HAL open science

Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests

Antoine Lejay, Géraldine Pichot

To cite this version:

Antoine Lejay, Géraldine Pichot. Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests. Journal of Computational Physics, 2016. hal-01003853v2

HAL Id: hal-01003853

https://inria.hal.science/hal-01003853v2

Submitted on 23 Feb 2016 (v2), last revised 5 Jul 2017 (v4)

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

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

Simulating Diffusion Processes in Discontinuous Media: Benchmark Tests

Antoine Lejay[*†‡§] Géraldine Pichot[¶‖]

February 19, 2016

Abstract

We present several benchmark tests for Monte Carlo methods simulating diffusion in one-dimensional discontinuous media. These benchmark tests aim at studying the potential bias of the schemes and their impact on the estimation of micro- or macroscopic quantities (repartition of masses, fluxes, mean residence time, ...). These benchmark tests are backed by a statistical analysis to filter out the bias from the unavoidable Monte Carlo error. We apply them on four different algorithms. The results of the numerical tests give a valuable insight of the fine behavior of these schemes, as well as rules to choose between them.

Keywords. Monte Carlo methods for discontinuous media; Fick's law; breakthrough curve

[*]Université de Lorraine, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54500, France

[†]CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54500, France

[‡]Inria, Villers-lès-Nancy, F-54600, France

[§]Contact: IECL, BP 70238, F-54506 Vandœuvre-lès-Nancy CEDEX, France. Email: `Antoine.Lejay@univ-lorraine.fr`

[¶]Inria, Rennes, France

[‖]Contact: Inria de Paris, Inria, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France. Email: `Geraldine.Pichot@inria.fr`

1 Introduction

Many diffusion models arising in geophysics, population ecology and biology involve second-order operators of type $\nabla(D\nabla\cdot)$ with a discontinuous diffusivity D.

Monte Carlo methods provide simple ways to solve diffusion problems. Random walk techniques are popular in the geophysical community [55, 57]. Basically, the physical quantity of interest (*e.g.* the concentration of a solute) is approximated by averaging a suitable function over the positions of a large cloud of particles. For this, we need a rule for moving the particles during a small (deterministic or random) time step in a way which respects the physics. For linear equations, the particles move independently. Besides, their random future positions depend only on their current ones as well as on their immediate environments.

A simple technique when D is differentiable is to move the particle during a time step δt from its position x to

$$x + \xi\sqrt{2D(x)\delta t} + \nabla D(x)\delta t, \tag{1}$$

where ξ follows the unit, centered normal distribution.

This scheme no longer works when D is discontinuous (see *e.g.* [19, 27] for numerical tests). The latter case is still a challenging problem. However, the dynamic of the particles is now well understood for one-dimensional media.

Many interpretations and simulation techniques have been proposed during the last twenty years. Some numerical methods consider only the mathematical aspect of the simulation [12–14, 30, 31, 36–38] while others are driven by applications in a specific field: in geophysics [1–4, 9, 11, 22, 24–27, 33, 35, 43–46, 48, 55], fluid/gas dynamics [21], ecology [8, 42, 47], brain imaging [16], astrophysics [34, 56], meteorology [54], oceanography [19, 20, 53], molecular dynamics [7, 39], among others.

To validate the numerical methods, different benchmark tests have been developed. In [26, 27], several schemes are tested by comparing the concentration with analytic solutions and by checking that the proportion of particles on each side of the interface is correct in the steady state regime. In [26], the different schemes are also qualitatively evaluated for symmetry. In [52] and [19], which are related to oceanography, density and residence times are compared with respect to known values. In [4, 47], the first and second spatial moments in a two-layer aquifer system are estimated in a long time regime.

Our approach is different. While most of the benchmark tests in the literature aim at being realistic as *e.g.* in [47] or the Couplex test cases [6] or [4, Test scenarios S2-1, S2-2], the benchmark tests we propose here do not fulfil the same goal. Our objective is to quantify the *bias* of the schemes. The *bias* is the error

induced by the approximation schemes. The smaller, the better. However, the bias may be small in front of the Monte Carlo error. Monte Carlo simulations, justified by the law of large numbers, comes with an error in general of order $\mathrm{O}(N^{-1/2})$, with N the number of particles. In the benchmark tests proposed here, N is very large — from 10^5 to 4×10^6 particles — so that the Monte Carlo error is small. Besides, we quantify it to detect potential bias of the schemes by using confidence intervals. For a given time step, the size of the domain is chosen small enough so as to maximize the number of passages through the interfaces. It is also chosen so as to easily test new schemes as the approximation scheme is applied only in a boundary layer around the interface of discontinuity. A test is *passed* if one cannot distinguish the bias from the Monte Carlo error. Otherwise the test *failed*. Invalidating a scheme does not means it should be ruled out. A scheme could be fair enough for computing some macroscopic parameters but not for dealing with microscopic ones.

As we are interested in the behavior of schemes taking the discontinuities of the diffusivity into account, we consider that the diffusivity is piecewise constant over the medium. We do not consider situations where the diffusivity varies regularly not to add supplementary approximation errors. Our benchmark tests concern both the transient and the steady state regime. We propose five benchmark tests:

- Density:
  - Medium description: an infinite medium with an interface at $x_I = 0$, and diffusivities D^- at the left and D^+ at the right of it.
  - Test: check if the particles has the correct distribution in a bimaterial infinite medium.
- Layer:
  - Medium description: a periodic medium $[0, L]$ of diffusivity D_0, excepted on a layer of diffusivity D_m on $[L/2 - \ell, L/2 + \ell]$.
  - Test: check if the deviation from the uniform distribution is significant or not. Indeed, in the steady state regime, with periodic boundary conditions, whatever D, the particles should be uniformly distributed.
- Bimaterial:
  - Medium description: a medium $[0, L]$ with one interface at $x_I = L/2$, a diffusivity D^- (resp. D^+) on $[0, x_I]$ (resp. $[x_I, L]$) and reflecting boundary conditions (BC).
  - Test: check if the proportion of particles in the right-hand side of the medium is accurate, as this quantity may be analytically computed.
- Bimaterial absorbing I & II:
  - Medium description: medium $[0, L]$ with one interface at $x_I = L/2$, a diffusivity D^- (resp. D^+) on $[0, x_I]$ (resp. $[x_I, L]$) and a reflecting BC at 0 and an absorbing BC at L.

  - Test: check the accuracy of the loss of mass when one of the boundary is absorbing.
- SYMMETRY:
  - Medium description: same medium and same BC as BIMATERIAL.
  - Test: check whether or not the density $\mathfrak{q}(t, x, y)$, that is the density of the probability of a particle to go from x to a small volume around y during the time t, is symmetric in x and y. The more the scheme respects this property of symmetry, the better.

We apply those benchmark tests on four schemes with constant time steps (our framework is not the one of Continuous Time Random Walks, which consider random time steps, see *e.g.* [35] and references within), namely,

- The exact density-based, constant time step algorithm based on the exact method proposed by [28],
- The algorithm based on the approximation method proposed by Uffink [55],
- The algorithm based on the approximation method proposed by Hoteit *et al.* [22],
- A simpler version of the exact density-based algorithm with a linear interpolation for the time in case of crossing [28].

The numerical studies show the accurate or odd behavior of each schemes when computing the steady state and the transient regime.

Outline. A reminder on some theoretical results about stochastic processes are given in Section 2. Section 3 presents the five benchmark tests together with their theoretical foundations. We illustrate the use of these benchmark tests on four algorithms presented in Section 4. The results of the numerical simulations are presented in Section 5. Finally, we expose our conclusions in Section 6. Notice the two first Sections have been written so that the benchmark tests can be easily reused to test other schemes.

2 Theoretical results on diffusion processes, assumptions and methods

We present very briefly the results regarding stochastic processes on which our benchmark tests are based. The Monte Carlo methods are built on the simulation of these processes.

2.1 Stochastic processes and Fokker-Planck equations

We consider only one-dimensional medium $[0, L]$ of finite size with periodic, reflecting or absorbing boundary conditions (BC). The medium is defined by its diffusivity D on $[0, L]$ and the BC at 0 and L.

The particles are initially distributed with a probability ν. At time t, they are distributed with a density $f(t,\cdot)$ solution to

$$\begin{cases} \partial_t f(t,y) = \nabla(D(y)\nabla f(t,y)), \\ f(t,\cdot) \xrightarrow[t\to 0]{\text{weakly}} \nu, \\ D(0)\nabla f(t,0) = D(L)\nabla f(t,L) = 0 \text{ for reflecting BC at 0 and } L, \\ \text{ or } f(t,0) = f(t,L) \text{ for periodic BC}, \\ \text{ or } \begin{cases} D(0)\nabla f(t,0) = 0 \\ f(t,L) = 0 \end{cases} \quad \text{for reflecting BC at 0 and absorbing BC at } L. \end{cases} \tag{2}$$

This framework can be applied to many different diffusion problems by relating the particle density to the physical quantity of interest (e.g. the concentration of a solute).

The positions of the particles are appropriately defined by the paths of a stochastic process $(X_t)_{t\geq 0}$ indexed by the time on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The BC are taken into account in the distribution of $(X_t)_{t\geq 0}$. For example, the particle is stopped when reaching an absorbing BC.

The process follows the *Markov property*. This means roughly that for a given time $s > 0$, the distribution of $(X_t)_{t\geq s}$ of the future positions depends only on X_s and not on its prior positions $(X_r)_{r<s}$.

A numerical scheme provides us with an approximation of a path of $(X_t)_{t\geq 0}$. Justified by the Markov property, the simplest scheme consists in simulating $X_{t+\delta t}$ when X_t is known. This is a *constant time step scheme*. When D is smooth, the process is solution to the Stochastic Differential Equations SDE [18, 41]

$$X_t = x + \int_0^t \sqrt{2D(X_s)}\,\mathrm{d}W_s + \int_0^t \nabla D(X_s)\,\mathrm{d}s,$$

for a Brownian motion W. Therefore, the rule (1) with $x = X_t$ provides such a scheme [23, 40].

If D is discontinuous, this representation (2.1) is no longer valid. However, a diffusion process is associated to the divergence-form operator $\nabla(D(y)\nabla\cdot)$. A large amount of known results on the links between SDE and differential operators remain true, despite X is not solution to some SDE.

For numerical approximation, knowing the density of X_t given $X_s = x$ for each $t > s$ is sufficient. As the diffusivity is homogeneous in time, the density $\mathfrak{q}(t-s, x, y)$

of X_t given $X_s = x$, called the *fundamental solution* (or *Green function*) of $\nabla(D\nabla\cdot)$ is solution to the *Fokker-Planck* (or *Kolmogorov forward*) equation

$$\begin{cases} \partial_t \mathfrak{q}(t,x,y) = \nabla_y(D(y)\nabla_y \mathfrak{q}(t,x,y)), \\ \mathfrak{q}(t,x,y) \xrightarrow[t\to 0]{\text{weakly}} \delta_x(y), \\ \mathfrak{q}(t,x,\cdot) \text{ satisfies absorbing, reflecting or periodic BC.} \end{cases}$$

The density $f(t,y)$ solution to (2) is then equal to

$$f(t,y) = \int_0^L \nu(\,\mathrm{d}x)\mathfrak{q}(t,x,y). \tag{3}$$

The *probability current* or *flux* is $J(t,y) = -D(y)\nabla_y f(t,y)$. For each $t > 0$, $J(t,\cdot)$ is continuous over the medium, even in presence of discontinuities. At some point x_I of discontinuity of D, $J(t,x_I-) = J(t,x_I+)$ implies that

$$D(x_I-)\nabla f(t,x_I-) = D(x_I+)\nabla f(t,x_I+). \tag{4}$$

The proportion $p_{[a,b]}$ of particles in a box $[a,b] \subset [0,L]$ is $p_{[a,b]}(t) = \int_a^b f(t,y)\,\mathrm{d}y$. Integrating (2), its variation is

$$\partial_t p_{[a,b]}(t) = J(t,a) - J(t,b). \tag{5}$$

Hence, with N particles at positions $X_t^{(i)}$ at time t, $p_{[a,b]}(t)$ is easily approximated by

$$p_{[a,b]}(t) \approx \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{[a,b]}(X_t^{(i)}). \tag{6}$$

If the BC at 0 is the same as the BC at L, then $\mathfrak{q}(t,x,y) = \mathfrak{q}(t,y,x)$ for any $t > 0$ and $x, y \in [0,L]$. Thus $x \mapsto \mathfrak{q}(t,x,y)$ also satisfies (4).

2.2 Piecewise constant diffusivity

When the diffusivity is constant and equal to $D = 1/2$ over an infinite medium, the stochastic process X is simply the *Brownian motion* and $\mathfrak{q}(t,x,y)$ is nothing more than the Gaussian kernel

$$\mathfrak{g}(t, y-x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right).$$

With $D(x) = D^+$ if $x \geq 0$ and D^- if $x \leq 0$, the density transition function of the process X is (see *e.g.* [28, 55])

$$\mathfrak{q}(t, x, y) = \frac{1}{\sqrt{2D(y)}} \mathfrak{p}_\theta \left(t, \frac{x}{\sqrt{2D(x)}}, \frac{y}{\sqrt{2D(y)}} \right), \tag{7}$$

with

$$\theta = \frac{\sqrt{D^+} - \sqrt{D^-}}{\sqrt{D^+} + \sqrt{D^-}}. \tag{8}$$

Here $\mathfrak{p}_\theta(t, x, y)$ is the density transition function of the Skew Brownian motion of parameter θ [29] defined by

$$\mathfrak{p}_\theta(t, x, y) = \mathfrak{g}(t, y - x) + \mathrm{sgn}(y)\theta\mathfrak{g}(t, |y| + |x|).$$

In [28], we have constructed a scheme to simulate $X_{t+\delta t}$ from X_t by using (7). In a more general situation (finite media, presence of several discontinuities), there is no simple formula for the density transition function. Yet in short time, $\mathfrak{q}(\delta t, \cdot, \cdot)$ given by (7) could be used as an approximation of this density.

From the numerical point of view, the displacement of the particle during t and $t+\delta t$ is mostly influenced by the value of D close to X_t, in a region of size $\mathrm{O}(\sqrt{\delta t})$. For this reason, assuming a piecewise constant diffusivity is not restrictive at all provided that $\sqrt{\delta t}$ is small enough with respect to the distance between two discontinuities (see Section 3.6 as well as [28]).

2.3 Significance tests

Denote by X_t (resp. $\overline{X}_t$) the position of the particles at time t moved with the real dynamics (resp. when one of the scheme is used), and by $X_t^{(i)}$ (resp. $\overline{X}_t^{(i)}$) the position at time t of the i-th particle moved with the real dynamics (resp. when one of the schemes is used) when N independent particles paths are drawn.

A part of our methodology relies on the theory of *significance tests* (see *e.g.* [17, Chap. 12] or [10]). To be more precise, we estimate the distance between a value Λ and an empirical quantities Λ_N (resp. $\overline{\Lambda}_N$) constructed using N particle moving according to the true dynamics (resp. the scheme). Typically, $\Lambda = p_{\mathcal{V}}(t) = \mathbb{P}[X_t \in \mathcal{V}]$ for a volume $\mathcal{V}$. Its numerical approximation, combining the scheme and the Monte Carlo method, is $\overline{\Lambda}_N = N^{-1}\#\{i; \overline{X}_t^{(i)} \in \mathcal{V}\}$ for N independent realizations of the positions of the particles moving according to one of the schemes.

We place ourselves in situation where thanks to the Central Limit Theorem, for N large enough, $\sqrt{N}(\Lambda_N - \Lambda)$ is close in distribution to κG for a constant κ and G is

a unit, centered normal distribution $\mathcal{N}(0,1)$. We fix a confidence level α close to 1 and we set d_α so that

$$\mathbb{P}[G \in [-d_\alpha, d_\alpha]] = \alpha. \tag{9}$$

To draw a conclusion from our test, we then compare $\sqrt{N}|\overline{\Lambda}_N - \Lambda|$ with κd_α. We use for $\alpha = 99\,\%$, so that $d_\alpha = 2.57$. When $\Lambda(t)$ and $\kappa(t)$ depend on a parameter t, $t \mapsto \pm\kappa(t)d_\alpha$ is called a *confidence band*.

In the several situations we consider, alternative statistical tests could be constructed. However, for the sake of simplicity, we prefer simple procedures combined with graphical approaches.

3 Benchmark tests

A good benchmark test should be

- Physically relevant, *i.e.*, relative to a quantity of practical interest.
- Numerically relevant, *i.e.* sensitive to a quality or default of the scheme to replicate a physical phenomenon, a correct flux for example.
- Analytically relevant, *i.e.* the quantity of interest may be compared with an exact or a well approximated value.
- Statistically relevant as using empirical means over N particles leads to quantifiable fluctuations. This last point is important to discriminate the bias from the Monte Carlo error.

We consider five benchmark tests built on these criteria:

- Density: an infinite, bimaterial medium of diffusivity D^- (resp. D^+) on each side of the interface.
- Layer: a periodic medium $[0, L]$ of diffusivity D_0, excepted on a layer of diffusivity D_m on $[L/2 - \ell, L/2 + \ell]$.
- Bimaterial: a medium $[0, L]$ with one interface at $x_I = L/2$, a diffusivity D^- (resp. D^+) on $[0, x_I]$ (resp. $[x_I, L]$) and reflecting boundary conditions (BC).
- Bimaterial absorbing: medium $[0, L]$ with one interface at $x_I = L/2$, a diffusivity D^- (resp. D^+) on $[0, x_I]$ (resp. $[x_I, L]$) and a reflecting BC at 0 and an absorbing BC at L
- Symmetry: same conditions as Bimaterial.

3.1 Density benchmark test: check the distribution of the particles in an infinite bimaterial medium

In this benchmark test, we check if the particles has the correct distribution in a bimaterial, infinite medium. For this, we use the analytic density given by (7). We also compare the distribution functions (DF).

3.1.1 Density: Description

We consider an infinite medium with an interface at $x_I = 0$, and diffusivities D^- at the left and D^+ at the right of it (see Figure 1).

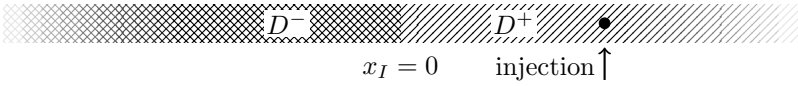


Figure 1: Medium for the Density test.

3.1.2 Theoretical results: the Kolmogorov-Smirnov distance

Let us consider two (one-dimensional) random variables X and Y with respective distributions functions (DF) F and G. We expect Y to be "close" to X.

Among the possible measures of the difference between Y and X is the *Kolmogorov-Smirnov distance*:

$$d_{\mathrm{KS}}(F, G) = \sup_{-\infty<y<+\infty} |F(x) - G(x)|.$$

In our situation, Y represents some variable obtained through a numerical scheme, for which we only know N independent samples $Y^{(i)}$, $i = 1, \ldots, N$. Thus, we consider the empirical DF $G_N(x) := N^{-1} \sum_{i=1}^{N} \mathbf{1}_{Y^{(i)} \leq x}$.

The Glivenko-Cantelli theorem states that $G_N(y) \xrightarrow[N\to\infty]{} G(y)$ for any y (see [17,50]). Besides, $G_N(y)$ follows a binomial distribution of parameter G and the following convergence holds for each y:

$$\sqrt{N}(G_N(y) - G(y)) \xrightarrow[N\to\infty]{\text{law}} B_b(G(y)), \tag{10}$$

where B_b is a Brownian bridge on $[0, 1]$ with $B_b(0) = B_b(1) = 0$. For each $y \in [0, 1]$, $B_b(y)$ follows the Gaussian distribution with mean 0 and variance $y(1 - y)$.

The normalized Kolmogorov-Smirnov distance $M_{\mathrm{KS}} := \sqrt{N} d_{\mathrm{KS}}(G_N, G)$ converges as N becomes large to the distribution of a maximum of a Brownian bridge [15,17,50]. Using for the *null hypothesis* that a DF G — known only through the empirical DF G_N — is equal to F, a hypothesis test is performed by using a threshold d_α for a confidence level α such that $\mathbb{P}[M_{\mathrm{KS}} \leq c_\alpha] = \alpha$. For all reasonable levels of α, $c_\alpha \leq 3$ (see the tables in [51] for tabulated values).

One advantage of this metric is that the asymptotic distribution of M_{KS} does not depend on a particular choice of F.

3.1.3 Density: Benchmark test definition

For the medium given above, the exact density of the positions X_t of the particles with $X_0 = x_0$ is $\mathfrak{q}(t, x_0, y)$ with $\mathfrak{q}$ follows the analytic formula given by (7).

To the density is associated the distribution function

$$F(y) := F(t, x_0, y) = \int_{-\infty}^{y} \mathfrak{q}(t, x_0, z)\,\mathrm{d}z = \mathbb{P}[X_t \leq z],\ -\infty < z < \infty.$$

Similarly, we denote by $\overline{\mathfrak{q}}_N(t, x_0, \cdot)$ and $\overline{F}_N := \overline{F}_N(t, x_0, \cdot)$ the empirical density and the empirical DF obtained by N samples of $\overline{X}$.

The statistics of interest is then the normalized Kolmogorov-Smirnov distance

$$M_{\mathrm{KS}} := \sup_{-\infty<y<+\infty} \sqrt{N}|\overline{F}_N(t, x_0, y) - F(t, x_0, y)| \tag{11}$$

for several values of t and x_0.

Density. Plot M_{KS} defined by (11) as a function of t and compare it with a threshold c_α of the Kolmogorov-Smirnov statistics for a confidence level α, or to the value $c_\alpha = 3$ which corresponds to a level of risk $1 - \alpha \approx 3 \times 10^{-7}$.

3.2 Layer benchmark test: check the distribution of the particles in the steady state regime

In this benchmark test, we check if the particles remain uniformly distributed after many steps in the steady state regime.

3.2.1 Layer: Description

We consider a periodic medium $[0, L]$ of diffusivity D_0, excepted on a layer of diffusivity D_m on $[L/2 - \ell, L/2 + \ell]$ (see Figure 2).

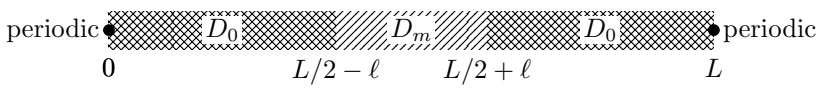


Figure 2: Medium for the Layer test.

3.2.2 Theoretical results: The steady-state regime — Invariant measure with periodic or reflecting BC

No mass is loss when the particle evolves on $[0, L]$ with either periodic or reflecting BC. Two facts are notable in this situation.

First, the Lebesgue measure $L^{-1}\,\mathrm{d}x$ is an *invariant measure* of the process: if the particles are uniformly distributed at initial time $t = 0$, then they remain uniformly distributed at any time $t > 0$. This case is referred to the *steady state*.

Second, the process is *ergodic* with respect to this measure. In particular, the distribution of X_t converges weakly to $L^{-1}\,\mathrm{d}x$.

We refer for example to [18] for a detailed account on these notions.

We use the notations introduced in Section 2.3. The empirical DF of X_t/L and $\overline{X}_t/L$ are, for $y \in [0,1]$,

$$F_N(t,y) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{X_t^{(i)} \le Ly} \text{ and } \overline{F}_N(t,y) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\overline{X}_t^{(i)} \le Ly}.$$

The DF of X_t/L and $\overline{X}_t/L$ are, for $y \in [0,1]$,

$$F(t,y) = \mathbb{P}[X_t \le Ly] = \mathbb{E}[F_N(t,y)] \text{ and } \overline{F}(t,y) = \mathbb{P}[\overline{X}_t \le Ly] = \mathbb{E}[\overline{F}_N(t,y)].$$

If the scheme is bias-free, then $\overline{F}(t,y) = F(t,y)$. Yet $\overline{F}(t,y)$ is only known through its empirical approximation $\overline{F}_N(t,y)$. From the results in Section 3.1.2, especially (10), it is expected that $\sqrt{N}(F_N(t,y) - F(t,y))$ behaves asymptotically for large N as a Brownian bridge $B_b(F(t,y))$. In particular, for $\alpha \in (0,1)$,

$$\mathbb{P}\big[B_b(u) \in [-d_\alpha\sqrt{u(1-u)}, d_\alpha\sqrt{u(1-u)}]\big] = \alpha,\ 0 < u < 1, \tag{12}$$

where d_α is defined by (9).

3.2.3 Layer: Benchmark test definition

As the uniform distribution is an invariant measure for the process X, if the particles are initially uniformly distributed over the medium, they should remain uniformly distributed as the time evolves, whatever the number of interfaces. If $X_0 \sim \mathcal{U}(0,L)$, then $X_t \sim \mathcal{U}(0,L)$ so that $F(t,y) = y$ for any $y \in [0,1]$.

In the steady state regime, the particles remains uniformly distributed over the medium. We then check whether or not this property is respected by the scheme. We test for t large enough (hence after many steps) the deviation of $\overline{F}_N(t,\cdot)$ from $F(t,\cdot)$ when the N particles are uniformly distributed at $t = 0$.

The statistic of interest is

$$K_N(t,y) := \sqrt{N}(\overline{F}_N(t,y) - F(t,y)) = \sqrt{N}(\overline{F}_N(t,y) - y). \tag{13}$$

The *null hypothesis* is that the empirical DF $\overline{F}_N(t,y)$ is a realization of $F_N(t,y) = y$ for any $y \in [0,1]$.

We could of course have used the Kolmogorov-Smirnov distance (see Section 3.1.2). We prefer a more graphical procedure relying on (12) and (13), and normal confidence intervals. With this approach, the correlations between $\overline{F}_N(t,x)$ and $\overline{F}_N(t,y)$ for $x \neq y$ are not taken into account. However, this statistical test is very simple to set up. Besides, it allows one to see where the bias of the schemes take their effect, if any, unlike the Kolmogorov-Smirnov test.

Layer. For a time t large enough, plot $K_N(t,y)$ and compare it with $y \in [0,1] \mapsto \pm d_\alpha \sqrt{y(1-y)}$ for a confidence level α with d_α is defined by (9).

This benchmark test is sensitive to the capacity of the numerical scheme to preserve the symmetry condition $\mathfrak{q}(\delta t, x, y) = \mathfrak{q}(\delta t, y, x)$ for any $x, y \in [0, L]$. As such, it is not restricted to this particular medium, and may be applied to media containing several layers for example.

3.3 Bimaterial benchmark tests: check the proportions of particles on each side of the interface in the steady state and transient regime

The third benchmark test checks the preservation of the flux condition (4) at the interface. It is a refinement of the one proposed by E. Labolle *et al.* [27]. It could be adapted to more general media, for example with multiple compartments.

3.3.1 Bimaterial: Description

We consider a medium $[0, L]$ with one interface at $x_I = L/2$ and reflecting BC (see Figure 3). The diffusivity is D^- (resp. D^+) on $[0, x_I]$ (resp. $[x_I, L]$). The particles start from an injection point at x with an initial mass equal to 1.

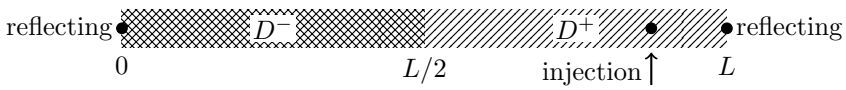


Figure 3: Medium for the Bimaterial test.

3.3.2 Theoretical results: Transient regime with reflecting boundary conditions at both endpoints

We set $\rho := D^-/D^+$.

There exists a family $\lambda_0 = 0 < \lambda_1 \leq \lambda_2 \leq \cdots$ as well as a family of functions $\{\phi_k\}_{k=0,1,\cdots}$ with $\phi_0(x) = 1/\sqrt{L}$ such that $\phi_k'(0) = \phi_k'(L) = 0$ and

$$\nabla(D\nabla\phi_k) = -\lambda_k^2 \phi_k \text{ and } \int_0^L \phi_k(x)\phi_j(x)\,\mathrm{d}x = \begin{cases} 0 & \text{for } j \neq k, \\ 1 & \text{for } j = k. \end{cases}$$

The eigenvalues $-\lambda_k^2$ are for any $k \geq 1$,

$$\lambda_k^2 = \frac{4\, z_k^2 D^+}{L^2} \text{ with } \frac{\tan(z_k)}{\sqrt{\rho}} = -\tan\left(\frac{z_k}{\sqrt{\rho}}\right) \tag{14}$$

where $z_0 = 0 < z_1 < z_2 < \cdots$.

The eigenfunctions are

$$\phi_k(x) = \frac{1}{\kappa_k} \begin{cases} \cos\left(\alpha_k^- x\right) & \text{if } x \in [0, L/2]\,, \\ \gamma_k \cos\left(\alpha_k^+ (L - x)\right) & \text{if } x \in [L/2, L]\,, \end{cases} \tag{15}$$

with

$$\alpha_k^\pm = \frac{\lambda_k}{\sqrt{D^\pm}},\ \gamma_k = \frac{\cos(\alpha_k^- L/2)}{\cos(\alpha_k^+ L/2)} \text{ and } \kappa_k^2 = \frac{L}{4}(1 + \gamma_k^2) + \frac{\sin(\alpha_k^- L)}{4\alpha_k^-} + \gamma_k^2 \frac{\sin(\alpha_k^+ L)}{4\alpha_k^+}.$$

For this family of eigenvalues and eigenfunctions,

$$\mathfrak{q}(t, x, y) = \frac{1}{L} + \sum_{k=1}^{+\infty} e^{-\lambda_k^2 t} \phi_k(x) \phi_k(y). \tag{16}$$

In particular, $\mathfrak{q}(t, x, \cdot)$ decreases exponentially fast to the uniform density over $[0, L]$. The value of λ_1^2 is the rate of convergence towards the steady state regime.

A tractable formula for the proportion of particles at a given time t in a volume $\mathcal{V}$ is deduced from (16) through $p_{\mathcal{V}}(t) = \int_0^L \mathrm{d}\nu(x) \int_{\mathcal{V}} \mathfrak{q}(t, x, y)\, \mathrm{d}y$.

The proportion of particles on $[0, x_I]$ (resp. $[x_I, L]$) at time t, follows from (3) so that

$$p_-(t) = p_{[0,x_I]}(t) = \int_0^{x_I} f(t, y)\, \mathrm{d}y = \int_0^{x_I} \mathfrak{q}(t, x, y)\, \mathrm{d}y$$

$$\text{resp. } p_+(t) = p_{[x_I,L]}(t) = \int_{x_I}^{L} f(t, y)\, \mathrm{d}y = \int_{x_I}^{L} \mathfrak{q}(t, x, y)\, \mathrm{d}y.$$

The spectral decomposition (16) of $\mathfrak{q}$ gives an analytic expression for the proportion of particles. In particular, after a short time, $p_\pm(t)$ converges to $1/2$ (since $x_I = L/2$) at time exponential rate $-\lambda_1^2$.

With (5), a scheme which respects well the flux $J(t, y\pm)$ should lead to a correct variation of $p_\pm(t)$.

Using (6), the idea of this benchmark is then to compare the theoretical evolution of $p_+(t)$ with the proportion of particles on the right-hand side of the medium.

In [27], it is shown that the repartition of masses is not correct if a simple Gaussian random walk is used. In the latter case, the gradient of f is continuous at the interface, not its flux.

ρ	2.5	5	7.5	10	12.5	15	17.5	20	100	250	500	750
z_1	1.838	1.934	1.966	1.982	1.991	1.998	2.002	2.005	2.024	2.027	2.027	2.028
z_2	3.987	4.503	4.663	4.735	4.775	4.800	4.818	4.831	4.897	4.907	4.910	4.911
z_3	5.651	6.647	7.256	7.532	7.664	7.737	7.853	7.854	7.951	7.968	7.973	7.975
z_4	7.738	8.412	9.023	9.677	10.18	10.487	10.655	10.753	11.045	11.070	11.078	11.080

Table 1: BIMATERIAL — Smallest positive solution $z_1, \ldots, z_4$ to (14) in function of $\rho = D^-/D^+$ and $D^- = 5$, from which the eigenvalues are computed.

3.3.3 BIMATERIAL: Benchmark test definition

We let N particles, starting from a fixed point, evolve in the medium with reflecting boundary conditions. We still use the notations of Section 2.3.

The statistics of interest are:

$$P_N(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{X_t^{(i)} \geq L/2} \text{ and } \overline{P}_N(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\overline{X}_t^{(i)} \geq L/2}. \tag{17}$$

The quantity $P_N(t)$ (resp. $\overline{P}_N(t)$) is the empirical mean number of particles moved with the real (resp. approximated) dynamic staying at the right side of the interface at time t, so that $P_N(t) \approx p_+(t)$.

With the spectral decomposition (16),

$$p_+(t) = \int_{L/2}^{L} \mathfrak{q}(t,x,y)\,\mathrm{d}y = \frac{1}{2} + \sum_{k=1}^{+\infty} c_k(x) e^{-\lambda_k^2 t} \tag{18}$$

where, for $\phi_k(\cdot)$ given by (15),

$$c_k(x) = \phi_k(x) \int_{L/2}^{L} \phi_k(y)\,\mathrm{d}y = \phi_k(x) \frac{\gamma_k}{\kappa_k \alpha_k^+} \sin\left(\alpha_k^+ \frac{L}{2}\right).$$

The first eigenvalue is $\lambda_0^2 = 0$ (corresponding to $z_0 = 0$) and $\phi_0 = 1/\sqrt{L}$. We report on Table 1 the smallest positive solutions z_k to (14) with $0 < z_1 \leq z_2 \leq \cdots$ for various ratios of $\rho = D^-/D^+$. The eigenvalues $-\lambda_k^2$ are then easily obtained.

The steady-state (or stationary) regime is reached when $\mathfrak{q}(t,x,y)$ is close to a constant function and then when $c_1(x)\exp(-\lambda_1^2 t)$ is close to 0.

Thanks to the exponential term, a truncated version of the sum up to order 4 provides a good approximation of $p_+(t)$.

For each time t, $NP_N(t)$ is a binomial random variable with N trials and a probability of success $p_+(t)$. Using the normal approximation, $\sqrt{N}(P_N(t) - p_+(t))$ is close to a normal distribution $\mathcal{N}(0, p_+(t)(1-p_+(t)))$.

Bimaterial. For a level of confidence α close to 1, check that for $t \in [0, T]$,

$$\sqrt{N}(\overline{P}_N(t) - p_+(t)) \in [-d_\alpha \sqrt{p_+(t)(1-p_+(t))}, d_\alpha \sqrt{p_+(t)(1-p_+(t))}]$$

with d_α defined by (9).

For t large enough, $p_+(t)$ is close to $1/2$, and $P_N(t)$ fluctuates around $1/2$ with variance $1/4N$.

3.4 Bimaterial absorbing benchmark test: check if the first exit time is correctly estimated

3.4.1 Bimaterial absorbing: Description

We now consider a medium with two compartments $[0, L/2]$ and $[L/2, L]$ of respective diffusivities D^- and D^+, and a reflecting BC at 0 and an absorbing one at L (see Figure 4). Again, we set $\rho = D^-/D^+$.

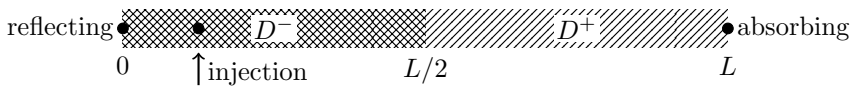


Figure 4: Medium for the Bimaterial absorbing test.

3.4.2 Theoretical results: Loss of mass with an absorbing boundary condition

An absorbing BC condition at L means that the particles are removed from the medium when reaching L, leading to a loss of mass. Let X be the stochastic process associated to $\nabla(D\nabla\cdot)$ in a medium with an absorbing BC.

Let τ be the random variable that gives the first time t at which X_t reaches the absorbing boundary.

If the particles are initially distributed according to the probability ν,

$$\mathbb{P}[\tau \le t] = 1 - \mathbb{P}[\tau > t] = 1 - \int_0^L \mathrm{d}\nu(x) \int_0^L \mathfrak{q}(t, x, y)\, \mathrm{d}y. \tag{19}$$

This quantity $\mathbb{P}[\tau > t]$ is the probability that the particle has not been absorbed before the time t. On the other hand, a spectral decomposition also holds for $\mathfrak{q}$: for the ordered eigenvalues $\widetilde{\lambda}_k$ and the corresponding eigenfunctions $\widetilde{\phi}_k$ with $\widetilde{\phi}_k'(0) = 0$ and $\widetilde{\phi}_k(L) = 0$,

$$\mathfrak{q}(t, x, y) = \sum_{k=0}^{+\infty} e^{-\widetilde{\lambda}_k^2 t} \widetilde{\phi}_k(x) \widetilde{\phi}_k(y) \text{ with } \int_0^L \widetilde{\phi}_k(x) \widetilde{\phi}_j(x)\, \mathrm{d}x = \begin{cases} 0 & \text{for } j \neq k, \\ 1 & \text{for } j = k. \end{cases} \tag{20}$$

The eigenvalues $-\widetilde{\lambda}_k^2$ are given by

$$\widetilde{\lambda}_k^2 = \frac{4\widetilde{z}_k^2 D^+}{L^2} \text{ with } \tan\left(\frac{\widetilde{z}_k}{\sqrt{\rho}}\right) = \frac{1}{\sqrt{\rho}\tan(\widetilde{z}_k)}, \tag{21}$$

where $0 < \widetilde{z}_0 < \widetilde{z}_1 < \cdots$. The corresponding eigenfunctions $\widetilde{\phi}_k$ are

$$\widetilde{\phi}_k(x) = \frac{1}{\widetilde{\kappa}_k}\begin{cases}\cos(\widetilde{\alpha}_k^- x) & \text{if } x \in [0, L/2],\\ \widetilde{\gamma}_k \sin(\widetilde{\alpha}_k^+(L-x)) & \text{if } x \in [L/2, L],\end{cases} \tag{22}$$

with

$$\widetilde{\alpha}_k^\pm = \frac{\widetilde{\lambda}_k}{\sqrt{D^\pm}},\ \widetilde{\gamma}_k = \frac{\cos\left(\frac{\widetilde{\alpha}_k^- L}{2}\right)}{\sin\left(\frac{\widetilde{\alpha}_k^+ L}{2}\right)},\ \widetilde{\kappa}_k^2 = \frac{L}{4}(1+\widetilde{\gamma}_k^2) + \frac{\sin\left(\widetilde{\alpha}_k^- L\right)}{4\widetilde{\alpha}_k^-} + \widetilde{\gamma}_k^2 \frac{\sin\left(\widetilde{\alpha}_k^+ L\right)}{4\widetilde{\alpha}_k^+}.$$

All the eigenvalues have multiplicity 1 and $\widetilde{\lambda}_0 \neq 0$. The eigenfunction $\widetilde{\phi}_0$ never vanishes in $(0, L)$. Hence $\mathfrak{q}(t, x, y)$ is exponentially fast decreasing and

$$\mathbb{P}[\tau \leq t] \approx 1 - \widetilde{\kappa} e^{-\widetilde{\lambda}_0^2 t} \text{ with } \widetilde{\kappa} = \int_0^L \mathrm{d}\nu(x) \int_0^L \widetilde{\phi}_0(x)\widetilde{\phi}_0(y)\,\mathrm{d}y. \tag{23}$$

The DF $G(t) = \mathbb{P}[\tau \leq t]$ of τ is the proportion of particles that have left the medium before time t. This function $G(t)$ is related to *breakthrough curves* [57] and *mean residence time* [5].

3.4.3 Bimaterial absorbing I: Benchmark test definition

For the particles starting from x, with (19) and (20),

$$G(t) = \mathbb{P}[\tau \leq t] = \sum_{k=0}^{+\infty} e^{-\widetilde{\lambda}_k^2 t}\widetilde{\phi}_k(x) \int_0^L \widetilde{\phi}_k(y)\,\mathrm{d}y \tag{24}$$

with $\widetilde{\phi}_k$ given by (22) and

$$\int_0^L \widetilde{\phi}_k(y)\,\mathrm{d}y = \frac{\sin\left(\widetilde{\alpha}^- L/2\right)}{\widetilde{\alpha}^- \widetilde{\kappa}_k} - \widetilde{\gamma}_k \frac{\cos\left(\widetilde{\alpha}^+ L/2\right)}{\widetilde{\alpha}^+ \widetilde{\kappa}_k} + \frac{\widetilde{\gamma}_k}{\widetilde{\alpha}^+ \widetilde{\kappa}_k}.$$

The values of the smallest positive roots $\widetilde{z}_0, \ldots, \widetilde{z}_4$ of (21) are given in Table 2.

We simulate for N particles the first time τ at which they reach the boundary where an absorbing BC holds.

We denote by $\overline{G}_N$ the empirical DF of the first exit time τ (see the entry *boundary layer* in Table 4 for the algorithm used to compute the exit time).

Bimaterial absorbing I. For a confidence level α, check if $\sqrt{N}(\overline{G}_N(t) - G(t))$ belongs to the confidence band $t \mapsto \pm d_\alpha \sqrt{G(t)(1-G(t))}$ with d_α defined by (9).

ρ	2.5	5	7.5	10	12.5	15	17.5	20	100	250	500	750
$\widetilde{z}_0$	0.830	0.845	0.850	0.853	0.854	0.855	0.856	0.856	0.860	0.860	0.860	0.860
$\widetilde{z}_1$	2.951	3.204	3.283	3.321	3.343	3.357	3.368	3.375	3.416	3.422	3.424	3.424
$\widetilde{z}_2$	4.840	5.698	6.024	6.159	6.229	6.271	6.299	6.319	6.416	6.429	6.433	6.435
$\widetilde{z}_3$	6.630	7.449	8.221	8.748	9.024	9.167	9.249	9.301	9.496	9.517	9.523	9.525
$\widetilde{z}_4$	8.799	9.617	10.002	10.487	11.051	11.551	11.902	12.115	12.599	12.628	12.637	12.640

Table 2: Bimaterial absorbing — Smallest positive solutions $\widetilde{z}_0, \ldots, \widetilde{z}_4$ to (21) in function of $\rho = D^-/D^+$ and $D^- = 5$, from which the eigenvalues are computed.

3.4.4 Bimaterial absorbing II: Benchmark test definition

Some applications requires only an accurate estimation of the rate of convergence towards 0 of the survival probability $1 - G(t)$. We propose a second benchmark test which relies on estimating this rate of convergence.

With (23) (see Figure 5),

$$\log(1 - G_N(t)) \approx \log(\widetilde{\kappa}) - \widetilde{\lambda}_0^2 t. \tag{25}$$

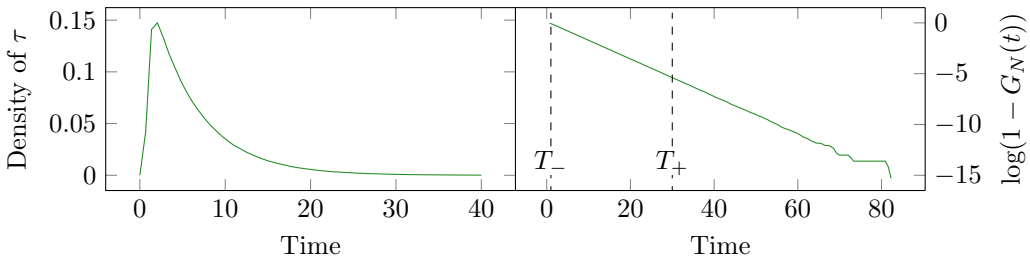


Figure 5: Bimaterial absorbing II — Estimation of the exponential convergence of $1 - \overline{G}_N(t)$ to 0.

We use as in [32] a linear least squares on $t \mapsto \log(1 - \overline{G}_N(t))$ on a window $[T_-, T_+]$. The choice of T_- and T_+ is crucial. On the one hand, T_- should be chosen large enough to avoid that $\exp(-\widetilde{\lambda}_0^2 t)$ dominates all the other terms. On the other hand, due to a rare event estimation problem, $\log(1 - \overline{G}_N(t))$ tends to oscillate for t large enough: The Monte Carlo error, of order $\sqrt{\overline{G}_N(t)(1 - \overline{G}_N(t))/N}$, is much more bigger than $1 - G(t)$. Thus, T_+ shall be chosen small enough to avoid the oscillations of $\log(1 - \overline{G}_N(t))$ (see Figure 5).

Bimaterial absorbing II. Estimate $-\widetilde{\lambda}_0^2$ from a linear least squares procedure on $t \in [T_-, T_+] \mapsto \log(1 - \overline{G}_N(t))$, where $\overline{G}_N$ is the empirical distribution function of the first exit time from the domain by the absorbing boundary.

3.5 Symmetry benchmark tests: check if the density transition function of the scheme is symmetric

3.5.1 Symmetry: Description

We use a bimaterial medium as in Figure 3.

3.5.2 Theoretical result: symmetry of the density

With periodic or reflecting BC at the endpoints 0 and L, the operator $\nabla(D\nabla\cdot)$ is self-adjoint and then $\mathfrak{q}(t,x,y) = \mathfrak{q}(t,y,x)$ for any $(x,y) \in [0,L]^2$. This means that the probability to go from x into a small volume $\mathrm{d}\mathcal{V}$ around y is the same as the probability to go from y into a small volume $\mathrm{d}\mathcal{V}$ around x. This property implies immediately that the Lebesgue measure is invariant measure for the dynamic.

We denote by p_{-+} (resp. p_{+-}) the proportion of particles during $[0,t]$ that goes from the compartment at the left ($x \leq L/2$) (resp. right) to the compartment at the right ($x \geq L/2$) (resp. left) of the discontinuity. If the particles are uniformly distributed on each compartment, then

$$p_{-+} = \frac{2}{L}\int_0^{L/2} \mathrm{d}x \int_{L/2}^{L} \mathrm{d}y\, \mathfrak{q}(t,x,y) = \frac{2}{L}\int_{L/2}^{L} \mathrm{d}x \int_0^{L/2} \mathrm{d}y\, \mathfrak{q}(t,x,y) = p_{+-}.$$

The symmetry property of $\mathfrak{q}(t,x,y)$ guarantees the correct exchange of particles between each compartment and the global equilibrium. Therefore, the more the scheme respects the symmetry of the density transition function, the better.

3.5.3 Symmetry: Benchmark test definition

For a box V of small size located around y, and a particle starting at x,

$$\mathbb{P}_x[X_t \in V] = \int_V \mathfrak{q}(t,x,y)\,\mathrm{d}y \approx |V| \times \mathfrak{q}(t,x,y).$$

We fix some integer n and we set $x_i = iL/n$ for $i = 0,\dots,n$, as well as $x_{i+1/2}$ the mid-point of $[x_i, x_{i+1}]$.

The domain $[0,L]$ is cut into small intervals $[x_i, x_{i+1}]$, $i = 0,\dots,n-1$. For N particles $X^{(1)},\dots,X^{(N)}$, set for $i,j = 0,\dots,n$,

$$\begin{aligned} Q_{ij}(t) &= \frac{n}{L}\frac{\#\{k \text{ such that } X_t^{(k)} \in [x_j, x_{j+1}] \text{ when } X_0 = x_{i+1/2}\}}{N} \\ &\approx \frac{n}{L}\int_{jL/n}^{(j+1)L/n} \overline{\mathfrak{q}}(t, x_{i+1/2}, y)\,\mathrm{d}y \approx \mathfrak{q}(t, x_{i+1/2}, x_{j+1/2}). \end{aligned} \tag{26}$$

Thus, $(Q_{ij}(t))_{i,j=0,\dots,n-1}$ is a discrete approximation of the density $\mathfrak{q}(t,\cdot,\cdot)$ at time t.

	Density	Layer
L	infinite	$L = 3$; $\ell = 0.5$
BC	none	periodic
X_0	$-0.5, 0, 0.5$	uniform
D	$\begin{cases} D^- = 5 \\ D^+ = D^-/\rho \end{cases}$	$\begin{cases} D_0 = 5 \\ D_m = D_0/\rho \end{cases}$
ρ	$\{5, 20\}$	$[2.5 : 750]$
T	1.5	10
δt	$\{0.01, 0.001\}$	0.001
N	$\{10^5, 10^6, 2 \times 10^6\}$	2×10^6

	Bimaterial	Bimaterial absorbing	Symmetry
L	$L = 2$	$L = 2$	$L = 4$
BC	reflecting/reflecting	reflecting/absorbing	reflecting/reflecting
X_0	1.5	0.1	$(i + 1/2) \times L/200$
D	$\begin{cases} D^- = 5 \\ D^+ = D^-/\rho \end{cases}$	$\begin{cases} D^- = 5 \\ D^+ = D^-/\rho \end{cases}$	$\begin{cases} D^- = 5 \\ D^+ = D^-/\rho \end{cases}$
ρ	$[2.5 : 20]$ | 100 | $[250 : 500]$ | 750	$[2.5 : 20]$ | 100 | 250 | 500 | 750	$\{5, 15\}$
T	15 | 100 | 300 | 400	100 | 250 | 750 | 1500 | 2000	0.2
δt	0.001	0.001	0.01
N	4×10^6	2×10^6	2×10^6

Table 3: Parameters for the benchmark tests.

Symmetry. Plot the absolute value of difference $\Delta_{ij}(t) = |Q_{ij}(t) - Q_{ji}(t)|$ as function of $i, j = 0, \ldots, n-1$.

The matrix $(\Delta_{ij}(t))_{i,j=0,\ldots,n-1}$ measures of how much numerically the density $\overline{\mathfrak{q}}(t, x, y)$ of the scheme deviates from being symmetric in x and y. We call it the *asymmetry measure*. By doing so, we refine the test proposed in [26, Sect. 4.3.3].

Here, we do not quantify the statistical fluctuations of $Q_{ij}(t)$, but it appears in our numerical results that it is not crucial as for the other benchmark tests.

3.6 Parameter settings: number of particles, time step and size of the domain

The parameters we use are given in Table 3. We set $\rho = D^-/D^+$.

3.6.1 Number of particles

We simulate the dynamic of N particles until the final time T is reached.

We choose a large number of particles so that the Monte Carlo error is small. Typically, Monte Carlo simulations comes with an error in general of order $\mathrm{O}(N^{-1/2})$, with N the number of particles. In the numerical simulations, we choose from 10^5 to 4×10^6 particles, so that the Monte Carlo error ranges from the order 3.2×10^{-3} to 5×10^{-4}.

3.6.2 Time step and size of the domain

The time step and the size of the domain are closely linked.

For a given time step. Assume we fix the time step δt as a given input. This time step determines the size of the domain so as to maximize the number of passage through the interface layer by maximizing the relative size of the interface layer within the medium.

Indeed, increasing the size of the domain without changing the time step hides artificially the potential bias of the scheme which appears only when the particle is in the interface layer.

As illustration, we plot in Figure 7 the distribution of 2×10^6 particles in the steady state regime at time $T = 10$ with a scheme used in the interface layer called `SBMlin` (see Algorithm 5 in Section 4 for more details). What is plotted is really an empirical estimation of the density of X_t at time $T = 10$ ($\delta t = 0.001$), with an initial uniform distribution in a bimaterial medium with $D^- = 5$ and $D^-/D^+ = \rho$ and reflecting BC. In Figure 7(a), for a domain size of 2 and a time step $\delta t = 0.001$, the bias of the scheme is visible around the interface. In Figure 7(b), for a domain size of 50 and the same time step $\delta t = 0.001$, the bias of the scheme is not visible.

For a given domain size. Assume we fix the size of the domain size as a given input. The time step must be chosen so as to keep a large amount of particles that cross the interface layer.

A caution must be observed if the time step is decreased while leaving the medium unchanged.

We report in Figure 6, the mean proportion of the number of steps done in the interface layer $n_{\text{interface}}$ over the total number of steps n_{steps} in a fixed, bimaterial medium. This proportion decreases quickly with the time step δt. If one choose another time step, the size of the domain should be resized accordingly.

Finally, let us mention that the mean proportion of steps in the interface layer varies from 32 % for $\rho = 2.5$ to 25 % for $\rho = 750$. This slight variation could contribute to the differences in the behavior of the schemes observed as ρ increases.

Therefore, a good benchmark test dedicated to emphasize the bias of schemes must have a domain size and a time step chosen accordingly so as to maximize the number of crossing of the interfaces.

3.6.3 Size of the domain with respect to the boundary layers

In the benchmark tests, the size of the domain is also chosen so that we can change the tested schemes and the boundary conditions independently. To do so, the

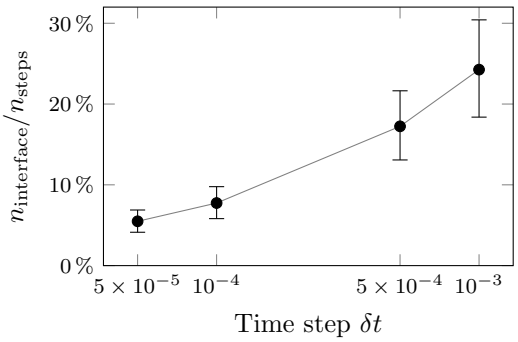


Figure 6: Influence of the time step for a given domain size $L = 2$: Mean proportion of steps performed in the interface layer as a function of the time step. Example with a bimaterial medium with $D^- = 5$, $D^+ = 0.25$ ($\rho = 20$) for the algorithm `sbm` (see Algorithm 5 in Section 4 below), $L = 2$ and one interface at $x_I = 1$, as well as reflections at both ends. The final time is $T = 10$ and the $N = 10{,}000$ particles are uniformly distributed at $t = 0$. The error bars represents the limits of the 1st and 3rd quartiles.

domain is large enough so that a particle close to a boundary does not "see" the discontinuity and a particle close to the discontinuity region does not "see" the boundaries.

To do so, the domain is split in three kinds of zones (see Figure 8):

- The interface layer I_{layer}, around a discontinuity at x_I,
- The boundary layer,
- The constant diffusivity zone.

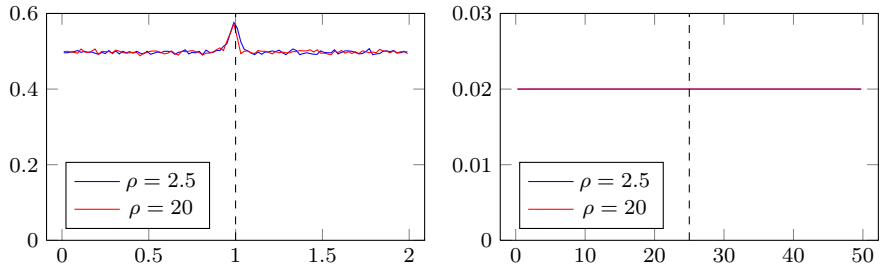
The size of the domain is chosen so that the interfaces and boundary layers fill most of the size of the domain without overlapping.

For a fixed the time step δt and a discontinuity at the position x_I, we define a zone around the interface, called the *interface layer*, as:

$$I_{\text{layer}}(x_I) = [x_I - d_\alpha\sqrt{2D^-\delta t}, x_I + d_\alpha\sqrt{2D^+\delta t}],$$

with $\mathbb{P}[|G| \leq d_\alpha] = \alpha$ for $G \sim \mathcal{N}(0,1)$. We choose $d_\alpha = 4$, corresponding to $\alpha = 1 - 6 \times 10^{-5} = 99.994\,\%$. This means that, when a particle is at a distance greater than $d_\alpha\sqrt{2D^\pm\delta t}$ from the discontinuity, it has a very small probability ($0.006\,\%$) to reach it.

Identically, the particle is within a so-called *boundary layer* if it is at the distance lower than $d_\alpha\sqrt{2D^\pm\delta t}$ from it.



(a) Size of the media: 2 (b) Size of the media: 50

Figure 7: Influence of the size of the domain for a given time step $\delta t = 0.001$: empirical estimation of the density of X_t. Example with a bimaterial medium in the steady state regime at time $T = 10$ and the algorithm `SBMlin` (see Algorithm 5 in Section 4 below) in the interface layer.

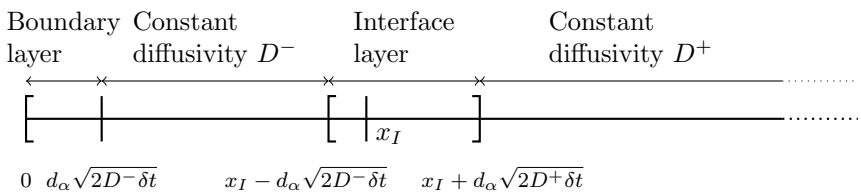


Figure 8: Three kind of zones: interface layer, boundary layer, zone of constant diffusivity

This means that, when a particle is at distance greater than $d_\alpha\sqrt{2D^\pm\delta t}$ from an interface or a boundary, we act as if the diffusivity is constant which defines the zones of *constant diffusivity*.

3.7 Algorithms

The particles are moved with a constant time step δt, with an algorithm which is function of the zones defined above.

3.7.1 Algorithms in the constant diffusivity zone

In the zone of *constant diffusivity*, we use Gaussian (see Algorithm 1) or Uniform steps (see Algorithm 2).

3.7.2 Algorithms in the boundary layer

In the *boundary layer*, the hitting time may be computed either exactly (see Algorithm 3: `ExactHittingTime`) or with a linear approximation (see Algorithm 4: `LinearHittingTimeUS` and `LinearHittingTimeGS`).

Data: The position x of the particle at time t, a time step δt and a diffusivity D.
Result: The position $X_{t+\delta t}$ at time $t+\delta t$ of the particle.
Draw a random variate $\xi \sim \mathcal{N}(0,1)$;
return $x + \sqrt{2D\delta t}\xi$;

Algorithm 1: `GaussianStep`$(x, \delta t, D)$: Gaussian step in a zone of constant diffusivity D.

Data: The position x at time t of the particle, a time step δt and a diffusivity D.
Result: The position $X_{t+\delta t}$ at time $t+\delta t$ of the particle.
Draw a random variate $U \sim \mathcal{U}(0,1)$;
return $x + \sqrt{6D\delta t}(2U-1)$;

Algorithm 2: `UniformStep`$(x, \delta t, D)$: Uniform step in a zone of constant diffusivity D.

Data: The position x at time t of the particle, a time step δt and a diffusivity D.
Result: The position (s, X_s) which is either $(\tau, 0)$ if $\tau < t+\delta t$ or $(t+\delta t, X_{t+\delta t})$, where $\tau = \inf\{s > t; X_s = 0\}$.
Draw a random variate $\xi \sim \mathcal{N}(0,1)$;
Set $z \leftarrow x/\sqrt{2D}$; `/* Normalize the position */`
Set $y \leftarrow z + \sqrt{\delta t}\xi$; `/* Try a first guess */`
if $\mathrm{sgn}(z) \neq \mathrm{sgn}(y)$ **then**
  `/* The boundary/interface has been crossed. */`
  Generate a random variate $\xi \sim \mathcal{IG}(|z|/|y|, z^2/\delta t)$;
  Set $\tau \leftarrow \delta t \times \xi/(1+\xi) + t$;
  return $(\tau, 0)$;
else
  `/* Check if the boundary/interface has been crossed. */`
  Generate a random variate $\xi \sim \mathcal{IG}(|z|/|y|, z^2/\delta t)$;
  Generate a random variate $U \sim \mathcal{U}(0,1)$;
  if $U < \exp(-2zy/\delta t)$ **then**
    `/* The boundary/interface has been crossed. */`
    Generate a random variate $\xi \sim \mathcal{IG}(|z|/|y|, z^2/\delta t)$;
    Set $\tau \leftarrow \delta t \times \xi/(1+\xi) + t$;
    return $(\tau, 0)$;
  else
    `/* The boundary/interface has not been crossed. */`
    return $(t+\delta t, y\sqrt{2D})$;
  end
end

Algorithm 3: `ExactHittingTime`$(t, x, \delta t, D)$: Exact simulation of the first hitting time of 0, where $\mathcal{IG}(\alpha, \beta)$ is the inverse Gaussian distribution of parameters (α, β) (see *e.g.* [28]).

Data: The position x at time t of the particle, a time step δt and a diffusivity D.
Result: The position (s, X_s) which is either $(\tau, 0)$ if $\tau < t + \delta t$ or $(t + \delta t, X_{t+\delta_t})$, where τ is an approximation of $\inf\{s > t; X_s = 0\}$
Draw a new position y according to `GaussianStep` $(x, \delta t, D)$ or to `UniformStep` $(x, \delta t, D)$;
if $\mathrm{sgn}(x) \neq \mathrm{sgn}(y)$ **then**
  `/* The boundary/interface has been crossed. */`
  Set $\tau \leftarrow \delta t \frac{|x|}{|x|+|y|} + t$;
  return $(\tau, 0)$
else
  `/* The boundary/interface has not been crossed. */`
  return $(t + \delta t, y)$
end

Algorithm 4: `LinearHittingTimeUS`$(t, x, \delta t, D)$ and `LinearHittingTimeGS`$(t, x, \delta t, D)$: Linear approximation of the first hitting time of 0 with a Uniform Step (US) or a Gaussian Step (GS).

3.7.3 Algorithms in the interface layer

In the interface layer $I_{\text{layer}}(x_I)$, around each interface with position x_I, comes into play the approximation method for which we would like to estimate the bias.

3.7.4 Summary of the possible algorithms in the different zones

For each particle, we generate a random sequence $\{\overline{X}_{k\delta t}\}_{k=0,1,2,\ldots}$ of approximations of $\{X_{k\delta t}\}_{k=0,1,2,\ldots}$. For this, $\overline{X}_{k\delta t}$ depends only on $\overline{X}_{(k-1)\delta t}$ according to a density $\overline{\mathfrak{q}}(\delta t, \overline{X}_{(k-1)\delta t}, \cdot)$, where $\overline{\mathfrak{q}}$ is an approximation of $\mathfrak{q}$.

Table 4 summarizes the possible algorithms according to the zone (interface layer, boundary layer, zone of constant diffusivity (see Figure 8) in which the particle is).

- Interface layer: scheme that has to be tested
- Boundary layer
  - ⋆ absorbing: either `ExactHittingTime` (Algorithm 3) or `LinearHittingTimeUS` (Algorithm 4).
  - ⋆ periodic: reinject the particle into the medium in a periodic way.
  - ⋆ reflecting: perform a reflection around the boundary point.
- Zone of constant diffusivities: either `GaussianStep` (see Algorithm 1) or `UniformStep` (see Algorithm 2).

Table 4: Possible algorithms for the three kinds of zone.

A caution must be observed when mixing the schemes as this may lead to bad

results. A general rule is to not mix Gaussian-like steps with uniform steps. An illustration is provided in Section 4.2.

4 Four numerical schemes in one dimensional discontinuous media

4.1 Algorithms

In the interface layer $I_{\text{layer}}(x_I)$, we consider four specific schemes:

- `SBM`: The exact density based, constant time step algorithm based on the exact method proposed by [28] (We warn that a normalization factor $1/2$ has been added for convenience in this reference in front of the diffusivity coefficient), see Algorithm 5.
- `Uffink`: The algorithm based on the approximation method proposed by Uffink [55], see Algorithm 6.
- `HMYLA`: The algorithm based on the approximation method proposed by Hoteit *et al.* [22], see Algorithm 7.
- `SBMlin`: A simpler version of the exact algorithm, with a linear interpolation for the time in case of crossing [28], see Algorithm 5.

Data: An initial position $X_t = x$ and a time $\delta t > 0$ in the interface layer $I_{\text{layer}}(x_I)$ of an interface at x_I.
Result: A position of $X_{t+\delta t}$ according to the SBM algorithm.
(SBM) Set $(s, y) \leftarrow$ `ExactHittingTime` $(t, x - x_I, \delta t, D^{\text{sgn}(x-x_I)})$
(SBMlin) Set $(s, y) \leftarrow$ `LinearHittingTimeGS` $(t, x - x_I, \delta t, D^{\text{sgn}(x-x_I)})$

```
if s < t + δt then
    /* A crossing occurred:  biased step                    */
    Generate a random variate U ∈∼ 𝒰(0,1);
    Generate a random variate G_2 ∼ 𝒩(0,1);
    if U < (1 + θ)/2 then
        return x_I + √(2D^+(t + δt − s)) |G_2|
    else
        return x_I − √(2D^−(t + δt − s)) |G_2|
    end
else
    /* No crossing occurred                                 */
    return x_I + y
end
```

Algorithm 5: `SBM` and `SBMlin` algorithms with the two-steps method.

From now, each method is associated to a color using the correspondence given in Table 5.

```
Data: An initial position X_t = x and a time step δt > 0 in the interface layer I_layer(x_I) of
      an interface at x_I
Result: A position X_{t+δt} according to Uffink's algorithm
Compute H_1 = √(D^- 6 δt) and H_2 = √(D^+ 6 δt);
Set z ← x − x_I ;                                        /* Shift the position */
if z + H_1 < 0 then
    /* the interface is not crossed:  uniform step                          */
    Generate a random variate V ~ U(0,1);
    return z + (2V − 1)H_1
else
    if z < 0 then
        /* The interface is crossed:  compute P_H                           */
        Compute x_L = z − H_1, x_M = −z − H_1 and x_R = (x + H_1) H_2/H_1;
        Compute P_H = 1/(2 H_1) (x_M − x_L);
    else
        if z − H_2 > 0 then
            /* the interface is not crossed:  uniform step                  */
            Generate a random variate V ~ U(0,1);
            return x + (2V − 1)H_2
        else
            /* The interface is crossed:  compute P_H                       */
            Compute x_L = x + (z − H_2) H_1/H_2, x_M = x_I − (z − H_2) and x_R = x + H_2;
            Compute P_H = (1 + θ)/(2 H_2) (x_M − x_L) with θ given by Eq. (8);
        end
    end
end
/* The interface is crossed:  biased step                                   */
Generate a random variate U ~ U(0,1);
Generate a random variate V ~ U(0,1);
if U ≤ P_H then
    return x_L + (x_M − x_L) V
else
    return x_M + (x_R − x_M) V
end
```

Algorithm 6: `Uffink` algorithm.

Data: An initial position $X_t = x$ and a time $\delta t > 0$ in the interface layer $I_{\text{layer}}(x_I)$ of an interface at x_I.
Result: A position $X_{t+\delta t}$ according to `HMYLA` algorithm
Compute $H_1 = \sqrt{D^- \, 6 \, \delta t}$ and $U_2 = \sqrt{D^+ \, 6 \, \delta t}$;
Set $z \leftarrow x - x_I$; `/* Shift the position */`
Set $(s, z_{\text{next}}) \leftarrow$ `LinearHittingTimeUS` $(t, z, \delta t, D^{\text{sgn}(z)})$;
if $s < t + \delta t$ **then**
  `/* The interface is crossed: biased step */`
  `/* Time splitting */`
  Compute $\delta t_2 = \delta t - s$;
  Generate a random variate $U \sim \mathcal{U}(0, 1)$;
  if $U < \dfrac{1-\theta}{2}$ **then**
    Compute $x_L = x_I - \sqrt{D^- \, 6 \, \delta t_2}$;
    Set $x_R = x_I$;
  else
    Set $x_L = x_I$;
    Compute $x_R = x_I + \sqrt{D^+ \, 6 \, \delta t_2}$;
  end
  Generate a random variate $W \sim \mathcal{U}(0, 1)$;
  return $x_L + (x_R - x_L) \, W$;
else
  `/* The interface is not crossed: uniform step */`
  return $x_I + z_{\text{next}}$
end

Algorithm 7: `HMYLA` algorithm.

`SBM`	$\rightsquigarrow$	Green	`HMYLA`	$\rightsquigarrow$	Coral
`Uffink`	$\rightsquigarrow$	Blue	`SBMlin`	$\rightsquigarrow$	Red

Table 5: Color convention for the schemes.

Figure 9 presents the empirical density $\overline{\mathfrak{q}}_N(\delta t, x, \cdot)$ of the schemes after one step $\delta t = 0.01$ of $N = 10^7$ particles for three values of the starting point x_0 in a medium with a diffusivity $D^- = 5$ on $[-2, 0]$ and $D^+ = 0.5$ on $(0, 2]$ and reflecting boundary conditions (such a medium is referred as a bimaterial one, see Section 3.3). It illustrates the effect of the interface layer on the density depending on the starting point.

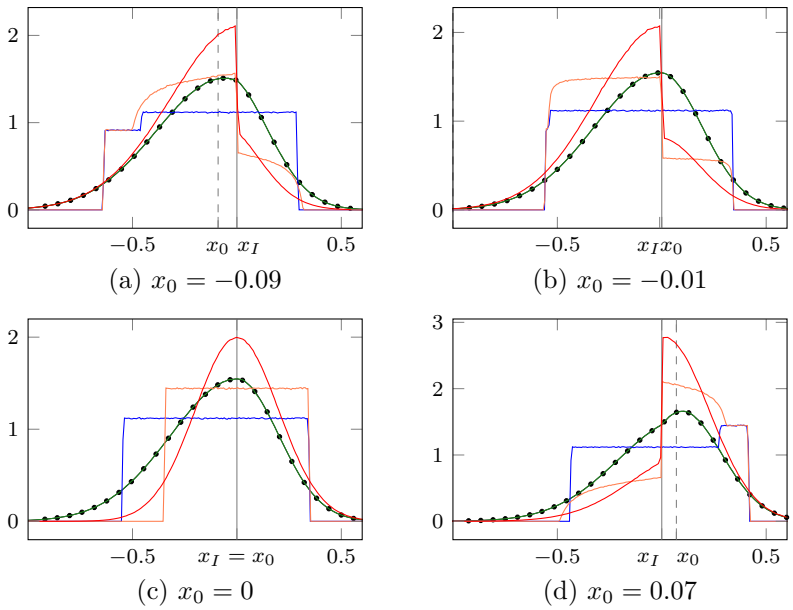
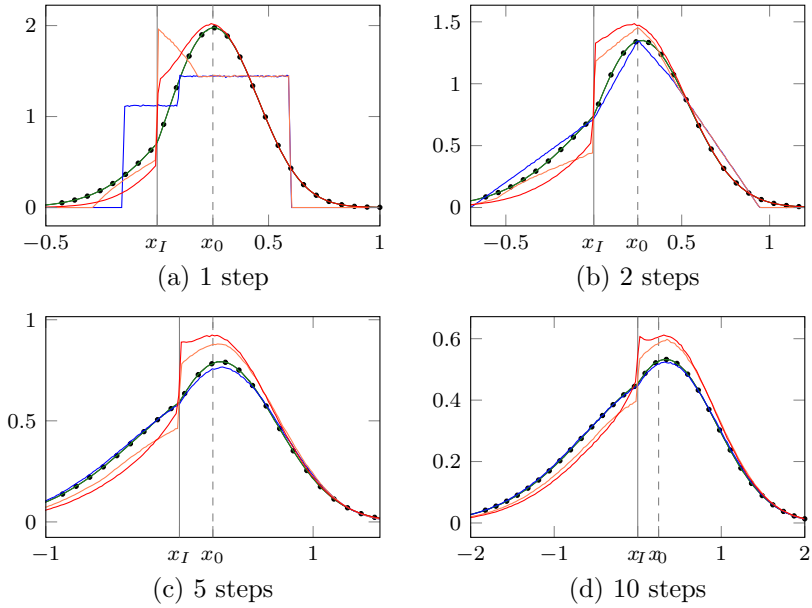


Figure 9: **After one time step for different starting points.** Density $\mathfrak{q}(t, x_0, \cdot)$ (black dots) and empirical densities $\overline{\mathfrak{q}}_N(t, x_0, \cdot)$ in an infinite bimaterial medium with $D^- = 5$, $D^+ = 2$ ($\rho = 2.5$) and an interface at $x_I = 0$. We use different starting points x_0 close to x_I. The time step is $\delta t = 0.01$. We used $N = 10^7$ particles. The colors of the lines follows the convention of Table 5. This graphs shows how the shape of the empirical densities may vary greatly with a starting point close to the interface.

In Figure 10, we plot the empirical densities of the four methods after 1, 2, 5 and 10 time steps for a starting point $x_0 = 0.25$, $N = 10^7$ particles, and a discontinuity at 0. We compare it with the exact density $\mathfrak{q}(t, x_0, y)$ with $\mathfrak{q}$ given by (7). We see that empirical densities of `SBMlin` and `HMYLA` have the same behavior near the

discontinuity, while `SBM` — as well as `Uffink` after a few time steps —, follows the right dynamic.



(a) 1 step (b) 2 steps

(c) 5 steps (d) 10 steps

Figure 10: **After a few time steps.** Density $\mathfrak{q}(t, x_0, \cdot)$ (black dots) and empirical densities $\overline{\mathfrak{q}}_N(t, x_0, \cdot)$ in an infinite bimaterial medium with $D^- = 5$, $D^+ = 2$ ($\rho = 2.5$) and an interface at $x_I = 0$. The starting point at $x_0 = 0.25$ and the time step $\delta t = 0.01$. We used $N = 10^7$ particles. The colors of the lines follows the convention of Table 5.

4.2 Combination of algorithms in the whole domain

Table 6 summarizes, for each algorithm tested in the interface layer, the recommended combination with the algorithms proposed in Table 4 of Section 3.7.4 (see Figure 8).

Other choices could have been performed. However, this leads to bad results. In Figure 11, we show the effect of mixing the `HMYLA` scheme in the interface layer with the `GaussianStep` outside the interface layer in a bimaterial medium. With a uniform repartition of the particles and reflected boundary conditions at 0 and 2, the particle shall remain uniformly distributed all over the medium. This is not the case around the limits of the interface layers. Similar results occurs with other ways of mixing the schemes.

Since `Uffink` and `HMYLA` rely on uniform approximations, they should be coupled

	SBM	Uffink	HMYLA	SBMlin
Interface layer	Algorithm 5	Algorithm 6	Algorithm 7	Algorithm 5
Boundary layer absorbing (if any)	`ExactHittingTime` Algorithm 3	`LinearHittingTimeUS` Algorithm 4	`LinearHittingTimeUS` Algorithm 4	`LinearHittingTimeGS` Algorithm 4
Constant diffusivity zone	`GaussianStep` Algorithm 1	`UniformStep` Algorithm 2	`UniformStep` Algorithm 2	`GaussianStep` Algorithm 1

Table 6: Combination of the algorithms for the three kinds of zone.

with schemes relying on uniform approximations to avoid bad behavior when the particle is moved from one zone to another.

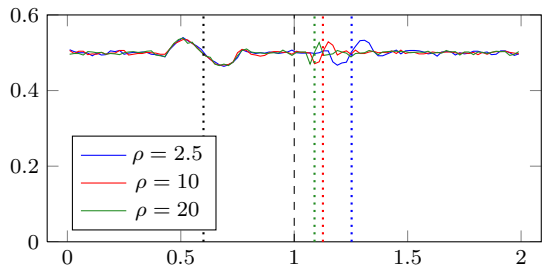
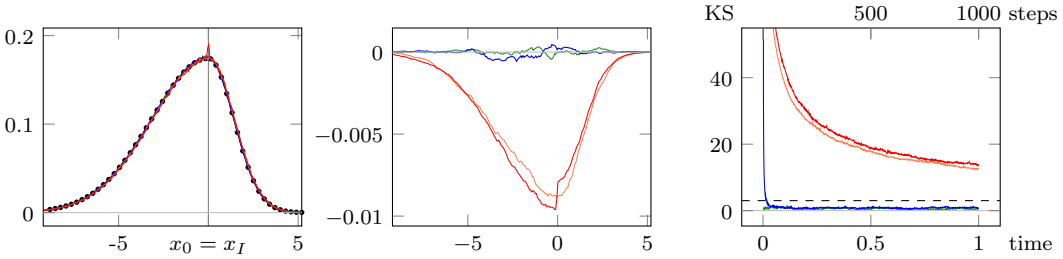


Figure 11: Histograms of the positions of 2×10^6 particles at time $T = 10$ ($\delta t = 0.001$) in a bimaterial medium with reflected BC for `HMYLA` coupled with `GaussianStep` outside the interface layer. The dotted lines represent the limits of the interface layers for three values of ρ, while the dashed line represents the interface.

5 Numerical results

In this section, we apply the benchmark tests described in Section 3 to the four schemes `SBM`, `Uffink`, `HMYLA` and `SBMlin` described the Section 4. We draw conclusions on the behavior of each method in the steady state and transient regime.

Unless stated, we use the parameters given in Table 3.



(a) Densities $\overline{\mathfrak{q}}_N$ and $\mathfrak{q}$ (b) $\overline{F}_N - F$ (c) $\sqrt{N} d_{\mathrm{KS}}(\overline{F}_N, F)$

$D^- = 5$, $\rho = 5$, $x_I = x_0 = 0$, $\delta t = 0.001$, $T = 1$, $N = 2 \times 10^6$

Figure 13: Plot of (a) the densities $\overline{\mathfrak{q}}_N$ and $\mathfrak{q}$, (b) the difference between $\overline{F}_N$ and F, and (c) the normalized Kolmogorov-Smirnov distance M_{KS} between the approximated and the true DF at time $T = 1$.

5.1 Benchmark test Density: Numerical results

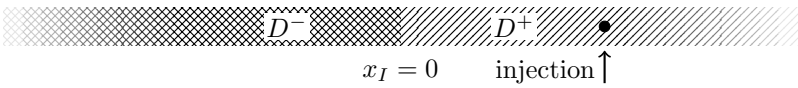
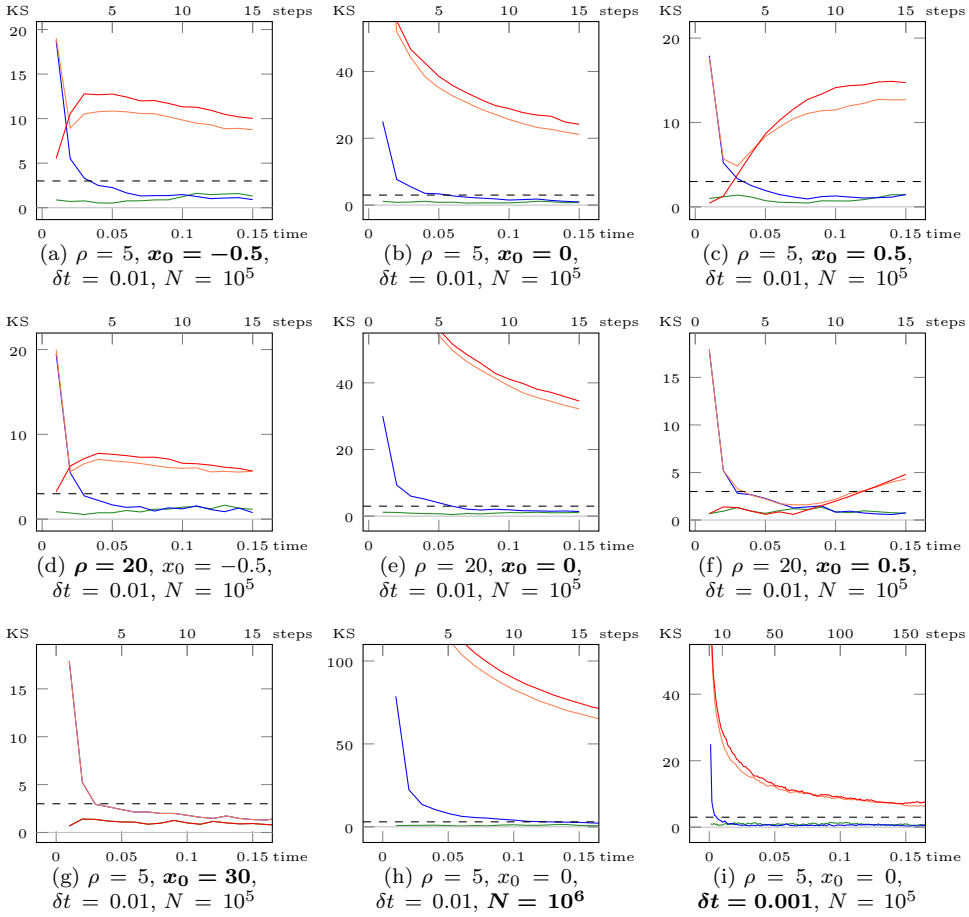


Figure 12: Medium for the Density test.

Density. Plot M_{KS} defined by (11) as a function of t and compare it with a threshold c_α of the Kolmogorov-Smirnov statistics for a confidence level α, or to the value $c_\alpha = 3$ which corresponds to a level of risk $1 - \alpha \approx 3 \times 10^{-7}$.

First, we represent on Figure 13(a) the empirical densities of the four schemes after 1000 time steps of 2×10^6 of particles ($\delta t = 0.001$). Qualitatively, `SBM`, `Uffink` and `HMYLA` seem to converge to the exact density (black dots). `SBMlin` presents a small peak at the discontinuity located at zero. However this sole qualitative comparison is not enough to quantify the bias of schemes. On Figure 13(b), we plot the differences between the approximated and the true DF. Now the bias of `HMYLA` and `SBMlin` is revealed. Figure 13(c) shows the evolution of the normalized Kolmogorov-Smirnov distance $M_{\mathrm{KS}} := \sqrt{N} d_{\mathrm{KS}}(\overline{F}_N, F)$ with time. `SBM` is always close to the true density, `Uffink` converges rather quickly to the true one. On the other hand, for `SBMlin` and `HMYLA` — which share the same interpolation technique for computing the first hitting time of the interface — M_{KS} is far above 3.

Now we represent on Figure 14 the evolution with time (and with number of steps) of the normalized Kolmogorov-Smirnov distance $M_{\mathrm{KS}} := \sqrt{N} d_{\mathrm{KS}}(\overline{F}_N, F)$ for several choices of the input parameters.



(a) $\rho = 5$, $\boldsymbol{x_0 = -0.5}$, $\delta t = 0.01$, $N = 10^5$

(b) $\rho = 5$, $\boldsymbol{x_0 = 0}$, $\delta t = 0.01$, $N = 10^5$

(c) $\rho = 5$, $\boldsymbol{x_0 = 0.5}$, $\delta t = 0.01$, $N = 10^5$

(d) $\boldsymbol{\rho = 20}$, $x_0 = -0.5$, $\delta t = 0.01$, $N = 10^5$

(e) $\rho = 20$, $\boldsymbol{x_0 = 0}$, $\delta t = 0.01$, $N = 10^5$

(f) $\rho = 20$, $\boldsymbol{x_0 = 0.5}$, $\delta t = 0.01$, $N = 10^5$

(g) $\rho = 5$, $\boldsymbol{x_0 = 30}$, $\delta t = 0.01$, $N = 10^5$

(h) $\rho = 5$, $x_0 = 0$, $\delta t = 0.01$, $\boldsymbol{N = 10^6}$

(i) $\rho = 5$, $x_0 = 0$, $\boldsymbol{\delta t = 0.001}$, $N = 10^5$

Figure 14: Normalized Kolmogorov-Smirnov distance M_{KS} for $D^- = 5$, $\rho := D^-/D^+ \in \{5, 20\}$ and a starting point $x_0 \in \{-0.5, 0, 0.5\}$. The colors of the lines follows the convention of Table 5. When $\delta t = 0.01$, the interface layer is $[-1.26, 0.56]$ for $\rho = 5$, and $[-1.26, 0.28]$ for $\rho = 20$. We used $N = 10^5$ particles. In the graph related to the Kolmogorov-Smirnov statistics, the dashed line has for abscissa 3, a choice justified in Section 3.1.2.

The method `SBM` is always close to the true density, whatever the parameters inputs.

In the interface layer, Figures 14(b,e,h,i) confirm that `Uffink` converges to the true density after a few time steps and that `SBMlin` and `HMYLA` do not.

Increasing the number of steps (or equivalently decreasing δt as shown on Figures 14(b,i)) decreases the bias of `Uffink`, `HMYLA` and `SBMlin`. The bias of `Uffink` will decrease to zero with the increase of number of steps (or decrease of δt) while the bias of `HMYLA` and `SBMlin` will not decrease to zero.

In the zones of constant diffusivity (Figures 14(g)), `SBM` and `SBMlin` are superimposed exactly as they move the particles with the same Gaussian dynamics. They replicate the exact dynamics of the particle, up to events of exponentially small probability (see the discussion on the interface layer above). Similarly, `Uffink` and `HMYLA` are superimposed as they both move the particles with a uniform step. They replicate the exact dynamics after a few time steps.

The change of zones is illustrated on Figures 14(a,c,d,f). The bias of `SBMlin` and `HMYLA` is revealed while more particles enter the interface layer (increase of M_{KS}).

Increasing the number of particles increases the bias of the schemes as illustrated on Figure 14(h).

This DENSITY benchmark test is important as it reveals the bias of schemes. However it does not provide any information on the effects of the bias on the computations of micro- and macroscopic quantities of interest. Moreover it does not allow to distinguish between `SBMlin` and `HMYLA` schemes.

5.2 Benchmark test LAYER: Numerical results

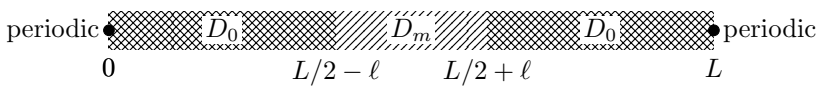


Figure 15: Medium for the LAYER test.

LAYER. For a time t large enough, plot $K_N(t,y)$ defined by (13) and compare it with $y \in [0,1] \mapsto \pm d_\alpha \sqrt{y(1-y)}$ for a confidence level α with d_α defined by (9).

In Figure 16, we plot $K_N(t,\cdot)$ for the values of $\rho = D_0/D_m$ given by Table 3.

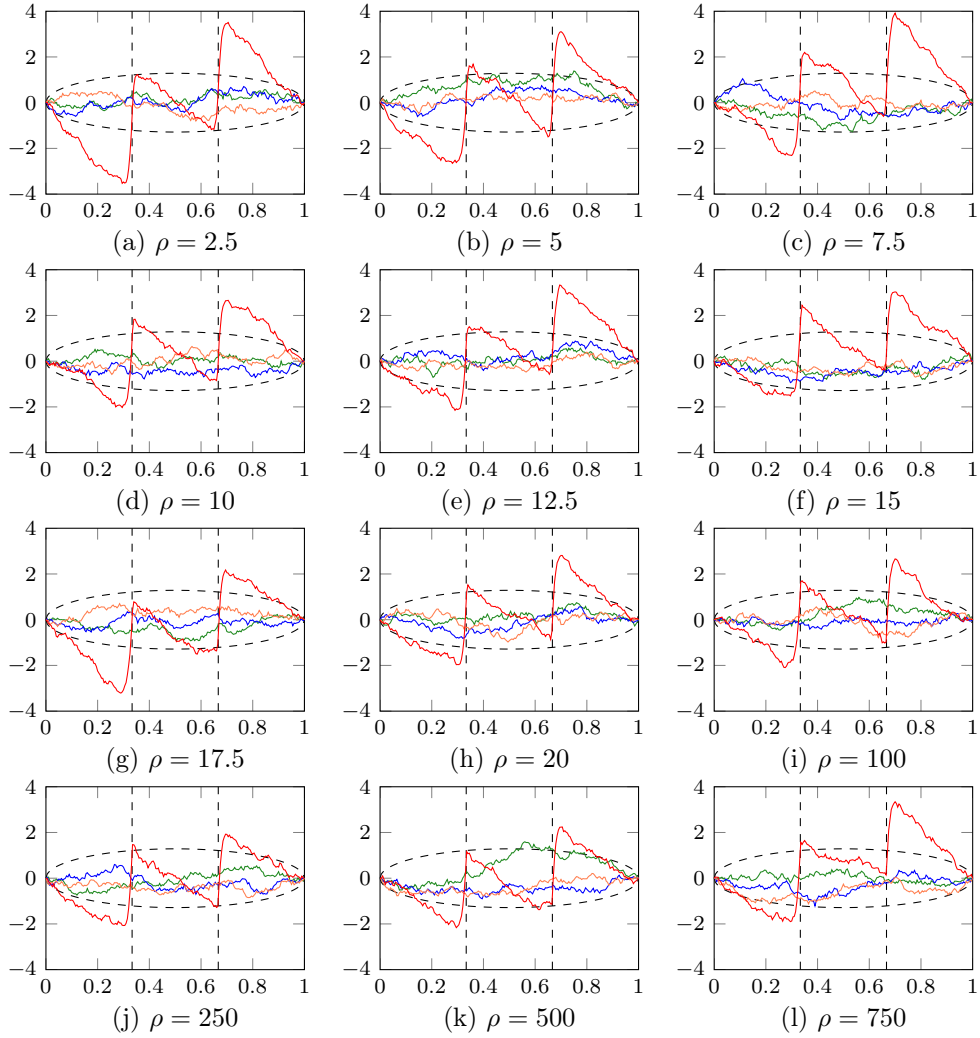


Figure 16: LAYER — Plot of $x \mapsto K_N(t,x)$ for the four methods: SBM, Uffink, HMYLA and SBMlin and comparison with the 99 %-confidence band $y \mapsto \pm d_{0.99}\sqrt{y(1-y)}$ at $T = 10$.

SMBlin method fails in preserving the uniform distribution of the particles with time. On the contrary, SBM, Uffink and HMYLA methods pass this test.

5.3 Benchmark test BIMATERIAL: Numerical results

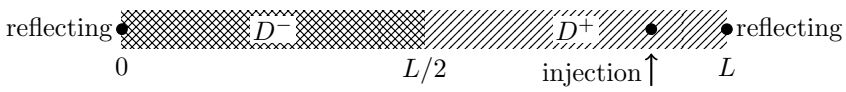


Figure 17: Medium for the BIMATERIAL test.

Bimaterial. For a level of confidence α close to 1, check that for $t \in [0, T]$,

$$\sqrt{N}(\overline{P}_N(t) - p_+(t)) \in [-d_\alpha \sqrt{p_+(t)(1-p_+(t))}, d_\alpha \sqrt{p_+(t)(1-p_+(t))}]$$

with d_α given by (9), $\overline{P}_N(t)$ given by (17) and $p_+(t)$ given by (18).

The evolution of the positive probability $p_+(t)$ and its estimation is shown Figure 18 for $\rho = 20$. We see that all the schemes respect the global behavior of $p_+(t)$ and fluctuates when the steady state regime is reached. However, in the transient regime, `HMYLA` and `SBMlin` are slightly out of the confidence interval, as seen in the right plot of Figure 18. Moreover, for `SBMlin`, $\overline{P_N}(t)$ fluctuates around a value which is smaller than $1/2$, which means an incorrect repartition of the particles.

For each ratio, we plot in Figure 19 the normalized difference $t \mapsto \sqrt{N}(\overline{P_N}(t) - p_+(t))$ as well as the 99 %-confidence band. In the steady state regime, for `SBM`, `Uffink` and `HMYLA`, the fluctuations of $\sqrt{N}(\overline{P_N}(t) - p_+(t))$ lies in the confidence interval, which means that no bias could be distinguished from the Monte Carlo error. For `SBMlin`, the fluctuations are in the right order, but there are in average more particles in the left-hand side of the medium, where the diffusivity is higher. In the transient regime, `HMYLA` and `SBMlin` do not perform well. They over estimate the number of particles staying at the right side of the interface. However, the error is rather small. This over estimation is due to the phenomena observed in Figure 2 for the Layer test close to the interface. The symmetry benchmark test also gives some insight about this (see Section 5.6).

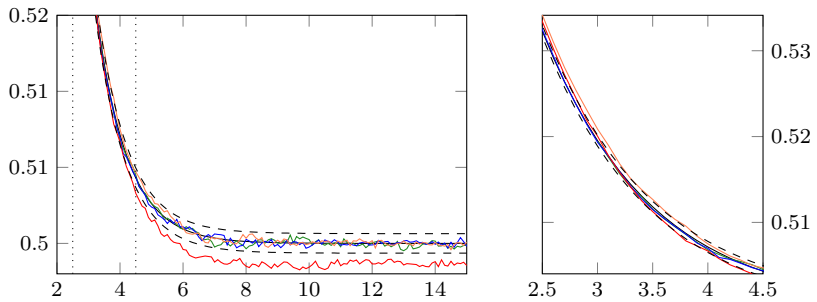


Figure 18: Bimaterial — Plots of $t \mapsto p_+(t)$ (black), 99 %-confidence band $t \mapsto p_+(t) \pm N^{-1/2} d_{0.99} \sqrt{p_+(t)(1-p_+(t))}$ (dashed) and $t \mapsto \overline{P}_N(t)$ for $\rho = 20$ and the four methods: `SBM`, `Uffink`, `HMYLA` and `SBMlin`. The plot on the right is a zoom for $t \in [2.5, 4.5]$.

Although there is no benchmark test attached to it, we plot in Figure 20 how evolves the fraction of particles that cross the interface from the right part to the left part of the medium during a single time step δt. This quantity is δt times the flow of particles from the right to the left.

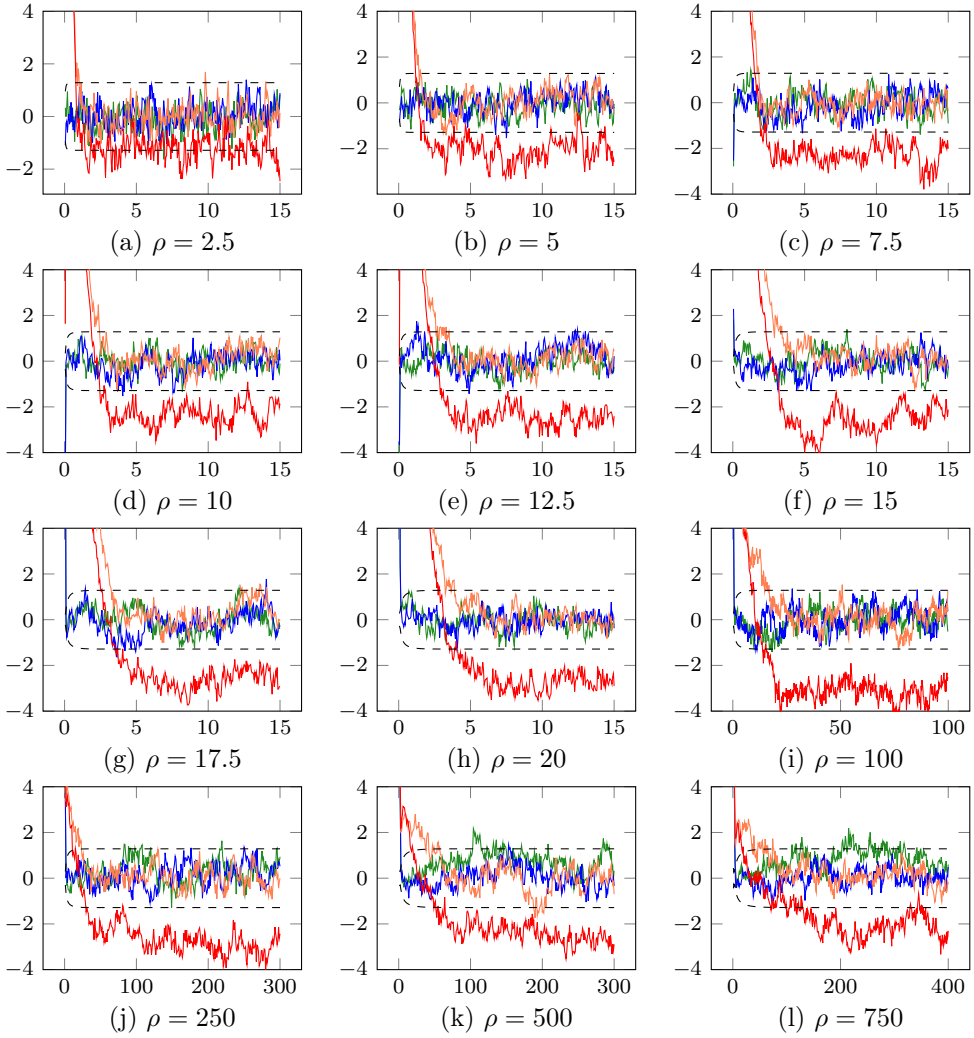


Figure 19: BIMATERIAL — Plot of $t \mapsto \sqrt{N}(\overline{P}_N(t) - p_+(t))$ for the four methods: SBM, Uffink, HMYLA and SBMlin. The dotted lines represent the 99 %-confidence band $t \mapsto \pm d_{0.99}\sqrt{p_+(t)(1-p_+(t))}$.

We have also checked that the same fraction of particles go from the right to the left and from the left to the right once the steady state regime is reached. However the flow of particles for HMYLA and SBMlin is around half of the flow of particles for SBM and Uffink. Since SBM is a method in which the particles are replaced according to the exact dynamic [28] (up to the error quantified by the interface layer), Uffink slightly overestimates the flow of particles, while HMYLA and SBMlin underestimate this quantity. This could explain why for these two methods, $\overline{P}_N(t)$ differs significantly from $p_+(t)$ in the transient regime (see Figure 18 and Figure 19).

However, for large values of ρ, all the values tend to converge quickly towards a very small probability of passage.

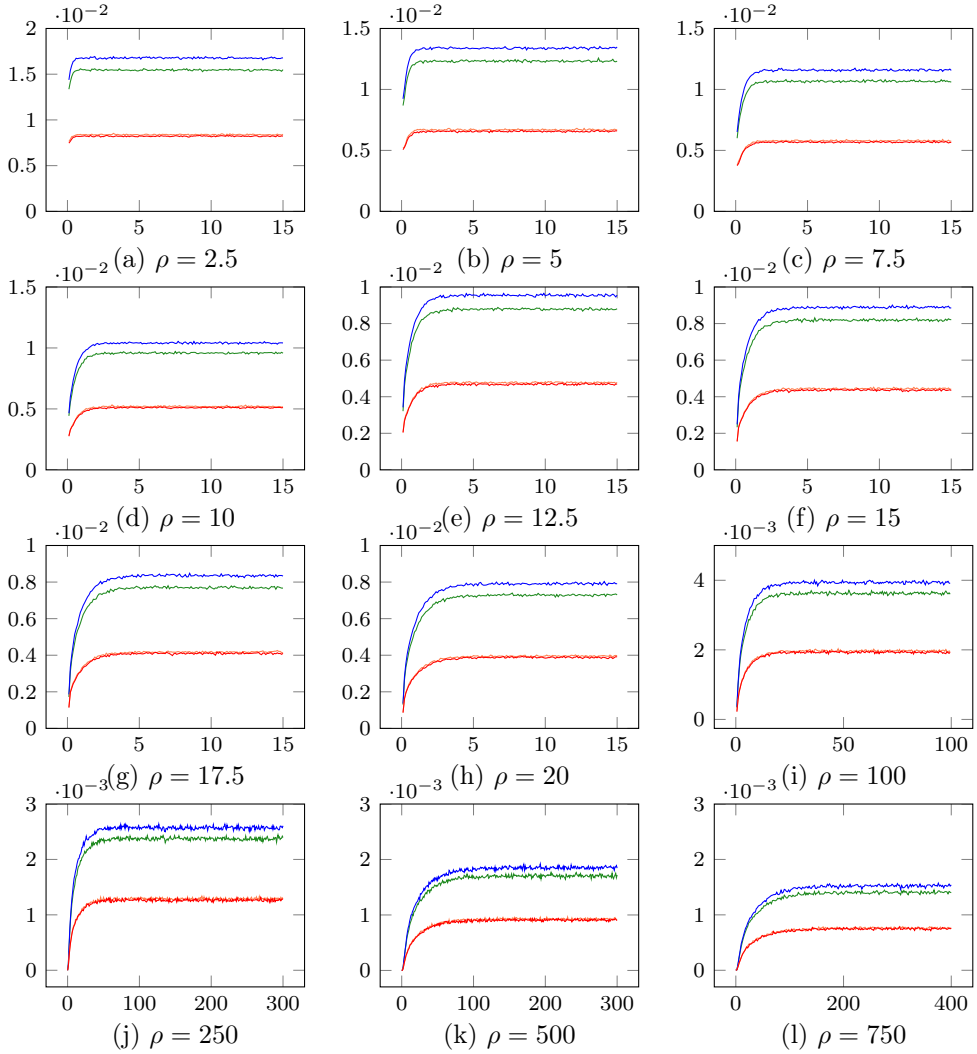


Figure 20: BIMATERIAL — Fraction of the particles going from the right part to the left part over all the particles moving between t and $t + \delta t$ for the four methods: SBM, Uffink, HMYLA and SBMlin

.

5.4 Benchmark test Bimaterial absorbing I: Numerical results

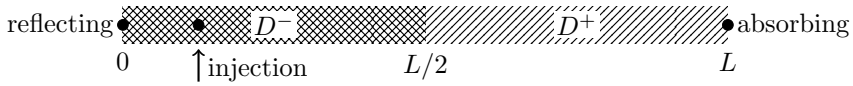


Figure 21: Medium for the Bimaterial absorbing test.

Bimaterial absorbing I. For a confidence level α, check if $\sqrt{N}(\overline{G}_N(t)-G(t))$ belongs to the confidence band

$$t \mapsto \pm d_\alpha \sqrt{G(t)(1-G(t))},$$

with $G(t)$ given by (24) and $\overline{G}_N$ be the empirical distribution function of the first exit time τ.

We plot in Figure 22 the difference between the true and the empirical DF of τ. Excepted for `SBM+ExactHittingTime`, the difference $\sqrt{N}(\overline{G}_N(t) - G(t))$ is too important to come from the Monte Carlo error. This is not surprising. From their very construction, the `LinearHittingTimeUS` and `LinearHittingTimeGS` methods overestimate the first exit time τ as indicated by the negative profiles. Indeed, with these methods, the first guess is accepted as a new position if it remains in the domain. This neglects the event that the particle may leave the domain during the time step. Nevertheless, since `Uffink` and `HMYLA` are combined with the same method `LinearHittingTimeUS`, this benchmark test demonstrates that the schemes at the interface have really an influence of the whole quality of the breakthrough curve and that `Uffink` performs better than `HMYLA`. Better schemes than `LinearHittingTimeUS` may be sought for dealing with the exit time. However, caution must be observed when designing such schemes as mixing schemes in the different zones may introduce additional errors as illustrated in Section 4.2. The way we combined the algorithms in Table 6 follows this recommendation.

Let us note also that the deviation of $\overline{G}_N$ from G is less important for large ratios.

5.5 Benchmark test Bimaterial absorbing II: Numerical results

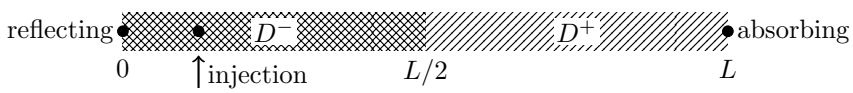


Figure 23: Medium for the Bimaterial absorbing test.

Bimaterial absorbing II. Estimate $-\widetilde{\lambda}_0^2$ defined by (25) from a linear least squares procedure on $t \in [T_-, T_+] \mapsto \log(1-\overline{G}_N(t))$, where $\overline{G}_N$ is the empirical

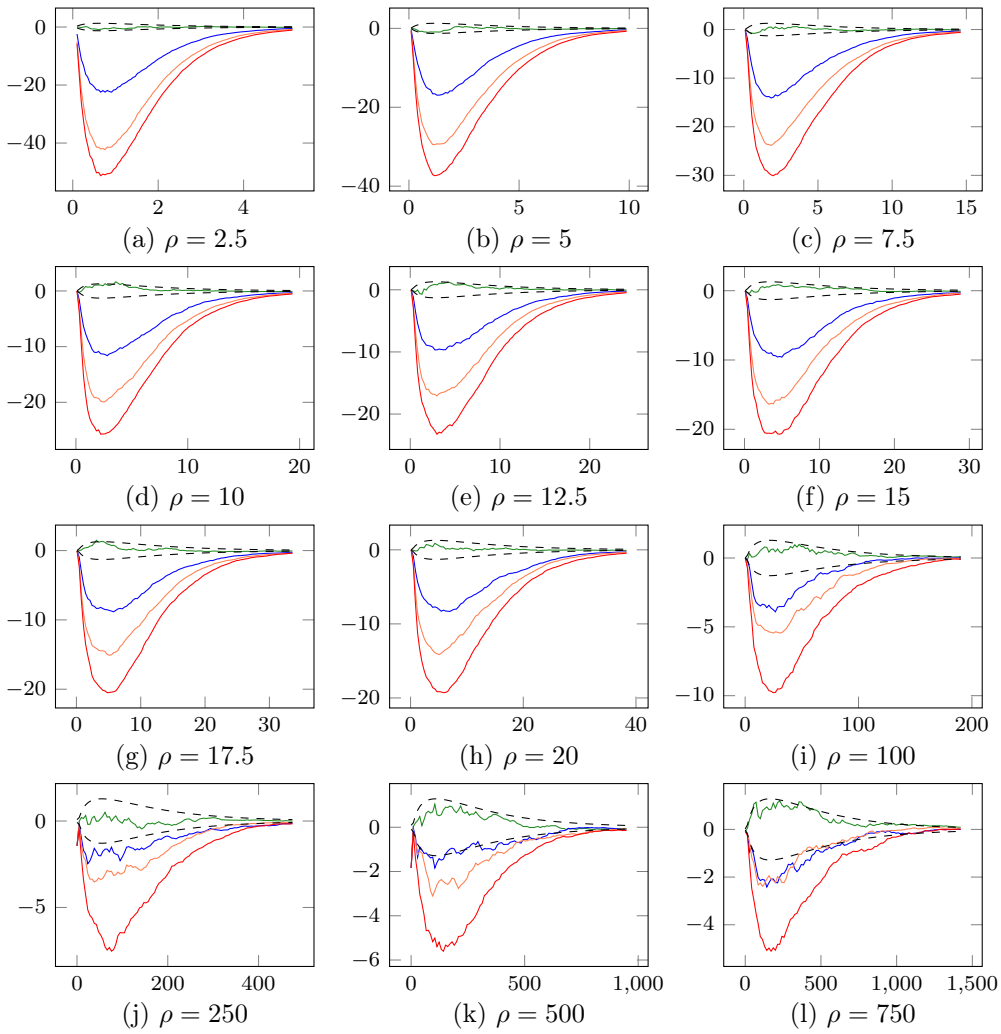


Figure 22: BIMATERIAL ABSORBING I — Plots of $\sqrt{N}(\overline{G}_N(t) - G(t))$ for the true DF G of the exit time τ and the empirical DF of the exit time using the schemes `SBM+ExactHittingTime`, `Uffink+LinearHittingTimeUS`, `HMYLA+LinearHittingTimeUS` and `SBMlin+LinearHittingTimeGS`. The dashed lines represents the 99 % confidence band $t \mapsto \pm d_{0.99}\sqrt{G(t)(1-G(t))}$.

distribution function of the first exit time from the domain by the absorbing boundary using one of the scheme.

For each method and our choice of ratios, $t \mapsto \log(1-\overline{G}_N(t))$ is plotted in Figure 24, and relative errors of the estimation of $-\widetilde{\lambda}_0^2$ are given.

All the methods exhibit the predicted exponential tail behavior of the DF of the first exit time, as shown by the linear aspect of the graphs. Besides, they all provide

a similar behavior.

`SBM+ExactHittingTime` performs the best and provides us with an accurate estimation of the exponential rate. The other combined algorithms slightly underestimate this quantity, especially for small ratio, as indicated by the negative percentage on Figure 24. It means that the particles are killed less quickly as they should be for `Uffink+LinearHittingTimeUS`, `HMYLA+LinearHittingTimeUS` and `SBMlin+LinearHittingTimeUS`. This is in agreement with the negative profiles of the plot of $\sqrt{N}(\overline{G}_N(t) - G(t))$ on Figure 22 for these three combined algorithms.

At small ratios (up to 100), the values of the percentages also confirms that `Uffink` performs better than `HMYLA` and `SBMlin`. At large ratio, it is more difficult to differentiate between the different methods. This phenomena was already observed in Bimaterial absorbing I.

5.6 Benchmark test Symmetry: Numerical results

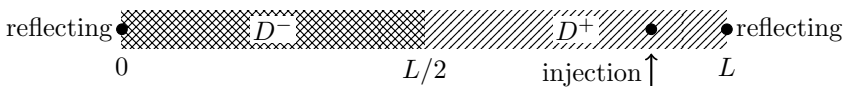


Figure 25: Medium for the Symmetry test.

Symmetry. Plot the absolute value of difference $\Delta_{ij}(t) = |Q_{ij}(t) - Q_{ji}(t)|$ as function of $i, j = 0, \ldots, n-1$, with $(Q_{ij}(t))_{i,j=0,\ldots,n-1}$ defined by (26) as a discrete approximation of the density $\mathfrak{q}(t, \cdot, \cdot)$ at time t.

The discrete density $Q_{ij}(T)$ and the symmetric difference $\Delta_{ij}(T)$ for the four scheme are plotted as function of (i, j) in Figures 26 and 27 at times $T = 0.01$ (after 1 time step) and $T = 0.2$ (after 20 time steps).

To estimate the asymmetry, we also compute the sup-norm $\|\Delta(t)\|_\infty := \max_{i,j=1,\ldots,n} |\Delta_{ij}(t)|$ as well as the L^2-norm $\|\Delta(t)\|_2 := \frac{n}{L}\left(\sum_{i,j=1,\ldots,n} |\Delta_{ij}(t)|^2\right)^2$.

The symmetry is well preserved by the exact density-based scheme `SBM`, even after many time steps. After one time step, `Uffink` performs better than `HMYLA` and `SBMlin`. However, after two time steps, the symmetric difference is mainly visible only for `SBMlin`. For `SBM`, `Uffink` and `HMYLA`, both the sup-norm and the L^2-norm is close to 0, which implies that the symmetry property is well preserved.

This could explain the bad performance of `SBMlin` in regards to the other schemes for our benchmark tests Bimaterial and Layer.

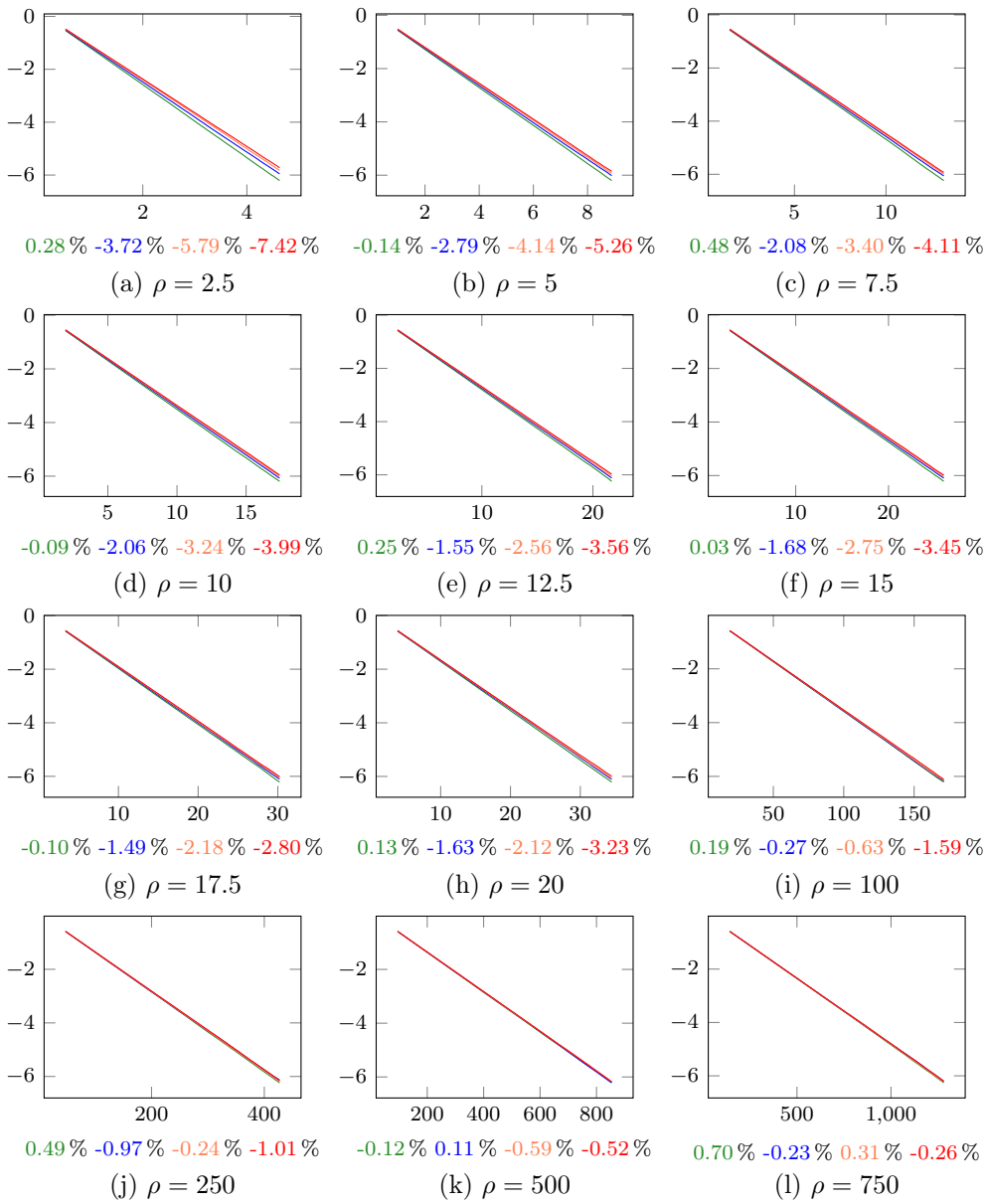


Figure 24: BIMATERIAL ABSORBING II — Plot of $t \mapsto \log(1 - \overline{G}_N(t))$ for $t \in [T/10, 9T/10]$. The relative error $(\lambda_0^2 - a)/\lambda_0^2$, where $-a$ is the slope of $t \in [T/10, 9T/10] \mapsto \log(1 - \overline{G}_N(t))$ estimated by a least squares procedure, is reported at the bottom of each graph for each of the methods `SBM+ExactHittingTime`, `Uffink+LinearHittingTimeUS`, `HMYLA+LinearHittingTimeUS` and `SBMlin+LinearHittingTimeGS` using the color code of Table 5.

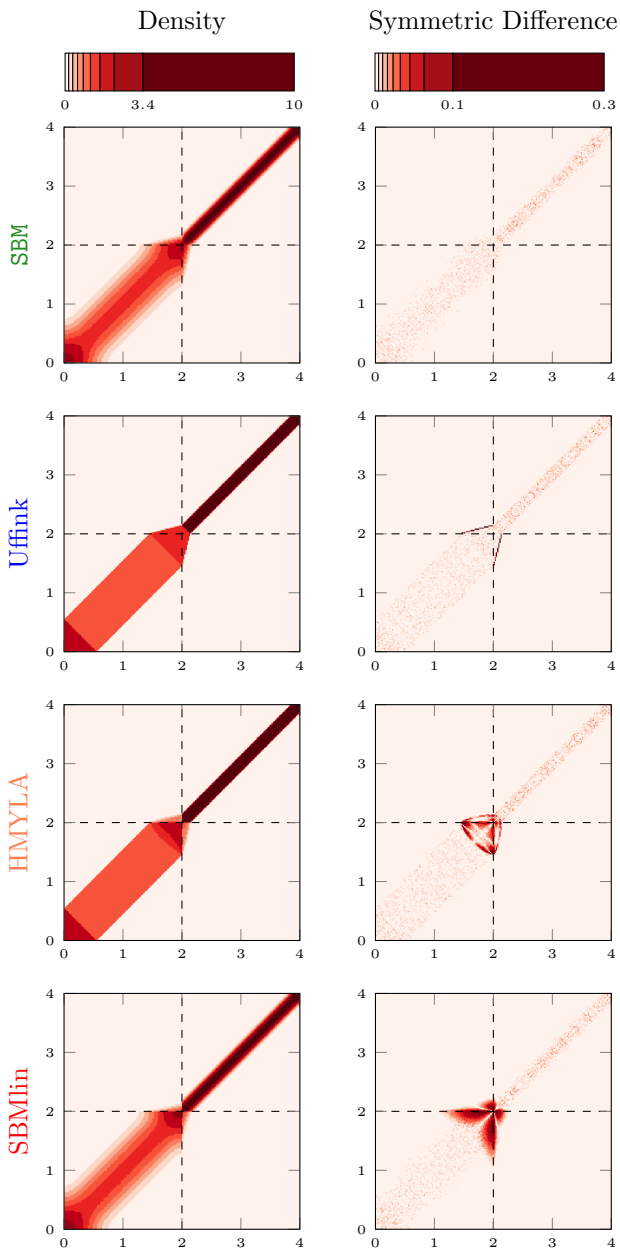


Figure 26: SYMMETRY — Discrete density $(Q_{ij}(\delta t))_{i,j=0,\ldots,n-1}$ and asymmetry measure $(\Delta_{ij}(\delta t))_{i,j=0,\ldots,n-1}$ after one time step $\delta t = 0.01$ in a bimaterial medium with $D^- = 5$ and $D^+ = 1/3$ ($\rho = 15$), an interface at $x_I = 2$ and reflecting boundary conditions at $x = 0$ and $x = 4$. To get higher contrasts, we used a discrete, non-uniform palette.

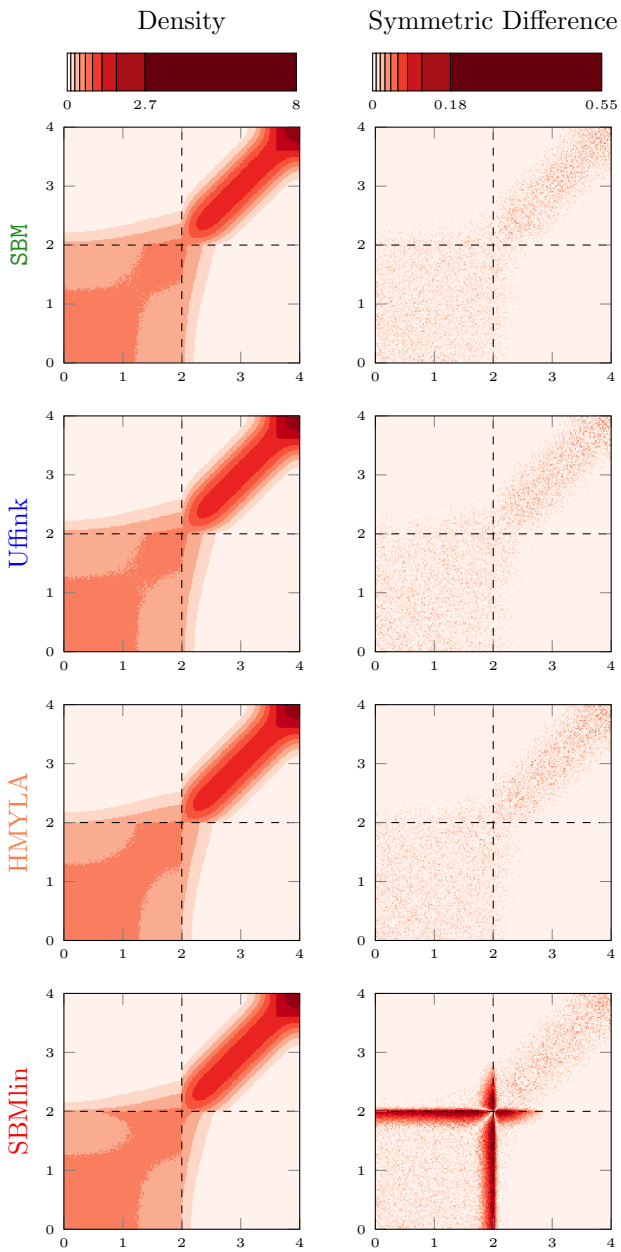
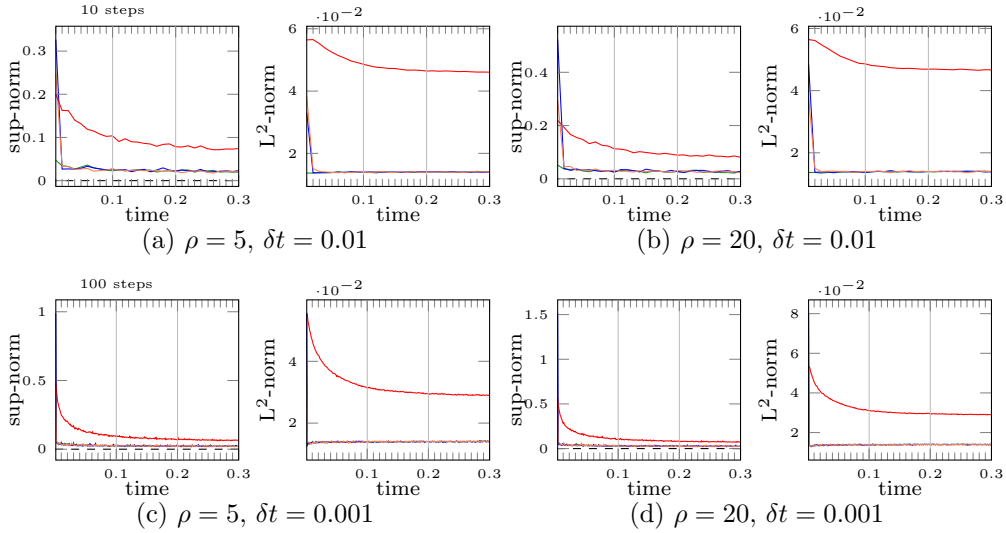


Figure 27: SYMMETRY — Discrete density $(Q_{ij}(T))_{i,j=0,\ldots,n-1}$ and asymmetry measure $(\Delta_{ij}(T))_{i,j=0,\ldots,n-1}$ after 20 time steps ($T = 0.2$) with $\delta t = 0.01$, in a bimaterial medium with $D^- = 5$ and $D^+ = 1/3$ ($\rho = 15$), an interface at $x_I = 2$ and reflecting boundary conditions at $x = 0$ and $x = 4$. To get higher contrasts, we used a discrete, non-uniform palette.



(a) $\rho = 5$, $\delta t = 0.01$ (b) $\rho = 20$, $\delta t = 0.01$

(c) $\rho = 5$, $\delta t = 0.001$ (d) $\rho = 20$, $\delta t = 0.001$

Figure 28: SYMMETRY — Plot of $k \mapsto \|\Delta(t)\|$ for the sup-norm and the L^2-norm with $\delta t = 0.01$ and $\delta t = 0.001$ in a bimaterial medium with $D^- = 5$ and (a) $D^+ = 1$ ($\rho = 5$) or (b) $D^+ = 1/3$ ($\rho = 15$), an interface at $x_I = 2$ and reflecting boundary conditions at $x = 0$ and $x = 4$. We draw the result for the methods `SBM`, `Uffink`, `HMYLA` and `SBMlin` using the color code of Table 5. The relaxation rate is much more faster for $\delta t = 0.001$ as it actually depends on the number of steps performed by the walk.

6 Conclusion

We have set up benchmark tests for Monte Carlo algorithms simulating diffusion in one-dimensional discontinuous media. These tests are physically and numerically relevant and backed by a statistical methodology to discriminate between the possible bias and the Monte Carlo error. They give the fine behavior of schemes and do not aim at being realistic. Indeed, the number of particles is chosen very large (from 10^5 to 4×10^6 particles) so that the Monte Carlo error is small. These tests could be used to motivate the choice for an algorithm in more realistic media depending on the accuracy required on the micro- or macro- quantities of interest. We have considered simple media, but our methodology extends well to media with multiples layers. Moreover the LAYER and BIMATERIAL benchmark tests could be generalized to higher dimensions.

The DENSITY benchmark test reveals the bias of the scheme. One can not distinguish the bias from the Monte-Carlo error (which is very small according to the number of particles used) for `SBM` and, after a few time steps, for `Uffink`. This fact should

be appreciated in regards of their respective computational times (see A). The number of steps required for `Uffink` to converge to the true density is linked to the convergence of uniform random variables to Gaussian ones as explained in B. On the other hand, `SBMlin` and `HMYLA` — which share the same interpolation technique for computing the first hitting time of the interface — always have a bias. Considering a linear time splitting for interfaces induces some systematic errors. This fact was already observed in [4, § [90]]. The DENSITY test case does not allow to differentiate between `SBMlin` and `HMYLA`. Moreover studying solely the density of the schemes with a fixed starting point does not reflect the global behavior of the schemes. This global behavior is rather emphasized by the four other benchmark tests, namely the LAYER, BIMATERIAL, BIMATERIAL ABSORBING and SYMMETRY benchmark tests.

The scheme `SBMlin` fails all the tests and should then be ruled out.

In the steady state regime, `SBM`, `Uffink` and `HMYLA` can be chosen as they all give a correct global repartition of the particles after many steps. However `HMYLA` should be used with caution as it under estimates the flow of particles at the interface as indicated by the BIMATERIAL test case.

In the transient regime, `SBM` performs the best as it provides a correct rate of convergence and it preserves the symmetry of the density. It should be preferred for applications requiring fine behaviors (mass flow, positions of particles close to the interface, breakthrough curve, ...) as for example, when the particles are associated to chemical species that react with the medium or interact with each other. Then a preference goes to `Uffink` which provides better results than `HMYLA` regarding the flow of particles at the interface and the repartition of particles in the domain, as shown in the BIMATERIAL test case. The rates of convergence towards the steady state regime are correct for `Uffink` and `HMYLA` as shown in the BIMATERIAL ABSORBING but again `Uffink` provides better results than `HMYLA`. Nevertheless, they both over estimate the first exit time mainly because they are combined with `LinearHittingTimeUS`. With absorbing BC, we recommend to look for a better scheme than the simple `LinearHittingTimeUS` with the caution that the schemes are difficult to mix. Namely when using `Uffink` or `HMYLA`, one should rather use uniform random variables despite it adds some slight errors.

Our empirical finding is that a good scheme should respect as much as possible the symmetry of the density transition function. `SBM` preserves well the symmetry property. `Uffink` performs slightly better than `HMYLA` regarding the symmetry property after one step. After a few time steps, both `Uffink` and `HMYLA` preserve the symmetry property, contrary to `SBMlin`. This explains why `HMYLA` performs better than `SBMlin` on all the benchmark tests. The symmetry property is therefore very important to preserve. This should be probably true in higher dimension and

could then guide the design of new schemes.

A future work is to apply those benchmark tests to other schemes available in the literature as the one proposed in [4]. Another forthcoming work is to properly define new benchmark tests with the same statistical methodology in presence of an advection term, where the infinitesimal generator of X is self-adjoint, but not with respect to the Lebesgue measure.

A Computational cost

We report here the average computational cost for one-step of each method, when the starting point is drawn uniformly on the interface layer, and the ratio ρ is chosen randomly. We also compare it with the time for `UniformStep` (see Algorithm 2) and `GaussianStep` (see Algorithm 1).

`UniformStep`	`GaussianStep`	`SBM`	`Uffink`	`HMYLA`	`SBMlin`
30	35 ($\times$1.1)	43 ($\times$1.4)	31 ($\times$1)	31 ($\times$1)	37 ($\times$1.2)

Table 7: Computational cost (in seconds) for 10^8 calls pf each method, and ratio with respect to `UniformStep`.

B Convergence of uniform random variables to Gaussian ones

Let $Y_k := \sum_{i=1}^{k} U^{(i)}$ be a random variable defined by the sum of independent uniform random variables $U^{(i)}$ on $[-1, 1]$. After k steps in a zone of constant diffusivity D, the the particles is displaced from x to $x + \sqrt{6D\delta t}Y_k$ using the methods `Uffink` and `HMYLA`. Let $H_{k,N}$ be the empirical DF of N particles moved according to these schemes after k steps, and $H_{k,\infty}$ be the one of Y_k.

On the one hand, the Kolmogorov-Smirnov statistics asserts that $d_{\mathrm{KS}}(H_{k,N}, H_{k,\infty})$ is asymptotically a random variable distributed as $M_{\mathrm{KS}}/\sqrt{N}$, where M_{KS} is the maximum of a Brownian bridge.

On the other hand, some improvements [49] of the Berry-Esséen bounds quantifies the maximal distance $d_{\mathrm{KS}}(H_{k,\infty}, G)$, where G is the DF of the Gaussian random variable with the same variance as Y_k. Basically, this distance is of order $\mathrm{O}(1/k)$.

If $d_{\mathrm{KS}}(H_{k,\infty}, G)$ and $d_{\mathrm{KS}}(H_{k,N}, H_{k,\infty})$ have the same order of magnitude, then the empirical DF of Y_k with N particles cannot be distinguished from the one of N

Gaussian random variables. This means that k steps with `Uffink`/`HMYLA` provides us with results close to one step of `SBM`/`SBMlin` in zone of constant diffusivities.

As the probability that the normalized Kolmogorov-Smirnov distance M_{KS} is higher than 3 is negligible, a rule of thumbs for choosing k_0, the minimal value of k such that $d_{\mathrm{KS}}(H_{k,\infty}, G) \approx 3/\sqrt{N}$. We report in Table 8 the values of k_0 in function in N found according to this rule when the approximation given in [49] is used.

N	10^4	5×10^4	10^5	5×10^5	10^6	5×10^6	10^7
k_0	4	5	6	7	8	10	12

Table 8: Choice of the number of steps k_0 to make the distance between the DF of sum of k_0 uniform random variables and the DF of a Gaussian one of the same order as the empirical DF of N Gaussian random variables and the DF of a Gaussian one given by the Kolmogorov-Smirnov statistics. This time is obtained by solving $\epsilon_{k_0} = 3/\sqrt{N}$ with ϵ_{k_0} given by (12) in [49].

Our numerical tests (not reported here) are in agreement with the theoretical values of k_0 given in Table 8 (see Figure 14(g) and (h) by looking at the number of steps used required by `Uffink` to cross the dotted line at 3). Besides, in presence of discontinuities, the number of steps required so that `Uffink` cannot be distinguished from `SBM` seems to be lower than these theoretical bounds.

Acknowledgments. This work was supported by ANR-MN, with the H2MNO4 project. Computer simulations related to this work were performed on the ADA cluster at the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS), Orsay, France.

Bibliography

[1] P. Ackerer and R. Mose, *Comment on "Diffusion theory for transport in porous media: Transition-probability densities of diffusion processes corresponding to advection-dispersion equations" by Eric M. LaBolle et al.*, Water Resour. Res. **36** (2000), no. 3, 819–821, DOI 10.1029/1999WR900325.

[2] T.A. Appuhamillage, V.A. Bokil, E. Thomann, E. Waymire, and B. D. Wood, *Solute transport across an interface: A Fickian theory for skewness in breakthrough curves*, Water Resour. Res., **46** (2010), W07511, DOI 10.1029/2009WR008258.

[3] T.A. Appuhamillage, V.A. Bokil, E. Thomann, E. Waymire, and B. D. Wood, *Occupation and local times for Skew Brownian motion with application to dispersion across an interface*, Ann. Appl. Probab. **21** (2011), no. 1, 183–214, DOI 10.1214/10-AAP691.

[4] M. Bechtold, J. Vanderborght, O. Ippisch, and H.V. Vereecken, *Efficient random walk particle tracking algorithm for advective-dispersive transport in media with discontinuous dispersion coefficients and water contents*, Water Resour. Res., **47** (2011), W10526, DOI 10.1029/2010WR010267.

[5] A.M. Berezhkovskii, V. Zaloj, and N. Agmon, *Residence time distribution of a Brownian particle*, Physical Review E **57** (1998), no. 4, 3937–3947, DOI 10.1103/PhysRevE.57.3937.

[6] A. Bourgeat, M. Kern, S. Schumacher, and J. Talandier, *The* COUPLEX *Test Cases: Nuclear waste disposal simulation*, Computational Geosciences **8** (2004), 83–98, DOI 10.1023/B:COMG.0000035097.89798.f9.

[7] M. Bossy, N. Champagnat, S. Maire, and D. Talay, *Probabilistic interpretation and random walk on spheres algorithms for the Poisson-Boltzmann equation in Molecular Dynamics*, ESAIM M2AN **44** (2010), no. 5, 997–1048, DOI 10.1051/m2an/2010050.

[8] R.S. Cantrell and C. Cosner, *Diffusion Models for Population Dynamics Incorporating Individual Behavior at Boundaries: Applications to Refuge Design*, Theor. Popul. Biol. **55** (1999), no. 2, 189–207, DOI 10.1006/tpbi.1998.1397.

[9] A. Cortis and A. Zoia, *Model of dispersive transport across sharp interfaces between porous materials*, Phys. Rev. E **80** (2009), 011122, DOI 10.1103/PhysRevE.80.011122.

[10] D. R. Cox and D. V. Hinkley, *Theoretical statistics*, Chapman and Hall, London, 1974.

[11] F. Delay, Ph. Ackerer, and C. Danquigny, *Simulating Solute Transport in Porous or Fractured Formations Using Random Walks Particle Tracking: A Review*, Vadose Zone J. **4** (2005), 360–379, DOI 10.2136/vzj2004.0125.

[12] P. Étoré, *On random walk simulation of one-dimensional diffusion processes with discontinuous coefficients*, Electron. J. Probab. **11** (2006), no. 9, 249–275, DOI 10.1214/EJP.v11-311.

[13] P. Étoré and A. Lejay, *A Donsker theorem to simulate one-dimensional processes with measurable coefficients*, ESAIM Probab. Stat. **11** (2007), 301–326, DOI 10.1051/ps:2007021.

[14] P. Étoré and M. Martinez, *Exact simulation of one-dimensional stochastic differential equations involving the local time at zero of the unknown process*, Monte Carlo Methods Appl. **19** (2013), no. 1, DOI 10.1515/mcma-2013-0002.

[15] W. Feller, *On the Kolmogorov-Smirnov limit theorems for empirical distributions*, Ann. Math. Statistics **19** (1948), 177–189, DOI 10.1214/aoms/1177730243.

[16] E. Fieremans, D. S Novikov, J. H. Jensen, and J. A. Helpern, *Monte Carlo study of a two-compartment exchange model of diffusion*, NMR Biomed. **23** (2010), 711–724, DOI 10.1002/nbm.1577.

[17] M. Fisz, *Probability theory and mathematical statistics*, John Wiley & Sons Inc., New York, 1963.

[18] C. Gardiner, *Stochastic methods. A handbook for the natural and social sciences*, 4th ed., Springer Series in Synergetics, Springer-Verlag, Berlin, 2009.

[19] U. Gräwe, E. Deleersnijder, S. H. A. M. Shah, and A. W. Heemink, *Why the Euler scheme in particle tracking is not enough: the shallow-sea pycnocline test case*, Ocean Dynamics **62** (2012), no. 4, 501–514, DOI 10.1007/s10236-012-0523-y.

[20] U. Gräwe, *Implementation of high-order particle-tracking schemes in a water column model*, Ocean Modelling **36**, no. 1-2, 80–89, DOI 10.1016/j.ocemod.2010.10.002.

[21] H.W. de Haan, M. Chubynsky, and G.W. Slater, *Monte Carlo Approaches for Simulating a Particle at a Diffusivity Interface and the "Ito–Stratonovich Dilemma"* (August 24, 2012), available at `arxiv:1208.5081v1`.

[22] H. Hoteit, R. Mose, A. Younes, F. Lehmann, and Ph. Ackerer, *Three-dimensional modeling of mass transfer in porous media using the mixed hybrid finite elements and the random-walk methods*, Math. Geology **34** (2002), no. 4, 435–456, DOI 10.1023/A:1015083111971.

[23] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992.

[24] E. M. LaBolle and Y. Zhang, *Reply to comment by D.-H. Lim on "Diffusion processes in composite porous media and their numerical integration by random walks: Generalized stochastic differential equations with discontinuous coefficients"*, Water Resour. Res. **42** (2006), W02602, DOI 10.1029/2005WR004403.

[25] E. M. LaBolle, J. Quastel, G. E. Fogg, and J. Gravner, *Diffusion processes in composite porous media and their numerical integration by random walks: Generalized stochastic differential equations with discontinuous coefficients*, Water Resour. Res. **36** (2000), 651–662, DOI 10.1029/1999WR900224.

[26] E. M. LaBolle, G.E. Fogg, and A.F.B. Thomson, *Diffusion theory for transport in porous media: Transition-probability densities of diffusion processes corresponding to advection-dispersion equations*, Water Resour. Res., **34** (1998), no. 7, 1685–1693, DOI 10.1029/98WR00319.

[27] E. M. LaBolle, G.E. Fogg, and A.F.B. Thomson, *Random-Walk Simulation of Transport in Heterogeneous Porous Media: Local Mass-Conservation Problem and Implementation Methods*, Water Resour. Res., **32** (1996), no. 3, 582–593, DOI 10.1029/95WR03528.

[28] A. Lejay and G. Pichot, *Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps*, J. Comput. Phys. **231** (2012), no. 21, 7299–7314, DOI 10.1016/j.jcp.2012.07.011.

[29] A. Lejay, *On the constructions of the Skew Brownian motion*, Probab. Surv. **3** (2006), 413–466, DOI 10.1214/154957807000000013.

[30] A. Lejay and M. Martinez, *A scheme for simulating one-dimensional diffusion processes with discontinuous coefficients*, Ann. Appl. Probab. **16** (2006), no. 1, 107–139, DOI 10.1214/105051605000000656.

[31] A. Lejay, *Simulating a diffusion on a graph. Application to reservoir engineering*, Monte Carlo Methods Appl. **9** (2003), no. 3, 241–256, DOI 10.1515/156939603322729003.

[32] A. Lejay and S. Maire, *Computing the principal eigenvalue of the Laplace operator by a stochastic method*, Math. Comput. Simulation **73** (2007), no. 3, 351–363, DOI 10.1016/j.matcom.2006.06.011.

[33] D.-H. Lim, *Comment on "Diffusion processes in composite porous media and their numerical integration by random walks: Generalized stochastic differential equations with discontinuous coefficients" by E. M. LaBolle, J. Quastel, G. E. Fogg, and J. Gravner*, Water Resour. Res. **42** (2006), W02601, DOI 10.1029/2005WR004091.

[34] A. Marcowith and F. Casse, *Postshock turbulence and diffusive shock acceleration in young supernova remnants*, Astron. Astrophys. **515** (2010), no. A90, DOI 10.1051/0004-6361/200913022.

[35] M. Marseguerra and A. Zoia, *Normal and anomalous transport across an interface: Monte Carlo and analytical approach*, Ann. Nucl. Energy **33** (2006), no. 17–18, 1396–1407, DOI 10.1016/j.anucene.2006.09.012.

[36] M. Martinez, *Interprétations probabilistes d'opérateurs sous forme divergence et analyse de méthodes numériques associées*, PhD thesis, Université de Provence / INRIA Sophia-Antipolis, 2004.

[37] M. Martinez and D. Talay, *Discrétisation d'équations différentielles stochastiques unidimensionnelles à générateur sous forme divergence avec coefficient discontinu*, C. R. Math. Acad. Sci. Paris **342** (2006), no. 1, 51–56, DOI 10.1016/j.crma.2005.10.025.

[38] M. Martinez and D. Talay, *One-Dimensional parabolic diffraction equations: Pointwise estimates and discretization of related stochastic differential equations with weighted local times*, Electron. J. Probab. **17** (2012), no. 27, DOI 10.1214/EJP.v17-1905.

[39] M. Mascagni and N. A. Simonov, *Monte Carlo methods for calculating some physical properties of large molecules*, SIAM J. Sci. Comput. **26** (2004), no. 1, 339–357, DOI 10.1137/S1064827503422221.

[40] G. N. Milstein and M. V. Tretyakov, *Stochastic numerics for mathematical physics*, Scientific Computation, Springer-Verlag, Berlin, 2004.

[41] B. Øksendal, *Stochastic differential equations: An introduction with applications*, 6th ed., Universitext, Springer-Verlag, Berlin, 2003.

[42] O. Ovaskainen and S. J. Cornell, *Biased movement at a boundary and conditional occupancy times for diffusion processes*, J. Appl. Probab. **40** (2003), no. 3, 557–580, DOI 10.1239/jap/1059060888.

[43] J. M. Ramirez, E. A. Thomann, and E. C. Waymire, *Advection–Dispersion Across Interfaces*, Stat. Sci. **28** (2013), no. 4, 487–509, DOI 10.1214/13-STS442.

[44] J. Ramirez, E. Thomann, E. Waymire, J. Chastenet, and B. Wood, *A Note on the Theoretical Foundations of Particle Tracking Methods in Heterogeneous Porous Media*, Water Resour. Res. **44** (2007), W01501, DOI 10.1029/2007WR005914.

[45] J. M. Ramirez, *Skew Brownian Motion and Branching Processes Applied to Diffusion-Advection in Heterogenous Media and Fluid Flow*, PhD thesis, Oregon State University, 2007.

[46] J. M. Ramirez, E. A. Thomann, E. C. Waymire, R. Haggerty, and B. Wood, *A generalized Taylor-Aris formula and skew diffusion*, Multiscale Model. Simul. **5** (2006), no. 3, 786–801, DOI 10.1137/050642770.

[47] P. Salamon, D. Fernàndez-Garcia, and J. J. Gómez-Hernández, *A review and numerical assessment of the random walk particle tracking method*, J. Contaminant Hydrology **87** (2006), no. 3-4, 277–305, DOI 10.1016/j.jconhyd.2006.05.005.

[48] K. Semra, *Modélisation tridimensionnelle du transport d'un traceur en milieux poreux saturé: évaluation des théories stochastiques*, PhD thesis, Université Louis Pasteur, Strasbourg, 1994.

[49] R. Sherman, *Error of the normal approximation to the sum of N random variables*, Biometrika **58** (1971).

[50] G. Shorack and J. Wellner, *Empirical processes with applications to statistics*, John Wiley & Sons, 1986.

[51] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, Ann. Math. Statistics **19** (1948), 279–281.

[52] D. Spivakovskaya, A. W. Heemink, and E. Deleersnijder, *Lagrangian modelling of multi-dimensional advection-diffusion with space-varying diffusivities: theory and idealized test cases*, Ocean Dynamics **57** (2007), no. 3, 189–203, DOI 10.1007/s10236-007-0102-9.

[53] D. Spivakovsakaya, A.W. Heemink, and E. Deleersnijder, *The backward Ito method for the Lagrangian simulation of transport processes with large space variations of the diffusivity*, Ocean Sci. **3** (2007), no. 4, 525–535, DOI 10.5194/os-3-525-2007.

[54] D.J. Thomson, W.L. Physick, and R.H. Maryon, *Treatment of Interfaces in Random Walk Dispersion Models*, J. Appl. Meteorol. **36** (1997), 1284–1295, DOI 10.1175/1520-0450(1997)036<1284:TOIIRW>2.0.CO;2.

[55] G. J. M Uffink, *Analysis of dispersion by the random walk method*, PhD thesis, Delft University, The Netherlands, February 6, 1990.

[56] M. Zhang, *Calculation of diffusive shock acceleration of charged particles by skew Brownian motion*, Astrophys. J. **541** (2000), 428–435.

[57] C. Zheng and G. Bennett, *Applied Contaminant Transport Modeling*, 2nd ed., Wiley-Interscience, 2002.