1}+(x_m-x_{m-1}),\sigma_{hm}^2}(y_m)\left(\mathbf{1}_{y_m\in\mathcal{D}} P_c + \mathbf{1}_{y_m\notin\mathcal{D}}\right), \end{aligned}$$

if $m = n$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)(1-P_c)\right) \\ &\quad \phi_{y_{n-1}+(x_n-x_{n-1}),\sigma_{hm}^2}(y_n)\left(\mathbf{1}_{y_n\notin\mathcal{D}} + \mathbf{1}_{y_n\in\mathcal{D}}\mathbf{1}_{x_n\in\mathcal{D}}(1-P_c) + \mathbf{1}_{y_n\in\mathcal{D}}\mathbf{1}_{x_n\notin\mathcal{D}}P_a\right), \end{aligned}$$

and if $m \geq n+1$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1}+(x_i-x_{i-1}),\sigma_{hm}^2}(y_i)(1-P_c)\right) \\ &\quad \mathbf{1}_{y_n\in\mathcal{D}} \phi_{y_{n-1}+(x_n-x_{n-1}),\sigma_{hm}^2}(y_n)\left(\mathbf{1}_{x_n\in\mathcal{D}} P_c + \mathbf{1}_{x_n\notin\mathcal{D}}(1-P_a)\right), \\ &\quad \prod_{i=n+1}^{m-1} \left(\mathbf{1}_{y_i\in\mathcal{D}} \phi_{y_{i-1},\sigma_{mk}^2}(y_i)(1-P_a)\right) \\ &\quad \phi_{y_{m-1},\sigma_{mk}^2}(y_m)\left(\mathbf{1}_{y_m\notin\mathcal{D}} + \mathbf{1}_{y_m\in\mathcal{D}}P_a\right). \end{aligned}$$

4.3. Description of the two-dimensional case and expression of the probability density functions

4.3.1. *Description of the neutron transport problem*

The monokinetic particle evolves in $\mathbb{R}^2$, and its birth takes place at the source point $s = (-p_{s,x}, 0)$, with $p_{s,x} > 0$. The domain of interest is a box $[-\frac{L}{2}, \frac{L}{2}]^2$ with $p_{s,x} < \frac{L}{2}$, in which there is an obstacle sphere with center 0 and radius l, with $l < \frac{L}{2}$. The obstacle sphere is the set $\left\{x \in \mathbb{R}^2 | |x| \leq l\right\}$, with $|x|$ the Euclidean norm of $x \in \mathbb{R}^2$.

The box is composed of two media. The obstacle sphere is composed of "poison" and the rest of the box is composed of "water". The exterior of the box is also composed of "water". Nevertheless, if the monokinetic-particle reaches this exterior, it is considered to have gone too far away, and subsequently is is absorbed at the first collision point in the exterior of the box.

We consider a detector, in the box, that is a sphere with center $(p_{det,x}, 0)$ and radius l_d. The detector is the set $\left\{x \in \mathbb{R}^2 | |x - (p_{det,x}, 0)| \leq l_d\right\}$. The detector is in the exterior of the obstacle sphere, that is $l < p_{det,x} - l_d$. The event of interest is that the monokinetic particle makes a collision in the detector, before being absorbed. The cemetery point Δ is, consequently, just chosen to be different from s and in the exterior of the detector.

Let $(x_i)_{i\in\mathbb{N}^*}$ be a trajectory of the monokinetic particle. When using the interacting-particle method of section 2, the event of interest is traduced by the event $\Phi(x) \geq 0$, with $\Phi(x) = l_d - \inf_{i\in\mathbb{N}^*; x_i\neq\Delta} |x_i - (p_{det,x}, 0)^t|$.

We now discuss the probabilities of absorption. If a collision takes place outside the box, then we have discussed that the monokinetic particle has left the domain of interest. With respect to the probability of

reaching the detector, this is equivalent to writing that the probability of absorption outside the box is one. This is what we do in the sequel. Then, the probability of absorption in the box, but outside the obstacle sphere, is written $P_{a,w}$ and the probability of absorption in the obstacle sphere is written $P_{a,p}$, with $P_{a,w} \leq P_{a,p}$.

Finally, let us discuss the distribution of the jumps between collision points. Following the neutron-transport models discussed in section 1, after a scattering, or birth, at X_n, of the monokinetic-particle, the direction toward which the monokinetic particle travels has isotropic distribution. This direction is here denoted u, with u an unit two-dimensional vector. Then, the sampling of the distance to the next collision point X_{n+1} is as follows. First, the distance τ is sampled from an exponential distribution with rate λ_w, if X_n is in the medium "water", or $\lambda_p > \lambda_w$ if X_n is in the medium "poison". Then, two case are possible. First, if the sampled distance is so that the monokinetic particle stays in the same medium while it travels this distance, then the next collision point is $X_{n+1} = X_n + \tau u$. Second, if between X_n and $X_n + \tau u$, there is a change of medium, then the monokinetic particle is virtually stopped at the first medium-change point between X_n and $X_n + \tau u$. At this point, the travel direction remains the same, but the remaining distance to travel is resampled, from the exponential distribution with the rate corresponding to the new medium. These resampling are iterated each time a sampled distance causes a medium-change. The new collision point X_{n+1} is the point reached by the first sampled distance that does not cause a medium change. Notice that, in this precise setting with two media, the maximum number of distance sampling between two collision points is three. This can happen in the case where the collision point X_n is in the box but not in the obstacle sphere, where the sampled direction points toward the obstacle sphere, and where toward this direction, the monokinetic particle enters and leaves the obstacle sphere.

The actual pdf, corresponding to the medium-change process described, of a collision point X_{n+1}, conditionally to a collision point X_n, is given in proposition 4.5.

Notice that the setting described does constitute a model for a shielding system in neutron transport. The source point corresponds to a neutron production area. This neutron production area is separated from a sensible area, modeled by the detector. The shielding system is constituted first by the obstacle sphere, which is placed between the source and the detector and has the largest probability of absorption $P_{a,p}$. Second, the standard "water" medium also constitutes a milder protection, because it also has a probability of absorption $P_{a,w}$.

We are interested in evaluating the number of monokinetic particles that reach the detector. Since the number of monokinetic-particles produced at the source point is approximately known, the problem is to evaluate the probability that a monokinetic-particle produced at the source reach the detector.

4.3.2. *Expression of the probability density function of a trajectory*

We first set some notations for the two-dimensional problem presented in section 4.3.1. We write B as the box $[-\frac{L}{2}, \frac{L}{2}]^2$. We write B_{ext} as the exterior of the box, $B_{ext} = \mathbb{R}^2 \backslash B$. The obstacle sphere is denoted S, with $S = \left\{x \in \mathbb{R}^2 | |x| \leq l\right\}$.

We write $|x|$ as the Euclidean norm of $x \in \mathbb{R}^2$. We write $[v, w]$ as the segment between two points $v, w \in \mathbb{R}^2$.

Consider two points $v, w \in \mathbb{R}^2$ so that v is strictly in the interior of S ($|v| < l$) and w is strictly in the exterior of S ($|v| > l$). Then $c_S(v, w)$ is defined as the unique point in the boundary of S that belongs to $[v, w]$.

Similarly, for $v, w \in \mathbb{R}^2 \backslash S$ and when $[v, w]$ has a non-empty intersection with S, we denote by $c_{S_1}(v, w)$ and $c_{S_2}(v, w)$ the two intersection points between $[v, w]$ and the boundary of S. The indexes 1 and 2 are so that $|v - c_{S,1}(v, w)| \leq |v - c_{S,2}(v, w)|$.

For $v, w \in \mathbb{R}^2 \backslash S$, we let $I_S(v, w)$ be equal to 1 if $[v, w]$ has a non-empty intersection with S, and 0 otherwise.

The computation of $c_s(v, w)$, $I_S(v, w)$, $c_{S,1}(v, w)$, and $c_{S,2}(v, w)$ are equally needed for a monokinetic-particle simulation, and for the computation of the corresponding pdf of proposition 4.6. The four quantities can be computed explicitly.

We now give the pdf of the collision point X_{n+1}, conditionally to a scattering or a birth point X_n.

Proposition 4.5. *Consider a scattering, or birth, point $x_n \in B$. Then, the pdf of the collision point X_{n+1}, conditionally to x_n, is denoted $q(x_n, x_{n+1})$ and is given by, if $x_n \in B\backslash S$*

$$\begin{aligned} q(x_n, x_{n+1}) &= \frac{1}{2\pi|x_n - x_{n+1}|}\mathbf{1}_{x_{n+1}\in\mathbb{R}^3\backslash S}(1 - I_S(x_n, x_{n+1}))\lambda_w e^{-\lambda_w|x_n - x_{n+1}|} \\ &+ \frac{1}{2\pi|x_n - x_{n+1}|}\mathbf{1}_{x_{n+1}\in\mathbb{R}^3\backslash S} I_S(x_n, x_{n+1}) e^{-\lambda_w|x_n - c_{S_1}(x_n,x_{n+1})|} \\ &\quad e^{-\lambda_p|c_{S_1}(x_n,x_{n+1}) - c_{S_2}(x_n,x_{n+1})|}\lambda_w e^{-\lambda_w|c_{S_2}(x_n,x_{n+1}) - x_{n+1}|} \\ &+ \frac{1}{2\pi|x_n - x_{n+1}|}\mathbf{1}_{x_{n+1}\in S} e^{-\lambda_w|x_n - c_S(x_n,x_{n+1})|}\lambda_p e^{-\lambda_p|c_S(x_n,x_{n+1}) - x_{n+1}|} \end{aligned} \tag{2}$$

and, if $x_n \in S$,

$$\begin{aligned} q(x_n, x_{n+1}) &= \frac{1}{2\pi|x_n - x_{n+1}|}\mathbf{1}_{x_{n+1}\in S}\lambda_p e^{-\lambda_p|x_n - x_{n+1}|} \\ &+ \frac{1}{2\pi|x_n - x_{n+1}|}\mathbf{1}_{x_{n+1}\in\mathbb{R}^3\backslash S} e^{-\lambda_p|x_n - c_S(x_n,x_{n+1})|}\lambda_w e^{-\lambda_w|c_S(x_n,x_{n+1}) - x_{n+1}|}. \end{aligned} \tag{3}$$

Proof. The proposition is obtained by using the properties of the exponential distribution, the definitions of $c_s(x_n, x_{n+1})$, $I_S(x_n, x_{n+1})$, $c_{S,1}(x_n, x_{n+1})$, and $c_{S,2}(x_n, x_{n+1})$ and a two-dimensional polar change of variables. The proof is straightforward but burdensome. □

Using proposition 4.5 above, we now give the pdf of the monokinetic-particle trajectories obtained from the sampling procedure of section 4.3.1.

Proposition 4.6. *The pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 4.8 and proposition 4.9, of a trajectory $(x_n)_{n\in\mathbb{N}^*}$, sampled from the procedure of section 4.3.1, is $f(x) = \sum_{n\in\mathbb{N}^*}\mathbf{1}_{A_{n+1}}(x)f_n(x)$, with*

$$f_n(x) = \prod_{i=1}^{n-1}\left(q(x_{i-1}, x_i)\left[\mathbf{1}_{x_i\in B\backslash S}(1 - P_{a,w}) + \mathbf{1}_{x_i\in S}(1 - P_{a,p})\right]\right) q(x_{n-1}, x_n)\left[\mathbf{1}_{x_n\notin\mathcal{B}} + \mathbf{1}_{x_n\in B\backslash S}P_{a,w} + \mathbf{1}_{x_n\in S}P_{a,p}\right],$$

where $x_0 = 0$ by convention, and with $q(x_{i-1}, x_i)$ and $q(x_{n-1}, x_n)$ as in proposition 4.5.

4.3.3. *Description of the trajectory perturbation method*

The perturbation method is parameterized by $\sigma_{hm}^2 > 0$, $0 < P_{c,w} < 1$ and $0 < P_{c,p} < 1$.

Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n.

As in section 4.2, the set of birth and collision points of the perturbed monokinetic-particle is an inhomogeneous Markov chain $(Y_i)_{i\in\mathbb{N}}$, so that $Y_0 = 0$. We modify independently the collision points of the initial trajectory, and, if the perturbed trajectory outsurvive the initial one, we generate the sequel with the initial distribution. Similarly to section 4.2.5, we perturb the absorption/non-absorption sampling by changing the initial values with probabilities $P_{c,w}$ and $P_{c,p}$, if the initial and perturbed collision points are both in $B\backslash S$ or both in S. If this is not the case, we sample the absorption/non-absorption for the perturbed monokinetic-particle with the initial probabilities $P_{a,w}$ and $P_{a,p}$.

More precisely, for $i \leq n-1$, and if the perturbed monokinetic particle has not been absorbed before collision point Y_i, it is absorbed at collision point Y_i with probability $P(x_i, Y_i)$ with

$$P(x_i, Y_i) = \begin{cases} 1 & \text{if } Y_i \in \mathbb{R}^2\backslash B \\ P_{a,w} & \text{if } Y_i \in B\backslash S \text{ and } x_i \in S \\ P_{a,p} & \text{if } Y_i \in S \text{ and } x_i \in B\backslash S \\ P_{c,w} & \text{if } Y_i \in B\backslash S \text{ and } x_i \in B\backslash S \\ P_{c,p} & \text{if } Y_i \in S \text{ and } x_i \in S \end{cases}. \tag{4}$$

Similarly to the one-dimensional case, by taking $P_{a,w}$ smaller than $\min(P_{a,w}, 1-P_{a,w})$, and $P_{a,p}$ smaller than $\min(P_{a,p}, 1-P_{a,p})$, we can modify rather mildly the initial trajectories.

If the perturbed monokinetic particle is not absorbed at collision point Y_i, its next collision point is $Y_{i+1} = x_{i+1} + \epsilon_{i+1}$, where the $(\epsilon_i)_{1\leq i\leq n}$ are independent and where ϵ_i follows a $\mathcal{N}(0, \sigma_{hm}^2 I_2)$ distribution, where I_2 is the 2×2 identity matrix. If the perturbed monokinetic particle has not been absorbed before collision point Y_n, then it is absorbed with probability $P(x_n, Y_n)$ given by

$$P(x_n, Y_n) = \begin{cases} 1 & \text{if } Y_n \in \mathbb{R}^2\backslash B \\ P_{a,w} & \text{if } Y_n \in B\backslash S \text{ and } x_n \in S \\ P_{a,w} & \text{if } Y_n \in B\backslash S \text{ and } x_n \in \mathbb{R}^2\backslash B \\ P_{a,p} & \text{if } Y_n \in S \text{ and } x_n \in B\backslash S \\ P_{a,p} & \text{if } Y_n \in S \text{ and } x_n \in \mathbb{R}^2\backslash B \\ 1-P_{c,w} & \text{if } Y_n \in B\backslash S \text{ and } x_n \in B\backslash S \\ 1-P_{c,p} & \text{if } Y_n \in S \text{ and } x_n \in S \end{cases}. \tag{5}$$

As in section 4.2, if the perturbed monokinetic particle has not been absorbed before collision point Y_n, the sequel of the trajectory of the perturbed monokinetic particle is sampled as the initial particle would be sampled if its collision point n was Y_n.

4.3.4. *Expression of the probability density function of a perturbed trajectory*

Proposition 4.7. *Let us consider an historical trajectory $(x_i)_{i\in\mathbb{N}^*}$, absorbed at collision n. Let $y_0 = x_0 = 0$ by convention. The conditional pdf, with respect to $(c_0, \mathcal{S}, \Pi)$ of definition 4.8 and proposition 4.9, of a trajectory $(y_n)_{n\in\mathbb{N}^*}$ sampled from the procedure of section 4.3.3, is $\kappa(x,y) = \sum_{m\in\mathbb{N}^*} \mathbf{1}_{A_{m+1}}(y) f_{n,m}(x,y)$ where the $f_{n,m}$ are given by the following. If $m \leq n-1$,*

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{m-1} \left(\phi_{x_i,\sigma_{hm}^2 I_2}(y_i)\left[1-P(x_i,y_i)\right]\right) \\ &\quad \phi_{x_m,\sigma_{hm}^2 I_2}(y_m) P(x_m, y_m), \end{aligned}$$

with $P(x_i,y_i)$ and $P(x_m,y_m)$ as in (4). If $m = n$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\phi_{x_i,\sigma_{hm}^2 I_2}(y_i)\left[1-P(x_i,y_i)\right]\right) \\ &\quad \phi_{x_n,\sigma_{hm}^2 I_2}(y_n) P(x_n, y_n), \end{aligned}$$

with $P(x_i,y_i)$ as in (4) and $P(x_n,y_n)$ as in (5). If $m \geq n+1$,

$$\begin{aligned} f_{n,m}(x,y) &= \prod_{i=1}^{n-1} \left(\phi_{x_i,\sigma_{hm}^2 I_2}(y_i)\left[1-P(x_i,y_i)\right]\right) \\ &\quad \phi_{x_n,\sigma_{hm}^2 I_2}(y_n)\left[1-P(x_n,y_n)\right] \\ &\quad \prod_{i=n+1}^{m-1} \left(q(y_{i-1},y_i)\left[\mathbf{1}_{y_i\in B\backslash S}(1-P_{a,w}) + \mathbf{1}_{y_i\in S}(1-P_{a,p})\right]\right) \\ &\quad q(y_{m-1},y_m)\left[\mathbf{1}_{y_m\notin B} + \mathbf{1}_{y_m\in B\backslash S} P_{a,w} + \mathbf{1}_{y_m\in S} P_{a,p}\right], \end{aligned}$$

with $P(x_i,y_i)$ as in (4), $P(x_n,y_n)$ as in (5) and $q(y_{i-1},y_i)$ and $q(y_{m-1},y_m)$ as in proposition 4.5.

4.4. **Proofs for section 4**

For the proofs for section 4, we first define the space of the monokinetic-particle trajectories, and the corresponding σ-algebra and measure, in definition 4.8 and proposition 4.9. These definitions are similar to those given in section 3. They are stated here so that section 4 is self-sufficient.

Definition 4.8. Define

$$c_0 = \{(u_n)_{n\geq 1} \in \left(\mathbb{R}^d\right)^{\mathbb{N}^*} : \exists n_0 \in \mathbb{N}^*, \forall n \geq n_0, u_n = 0\}.$$

We define $\mathcal{S}$ as the smallest sigma-algebra on c_0 containing the sets $\{x | x_1 \in A_1, ..., x_n \in A_n\}$, for $n \in \mathbb{N}^*$ and $A_i \in \mathcal{B}\left(\mathbb{R}^d\right)$, where $\mathcal{B}\left(\mathbb{R}^d\right)$ is the Borel sigma-algebra on $\mathbb{R}^d$.

Proposition 4.9. *There exists a unique measure Π on $(c_0, \mathcal{S})$ that verifies the following relation, for any $E_n = \{x | x \in A_{n+1}, (x_1, ..., x_n) \in A_1 \times ... \times A_n\}$, with $A_1, ..., A_n \in \mathcal{B}\left(\mathbb{R}^d\right)$ and $n \in \mathbb{N}^*$.*

$$\Pi(E_n) = \lambda(A_1 \times ... \times A_n), \tag{6}$$

with λ the Lebesgue measure.

Proof. The proof is carried out in the same way as in section 3. Alternatively, the Carathéodory extension theorem can be used. □

We now give the following general proposition 4.10, for the expression of pdf for inhomogeneous Markov chains that are absorbed in finite-time.

Proposition 4.10. *Consider a sequence of measurable applications $a_n : \mathbb{R}^d \to [0, 1]$, $n \in \mathbb{N}$, with $a_0 = 0$. Consider a sequence $(q_n)_{n\in\mathbb{N}^*}$ of conditional pdf, that is to say $\forall n, y_{n-1}$, $p_n(y_{n-1}, y_n)$ is a pdf on $\mathbb{R}^d$ with respect to y_n.*

Consider a probability space $(\Omega, \mathcal{F}, P)$. Consider a Markov Chain on $(\Omega, \mathcal{F}, P)$, $(Y_n)_{n\in\mathbb{N}}$, so that first $Y_0 = y_o$ a.s, when y_0 in a non-zero constant of $\mathbb{R}^d$. Second, Y_n has the non-homogeneous transition kernel defined by

$$k(y_{n-1}, dy_n) = \mathbf{1}_{y_{n-1}=0}\delta_0(dy_n) + \mathbf{1}_{y_{n-1}\neq 0}\left\{a_n(y_{n-1})\delta_0(dy_n) + \left[1 - a_n(y_{n-1})\right] q_n(y_{n-1}, y_n)dy_n\right\}. \tag{7}$$

Assume finally that, almost surely, the Markov Chain Y_n reaches 0 after a finite time. Then, the application $\omega \to (Y_i(\omega))_{i\in\mathbb{N}^*}$ is a random variable on $(c_0, \mathcal{S}, \Pi)$ (see definition 4.8 and proposition 4.9), with probability density function, for $y = (y_i)_{i\in\mathbb{N}^*}$, $f(y) = \sum_{n=1}^{+\infty} \mathbf{1}_{A_{n+1}}(y) f_n(y)$, with

$$f_n(y) = \prod_{i=1}^{n} \left[(1 - a_{i-1}(y_{i-1}))q_i(y_{i-1}, y_i)\right] a_n(y_n),$$

where s_0 is the constant value of Y_0 by convention.

Proof. Proposition 4.10 is proved in the same way as proposition 3.2. □

Let us now comment proposition 4.10.

Similarly to proposition 3.2, the value 0 corresponds to the cemetery point, symbolizing the absorption of the Markov chain. Thus, the dynamic (7) is that, at each time n, if the Markov chain with value Y_n is absorbed, it stays absorbed. If it is not absorbed, then is is absorbed with probability $a_n(Y_n)$, thus a probability both depending on position and time. This is consistent with the sampling of perturbed trajectories in sections 4.2.5 and 4.3.3, where these probabilities of absorption depend on the initial trajectories, and thus depend on time and position. If the Markov chain is not absorbed at Y_n, then its next value Y_{n+1} has, conditionally to Y_n, the pdf $q_{n+1}(Y_n, y_{n+1})$. This conditional distribution depends on time, as is the case in the perturbation procedures of sections 4.2.3, 4.2.5 and 4.3.3.

Hence, proposition 4.10 can be applied to calculate the pdf of initial and perturbed trajectories in propositions 4.2, 4.3, 4.4, 4.6 and 4.7.

Proof of proposition 4.2. We apply proposition 4.10 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \in D} P_a + \mathbf{1}_{y_i \notin D}$$

and

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1}, \sigma^2_{mk}}(y_i).$$

□

Proof of proposition 4.3. We denote $x = (x_i)_{i \in \mathbb{N}^*}$ the initial trajectory, so that $x \in A_{n+1}$, and $x_0 = 0$ by convention. We apply proposition 4.10 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \notin D},$$

for $i \geq 1$,

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1} + x_i - x_{i-1}, \sigma^2_{hm}}(y_i),$$

for $1 \leq i \leq n$ and

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1}, \sigma^2_{mk}}(y_i),$$

for $i \geq n+1$. □

Proof of proposition 4.4. We denote $x = (x_i)_{i \in \mathbb{N}^*}$ the initial trajectory, so that $x \in A_{n+1}$, and $x_0 = 0$ by convention. We apply proposition 4.10 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \in D} P_c + \mathbf{1}_{y_i \notin D},$$

for $1 \leq i \leq n-1$,

$$a_n(y_n) = \mathbf{1}_{y_i \in D} \left(\mathbf{1}_{x_i \in D}(1 - P_c) + \mathbf{1}_{x_i \notin D} P_a \right) + \mathbf{1}_{y_i \notin D},$$

$$a_i(y_i) = \mathbf{1}_{y_i \in D} P_a + \mathbf{1}_{y_i \notin D},$$

for $i \geq n+1$,

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1} + x_i - x_{i-1}, \sigma^2_{hm}}(y_i),$$

for $1 \leq i \leq n$ and

$$q_i(y_{i-1}, y_i) = \phi_{y_{i-1}, \sigma^2_{mk}}(y_i),$$

for $i \geq n+1$. □

Proof of proposition 4.6. We apply proposition 4.10 with

$$a_0(y_0) = 0,$$

$$a_i(y_i) = \mathbf{1}_{y_i \in S} P_{a,p} + \mathbf{1}_{y_i \in B \setminus S} P_{a,w} + \mathbf{1}_{y_i \in \mathbb{R}^2 \setminus B}$$

and

$$q_i(y_{i-1}, y_i) = q(y_{i-1}, y_i),$$

with $q(y_{i-1}, y_i)$ as in proposition 4.5. □

Proof of proposition 4.7. We denote $x = (x_i)_{i\in\mathbb{N}^*}$ the initial trajectory, so that $x \in A_{n+1}$, and $x_0 = 0$ by convention. We apply proposition 4.10 with

$$a_0(y_0) = 0,$$
$$a_i(y_i) = P(x_i, y_i),$$

for $1 \leq i \leq n-1$ and with $P(x_i, y_i)$ as in (4),

$$a_n(y_n) = P(x_n, y_n),$$

with $P(x_n, y_n)$ as in (5),

$$a_i(y_i) = \mathbf{1}_{y_i\in S} P_{a,p} + \mathbf{1}_{y_i\in B\setminus S} P_{a,w} + \mathbf{1}_{y_i\in\mathbb{R}^2\setminus B},$$

for $i \geq n+1$

$$q_i(y_{i-1}, y_i) = \phi_{x_i,\sigma_{hm}^2 I_2}(y_i),$$

for $1 \leq i \leq n$ and

$$q_i(y_{i-1}, y_i) = q(y_{i-1}, y_i),$$

for $i \geq n+1$ and with $q(y_{i-1}, y_i)$ as in proposition 4.5. □

5. Numerical results in dimension one and two

In this section 5, we present numerical results for the interacting-particle method of section 2, in the one and two-dimensional cases of section 4. We follow a double objective. First we aim at investigating to what extent the ideal results of the interacting-particle method hold (in term of bias and of theoretical confidence intervals). Second, we want to confirm that, when the objective probability is small, the method outperforms a simple-Monte Carlo method.

The simple-Monte Carlo method is parameterized by a number of Monte Carlo sample n_{mc}. It consists in generating n_{mc} independent trajectories $x_1, ..., x_{n_{mc}}$ and in estimating p by the empirical proportion of these trajectories that verify the small-probability event. We denote by $\hat{p}_{mc}$ the simple-Monte Carlo estimator of p.

5.1. **Numerical results in dimension one**

5.1.1. *Features of the interacting-particle method*

We first present a simple one-dimensional setting, with no-absorption ($P_a = 0$). We set for the domain $L_{inf} = -10$, $L_{sup} = 1$, and for the variance of the increments $\sigma_{mk}^2 = 1$. As a result, the probability p to estimate is not small. It is easily estimated to be $p = 0.13$ by the simple-Monte Carlo method.

For the perturbation method, we set $\sigma_{hm}^2 = 0.1^2$. This choice may not be optimal, but it is reasonable and can be considered as typical for the implementation of the interacting-particle method in this one-dimensional case.

The results we obtain for 100 independent estimations for the interacting-particle method are regrouped in figure 1. We have used $N = 200$ particles and $T = 300$ and $T = 30$ iterations in the HM algorithm 2.2. Let us first interpret the results for $T = 300$ iterations. In this case, we observe that the estimator is empirically non-biased. Furthermore, we also plot the theoretical 95% confidence intervals for the ideal estimator with $T = +\infty$, that are approximately (for N large) $I_p = \left[p\exp\left(-1.96\sqrt{\left(\frac{-\log p}{N}\right)}\right), p\exp\left(1.96\sqrt{\left(\frac{-\log p}{N}\right)}\right)\right]$. We also recall from the discussion after (1) that the events $\hat{p}_{ipm} \in I_p$ and $p \in I_{\hat{p}_{ipm}}$ are approximately equivalent when N is large. Hence the coverage probability of I_p for $\hat{p}_{ipm}$ is approximately the probability that $I_{\hat{p}_{ipm}}$ contains p, which is the practical quantity of interest. We see on figure 1 that I_p approximately matches the empirical distribution of the estimator $\hat{p}_{ipm}$. The overall conclusion of this case $T = 300$ is that there is a good agreement between theory and practice. This emphasizes the validity of using the interacting-particle method of algorithm 2.3, involving the HM algorithm, in a space that is not a subset of $\mathbb{R}^d$.

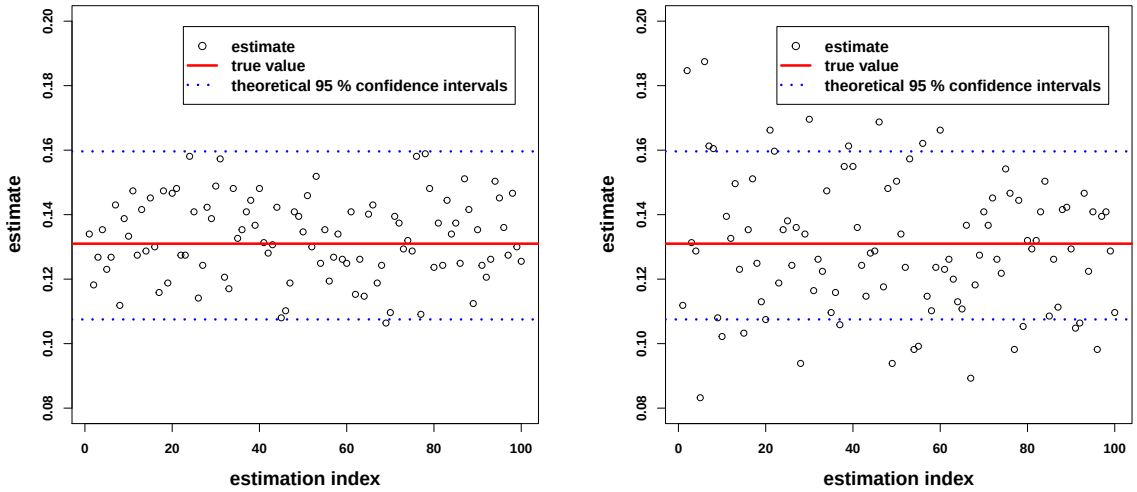

FIGURE 1. One-dimensional case. Plot of 100 independently estimated probabilities with the interacting-particle method 2.3, for number of particles $N = 200$, and number of iterations in the HM algorithm 2.2 $T = 300$ (left) and $T = 30$ (right). We also plot the theoretical 95% confidence intervals (1) of the case $T = +\infty$. The true probability $p = 0.13$ is evaluated quasi-exactly by a simple-Monte Carlo method. In both cases, the interacting-particle estimator is empirically unbiased. For $T = 300$, the theoretical confidence interval, obtained in the case $T = +\infty$ is adapted to the practical estimator. For $T = 30$ however, the estimator has more variance that the ideal estimator $T = +\infty$ has.

In figure 1, we also consider the case $T = 30$. The estimator is still empirically unbiased. However, its empirical variance is larger, so that the theoretical 95% confidence interval I_p is non negligibly too thin. This can be interpreted, because when T is small, a new particle at a given conditional sampling step of algorithm 2.3 is not independent of the $N - 1$ particles that have been kept. Thus, one can argue that, at each step of algorithm 2.3, the overall set of N particles has more interdependence, so that eventually the estimator has more variance. Nevertheless, on the other hand, an estimation with $T = 30$ is 10 times less time-consuming that an estimation with $T = 300$. We further discuss this trade-off problem in section 5.3.

Finally, for this case of a probability that is not small, we have used simple Monte Carlo as a mean to estimate it quasi-exactly. We have found that the interacting-particle method 2.3 requires more computation time than the Monte Carlo method, for reaching the same accuracy. We do not elaborate on this fact, since we especially expect the interacting-particle-method to be competitive for estimating a small probability. This is the object of section 5.1.2. For this case of a probability that is not small, we have just investigated the features of the interacting-particle method.

5.1.2. *Comparison with simple Monte Carlo in a small-probability case*

We now consider a case with possible absorption of the monokinetic particle. Thus we set $P_a = 0.45$. We keep the same values $\sigma^2_{mk} = 1$ and $L_{sup} = 1$ as in section 5.1.1, but we set $L_{inf} = -15$. As a result of these parameters for the monokinetic-particle transition kernel, the probability of interest is small. In fact, we have not estimated it with negligible uncertainty. With a simple-Monte Carlo estimation of sample size 10^9, the probability estimate is $\hat{p}_{vlmc} = 6.6 \times 10^{-8}$. We call this estimate the very large Monte Carlo (VLMC) estimate. Given that the number of successes in this estimate is 66, which is not very large, we are reluctant to use the Central Limit Theorem approximation for computing 95% confidence intervals. Instead, we use the

Clopper-Pearson interval [6], for which the actual coverage probability is always larger than 95%. This 95% confidence interval is there equal to $[5.1\times10^{-8}, 8.4\times10^{-8}]$. This uncertainty is small enough for the conclusions we will draw from this case. Finally, notice that this very large Monte Carlo estimate is not a benchmark for the interacting-particle method, because it is much more time consuming.

For the interacting-particle method, we set $N = 200$ particles, and for the HM algorithm, we set $T = 300$ iterations. We use $\sigma^2_{hm} = 0.1^2$ and $P_c = 0.2$ for the perturbation method. We still denote $\hat{p}_{ipm}$ the obtained estimator for p. We consider a third estimator, that we denote p_{mc} and that consists in the simple-Monte Carlo estimator with sample size 5×10^6. This sample size is appropriate to compare the efficiency of the interacting-particle and Monte Carlo method, as we will show below.

The first criterion for comparing the two estimators $\hat{p}_{ipm}$ and $\hat{p}_{mc}$ is their computation time. We have two possible ways to make this comparison. First, we can evaluate the complexities of the two methods. The Monte Carlo method requires to perform 5×10^6 monokinetic-particle simulations. For each proposed perturbation, the interacting-particle method requires to sample one perturbed trajectory, and to compute its unconditional and conditional pdf. This has to be done approximately $T \times \frac{\log \hat{p}_{vlmc}}{\log(1-\frac{1}{N})} \approx 10^6$ times. Thus, from this point of view, the costs of the two methods have the same orders of magnitude. We can not give a more precise comparison, since the trajectories sampled by the two methods do not necessarily have the same length in the mean sense. Furthermore, it is not obvious to compare the computational cost of an initial sampling, with the costs of a conditional sampling and pdf computations.

Hence, we will just compare the computational costs of the two methods by considering their actual computational times, for the implementation we have used. Averaged over all the estimations, the time for the interacting-particle method is 58% of the time for the Monte Carlo method. Hence, we confirm that the computational costs are of the same order of magnitude, the comparison being nevertheless beneficial to the interacting-particle method.

We now compare the accuracy of the two methods for estimating the true probability p. On figure 2, we plot the results of 100 independent estimations for $\hat{p}_{ipm}$ and 50 independent estimations for $\hat{p}_{mc}$. It appears clearly that the interacting-particle method is more precise in this small probability case. Especially, consider the empirical Root Mean Square Error criterion, for n independent estimates $\hat{p}^1, ..., \hat{p}^n$, for any estimator $\hat{p}$ of p: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^n (p-\hat{p}^i)^2}$. Regardless of the value of p in the very large Monte Carlo 95% confidence interval $[5.1\times10^{-8}, 8.4\times10^{-8}]$, the RMSE is smaller for $\hat{p}_{ipm}$ than for $\hat{p}_{mc}$. If we assume $p = \hat{p}_{vlmc}$, then the RMSE is 10^{-7} for $\hat{p}_{mc}$ and 2×10^{-8} for $\hat{p}_{ipm}$.

A comparison ratio for $\hat{p}_{ipm}$ and $\hat{p}_{mc}$, taking into account both computational time and estimation accuracy (in line with the efficiency in [11]), is the quality ratio defined by $\frac{\sqrt{TIME_{mc}}\times RMSE_{mc}}{\sqrt{TIME_{ipm}}\times RMSE_{ipm}}$, where the four notations $TIME_{mc}$, $TIME_{ipm}$, $RMSE_{mc}$ and $RMSE_{ipm}$ are self-explanatory. This ratio is 6.7 here. This is interpreted as: if the two estimation methods were set as to require the same computational time, then the interacting-particle method would be 6.7 times as accurate (in term of RMSE) as the Monte Carlo method.

Notice that, if we had done the comparison from the point of view of the relative estimation errors, instead of the absolute errors, it would have been even more beneficial to the interacting-particle method. Indeed, assuming again $p = \hat{p}_{vlmv} = 6.6\times10^{-8}$ for discussion, the interacting-particle method does a maximum relative error of 250%. On the other hand, the Monte Carlo estimator takes only 3 different values in figure 2. When it takes value $\frac{2}{5\times10^6}$ it does a relative error of 600%, when it takes value $\frac{1}{5\times10^6}$ it does a relative error of 300% and when it takes value 0 it does an infinite relative error. Alternatively, we can also say that, in the majority of the cases, the Monte Carlo estimator does not see any realization of the rare event, so that it can provide only an overly-conservative upper-bound for p.

5.2. **Numerical results in dimension two**

5.2.1. *Features of the interacting-particle method*

We now present the numerical results for the two-dimensional case. We set the absorption probability in the background $P_{a,w} = 0.2$, the absorption probability in the obstacle sphere $P_{a,p} = 0.5$, the dimension of the box

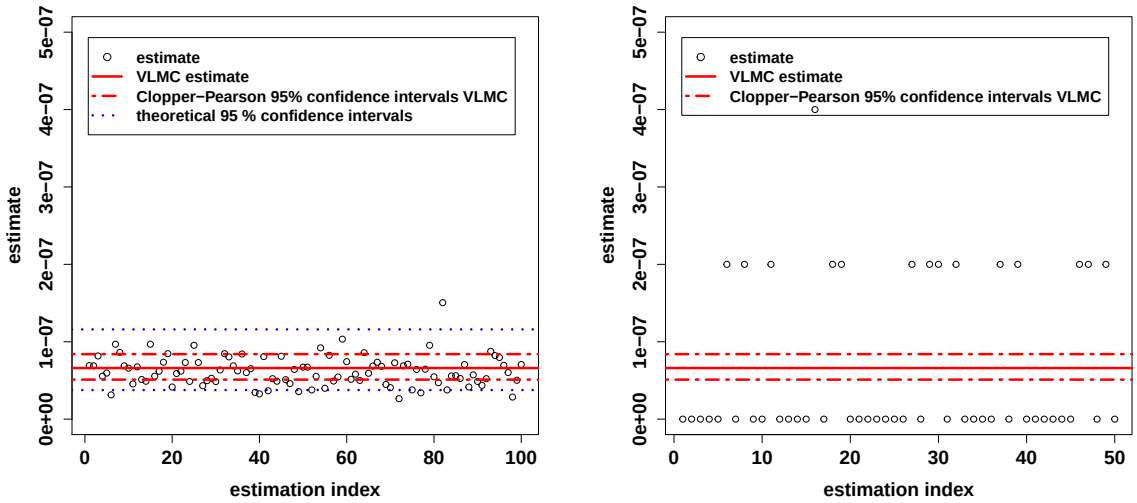

FIGURE 2. One-dimensional case. Plot of 100 independently estimated probabilities with the interacting-particle method 2.3 for $N = 200$ and $T = 300$ (left) and 50 independently estimated probabilities with the Monte Carlo method with sample size 5×10^6 (right). We plot a very large Monte Carlo estimate of the true probability $\hat{p}_{vlmc} = 6.6 \times 10^{-8}$, together with the associated Clopper-Pearson 95% confidence intervals. For the interacting-particle method, we also plot the theoretical 95% confidence intervals of the case $T = +\infty$, assuming the true probability is the VLMC estimate. The uncertainty on the VLMC estimate of the true value p of the probability is small enough for our conclusions to hold. These conclusions are that the interacting-particle method outperforms the Monte Carlo method (with sample size 5×10^6), both in term of computation time and of accuracy.

$[-\frac{L}{2}, -\frac{L}{2}]$ $L = 10$, the radius of the obstacle sphere $l = 2$, the radius of the detector $l_d = 0.5$. The positions of the detector and the source are given by $p_{det,x} = p_{s,x} = 3$. We set the rate of collisions in the water medium $\lambda_w = 0.2$ and in the poison medium $\lambda_p = 2$. As a result, the probability is $p = 2 \times 10^{-4}$ and is evaluated quasi-exactly by a Monte Carlo sampling, similarly to section 5.1.1.

This value is not very small, so that we do not compare the interacting-particle method with the Monte Carlo method. We just aim at showing that the interacting-particle method is valid in this two-dimensional setting. Indeed, this setting has many features that are representative of shielding studies with Monte Carlo codes. Namely, the setting involves absorption, the presence of two media with different collision rates and the presence of medium-border crossing phenomena.

For the HM perturbations of algorithm 2.2, we set the collision-point perturbation variance $\sigma^2_{hm} = 0.5^2$, the probability of changing the absorption/non absorption in the obstacle sphere $P_{c,p} = 0.1$ and in the rest of the box $P_{c,w} = 0.05$. As in section 5.1.1, these settings are reasonable, but are not tuned as to yield an optimal performance of the interacting-particle method.

In figure 3, we present the results for 50 independent estimations with the interacting-particle method. Empirically, the estimator is unbiased and the theoretical 95% confidence intervals are valid. This is the same conclusion as in section 5.1.1, and is again a validation of the HM algorithm in the space of the monokinetic-particle trajectories.

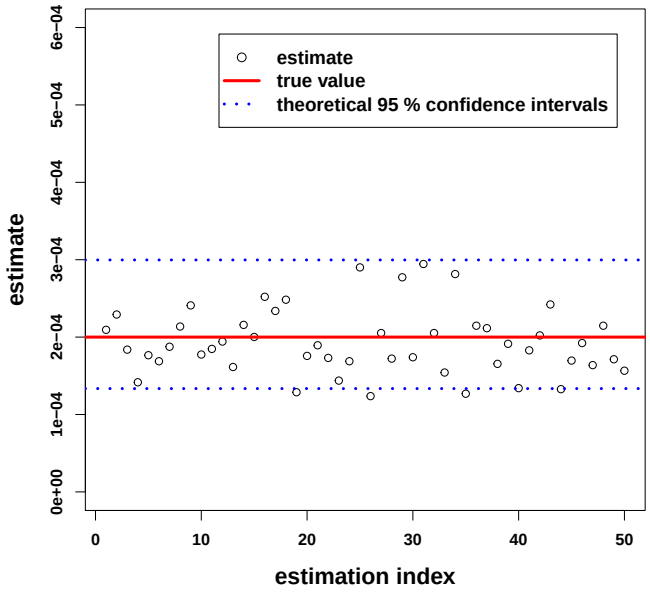

FIGURE 3. Two-dimensional-case. Plot of 50 independently estimated probabilities with the interacting-particle method 2.3, for number of particles $N = 200$ and number of iterations in the HM algorithm $T = 300$. We also plot the theoretical 95% confidence intervals (1) of the case $T = +\infty$. The true probability $p = 0.2 \times 10^{-4}$ is evaluated quasi-exactly by a simple-Monte Carlo method. The interacting-particle estimator is empirically unbiased and the 95% theoretical confidence interval, obtained in the case $T = +\infty$, is adapted to the practical estimator.

5.2.2. *Comparison with simple Monte Carlo in a small-probability case*

We now consider the case of a small probability. For this, we set the absorption probability in the obstacle sphere $P_{a,p} = 0.7$ and in the rest of the box $P_{a,w} = 0.5$, the dimension of the box $[-\frac{L}{2}, -\frac{L}{2}]$ $L = 10$, the radius of the obstacle sphere $l = 2.5$, the radius of the detector $l_d = 0.5$. The positions of the detector and the source are given by $p_{det,x} = p_{s,x} = 3$. We set the rate of collisions in the water medium $\lambda_w = 2$ and in the poison medium $\lambda_p = 3$. In essence, the obstacle sphere is larger than in subsection 5.2.1, the absorption probabilities are larger, and the collision rates are larger, thus yielding all the more frequent absorption.

The probability p is estimated by very large Monte Carlo with sample size 1.25×10^9. The estimate is $\frac{22}{1.25 \times 10^9} \approx 1.76 \times 10^{-8}$. Similarly to section 5.1.2, the Clopper-Pearson 95% confidence interval for the probability is $[10^{-8}, 2.5 \times 10^{-8}]$. It is also small enough for validating the discussion that follows.

We compare the estimators $\hat{p}_{ipm}$, with $N = 200$ particles and $T = 300$ iterations in the HM algorithm, and the estimator $\hat{p}_{mc}$ with sample size 5×10^6. We have found that the computation time for the estimator $\hat{p}_{ipm}$ is, on average, 88% of that of the estimator $\hat{p}_{mc}$.

Now, concerning estimation accuracy, the results are presented in figure 4. The interacting-particle method outperforms the simple-Monte Carlo method, to a greater extent that in figure 2. As a confirmation, the quality ratio $\frac{\sqrt{TIME_{mc}} \times RMSE_{mc}}{\sqrt{TIME_{ipm}} \times RMSE_{ipm}}$ is 10.5, against 6.7 in figure 2.

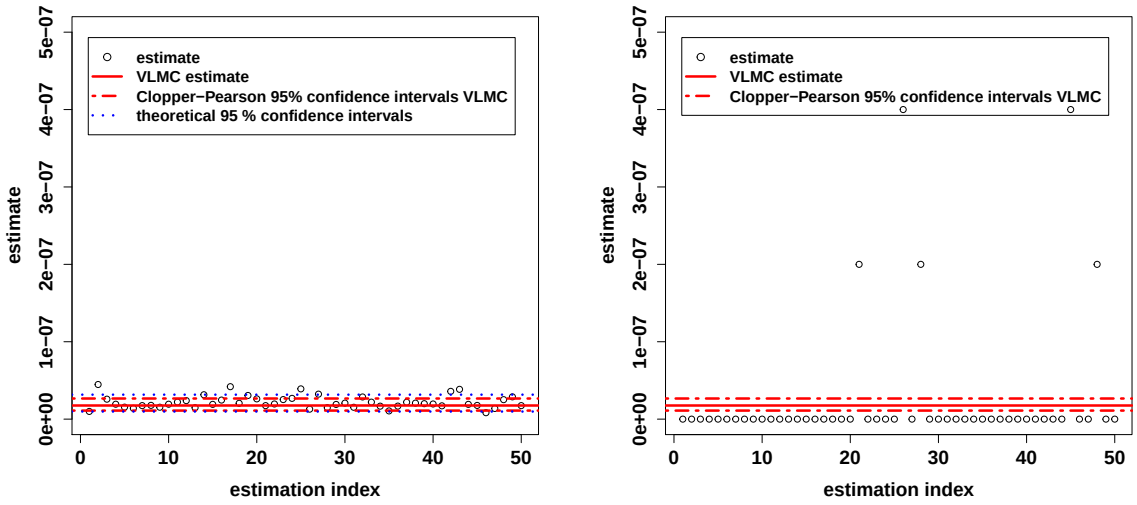

FIGURE 4. Two-dimensional-case. Plot of 50 independently estimated probabilities with the interacting-particle method of algorithm 2.3 for $N = 200$ and $T = 300$ (left) and with the simple-Monte Carlo method with sample size 5×10^6 (right). We plot a very large Monte Carlo estimate of the true probability $\hat{p}_{vlmc} = 1.76 \times 10^{-8}$, together with the associated Clopper-Pearson 95% confidence intervals. For the interacting-particle method, we also plot the theoretical 95% confidence intervals of the case $T = +\infty$, assuming the true probability is the VLMC estimate. The uncertainty on the VLMC estimate of the true value p of the probability is small enough for our conclusions to hold. These conclusions are that the interacting-particle method outperforms the Monte Carlo method (with sample size 5×10^6), both in term of computation time and of accuracy.

5.3. **Discussion on the numerical results**

We now discuss some conclusions on the numerical results of section 5. First, in two cases with a probability that is not small (figures 1 and 3), the interacting-particle method is empirically unbiased. The theoretical confidence intervals $T = +\infty$ are in agreement with the empirical distribution for finite T, provided that T is large enough. For the two cases of small probabilities (figures 2 and 4), we do not state conclusions on this question, in one sense or another, because we do not know the probability with negligible uncertainty.

However, for figures 2 and 4, the uncertainty on the probability is small enough to compare the performances of the simple-Monte Carlo and interacting-particle methods. The conclusion of this comparison is strongly unilateral, and is that, for a small probability, the interacting-particle method is preferable over a simple-Monte Carlo sampling.

We have not carried out numerical test for extremely small probabilities (say, under 10^{-10}). The reason for that is that we would not have an estimate of these probabilities similar to the very large Monte Carlo estimate $\hat{p}_{vlmc}$. That it to say an estimate that comes with confidence intervals with guaranteed coverage probability. Nevertheless, a simple-Monte Carlo method, with computational time similar to that of the interacting-particle method, would most likely never see the rare event, and thus only provide an overly conservative upper bound. Thus, the comparison would be even more in favor of the interacting-particle method than for figures 2 and 4.

In figure 1, we have mentioned the trade-off problem between the number of particles N and the number of HM iterations T. The average complexity of the interacting-particle method is proportional to the product NT. Naturally, increasing N improves the accuracy of the interacting-particle method. Especially, the variance

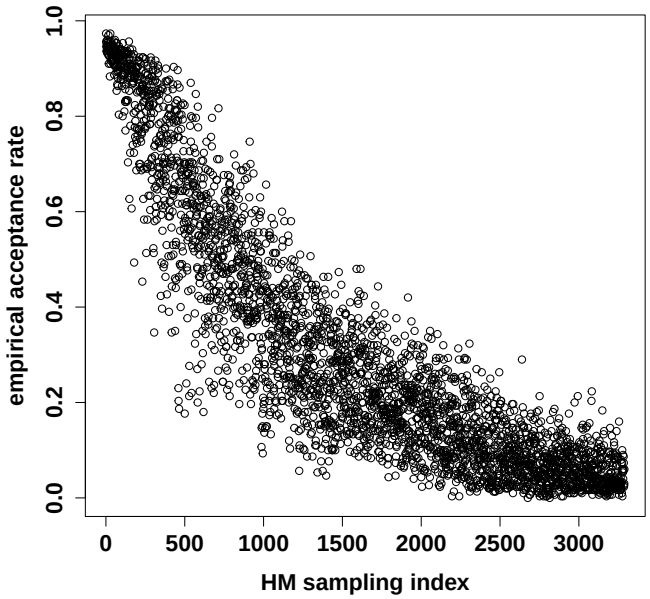

FIGURE 5. Same setting as in figure 2 for the interacting-particle method. We consider one estimation $\hat{p}_{ipm}$ of the interacting-particle method. We plot the empirical acceptance rate, over the $T = 300$ proposed perturbations, for each of the $\frac{\log \hat{p}_{ipm}}{\log\left(1-\frac{1}{N}\right)}$ HM samplings of algorithm 2.2. The acceptance rate decreases considerably when one gets closer to the rare-event.

is proportional to N when N is large, in the ideal case $T = +\infty$. We have seen in figure 2 that increasing T also reduces the variance, which is well interpreted. It is however quite difficult to quantify the dependence between T and the variance of the estimator. We think that the question of this trade-off between N and T would benefit from further investigation.

In our experiments, we have not optimized the choice of the perturbation method. This would naturally bring a potential additional benefit for the interacting-particle method. Perhaps less natural is the prospect of allowing the perturbation method to vary with the progression of the algorithm. For example, one could use a perturbation method that propose perturbed trajectories that are closer to the initial ones, when these trajectories are close to the rare event. The results we now present in figure 5 support this idea. In figure 5, we plot the acceptance rate in the HM algorithm 2.2 (by acceptance we mean that both the pdf ratio and the objective function conditions are fulfilled), as a function of the progression in the interacting-particle method. This acceptance rate is decreasing, and is small when the interacting-particle method is in the rare-event state. Notice that this was not the case in the experiments conducted in [10].

An other potential tuning of the interacting-particle method is the choice of the objective function Φ, for which the event "the monokinetic particle makes a collision in the detector" is equivalent to the event that Φ, evaluated on the trajectory of the monokinetic particle, exceeds a threshold. We have used as a function Φ the (opposite of the) minimum, over the collision points of the trajectory, of the Euclidean distance to the center of the detector. This choice could be improved. One natural possibility is to replace the Euclidean distance by the optical distance. That is to say the distance traveled in each medium would be weighted by the collision rate in the medium. For some neutron-transport problem, it is also possible to use more specific objective functions, by finding approximations of the importance function, see e.g. [9].

Conclusion

We have considered the adaptation of the interacting-particle method [10] to a small-probability estimation problem, motivated by shielding studies in neutron transport. The adaptation is not straightforward, because shielding studies involve working on probability distributions on a set of trajectories that are killed after a finite time.

The contribution brought by the paper it two-fold. First, it has been shown that probability density functions can be defined on this set. This enables to use the Hastings-Metropolis algorithm, which is necessary to implement the method [10] in practice. A convergence result has also been shown for the Hastings-Metropolis algorithm in this setting.

The second contribution of the paper is to give the actual probability density function equations, for implementing the interacting-particle method in an academic one-dimensional problem, and a simplified but realistic two-dimensional problem. In both cases, the method is shown to be valid. Furthermore, the method outperforms a simple-Monte Carlo estimator, for estimating a small probability.

Prospects are possible for both contributions. First, the proof of the convergence of the Hastings-Metropolis could be extended under more general assumptions. Second, several possibilities for practical improvement of the interacting-particle method are presented in section 5.3.

The authors are thankful to Josselin Garnier, for suggesting the adaptation of the interacting-particle method to Monte-Carlo codes in neutronic and for his advice. They acknowledge the role of Jean-Marc Martinez in the initiation and the conduct of the research project. They thank Fosto Malvagi and Eric Dumonteil for giving them an introduction to neutron-transport problems. Eric Dumonteil also contributed to the writing of the introduction to the neutronic problem. The authors are grateful to Nicolas Champagnat, for his advice during the research project.

References

[1] Agostinelli, S. & al, *Geant4: a simulation toolkit*, Nuclear instruments and methods in physics research section A 506, 250-303, 2003.
[2] Booth, T.E., *Comments on Monte Carlo probability of initiation estimates for neutron fission chains*, Nucl. Sci. Eng. 166, 175-178, 2010.
[3] Booth, T.E., *Common misconceptions in Monte Carlo particle transport*, Applied Radiation and Isotopes 70, 1042-1051, 2012.
[4] Both, J.P. & AL,, *Automated importance generation and biasing techniques for Monte-Carlo shielding techniques by the Tripoli-3-code*, Progress in Nuclear Energy 24, 273-281, 1990.
[5] Burn, K. W., *Complete optimization of space/energy cell importance with the dsa cell importance model*, Ann. Nucl. Energy 19(2), 65-98, 1992.
[6] Clopper, C. and Pearson, E. S., *The use of confidence or fiducial limits illustrated in the case of the binomial*, Biometrika 26(4), 404413, 1934.
[7] Del Moral, P. & Garnier, J., *Genealogical particle analysis of rare events* Ann. Appli. Probab. 15, 2496-2534, 2005.
[8] Diop, C. & al, *Tripoli-4: a 3D continuous-energy Monte Carlo transport code*, Trans. Am. Nuc. Soc. 95, 661, 2006
[9] Dumonteil, E., *On a New Variance Reduction Technique: Neural Network Biasing - A Study of Two Test Cases with the Monte Carlo Code Tripoli4*, Nuclear Technology 168(3), 793-798, 2009.
[10] Guyader, A., Hengartner, N., Matzner-Lober,E., *Simulation and Estimation of Extreme Quantiles and Extreme Probabilities*, Applied Mathematics & Optimization 64(2), 171-196, 2011.
[11] Hammersley, J. and Handscomb, D., *Monte Carlo methods*, Methuen, London, 1965.
[12] Hastings, W. K., *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika 57, 97-109, 1970.
[13] Hoogenboom, E., *Zero-variance Monte Carlo schemes revisited*, Nucl. Sci. Eng. 160, 1-22, 2008.
[14] Kahn, H., *Applications of Monte Carlo*, RM-1237-AEC, Rand Corporation, 1956.
[15] Los Alamos National Laboratory, *Monte Carlo Code Group*, https://mcnp.lanl.gov/.
[16] Meyn, S.P. & Tweedie, R. L., *Markov chains and stochastic stability*, Cambridge university press, 2009.
[17] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E., *Equations of state calculations by fast computing machines*, J. Chem. Phys. 21, 1087-32, 1953.
[18] Nummelin, E., *General irreducible Markov chains and non-negative operators*, Cambridge University Press, 1984.
[19] Tierney, L., *Markov chains for exploring posterior distributions*, The annals of statistics 22, 1702-1762, 1994.

[20] Zoia, A., & al, *Branching exponential flights: traveled lengths and collision statistics*, J. Phys. A: Math. Theor. 45, 425002, 2012.
[21] Zoia, A., & al, *Collision densities and mean residence times for d-dimensional exponential flights*, Physical Review E83, 041137, 2011.