Uncertain Darcy's problem and the stochastic particle transport

Giacomo Garegnani
Supervisors: Dr. Sebastian Krumscheid and Prof. Fabio Nobile

Contents

1 Introduction

In this project we analyse a model for describing the trajectory of particles of pollutants in groundwater. A successful model for describing underground flows is given by the uncertain Darcy's problem. Given a domain D such that its boundary ∂D is divided in three subsets $\Gamma_{in}, \Gamma_{out}, \Gamma_N$ such that

$$\Gamma_{in} \cup \Gamma_{out} \cup \Gamma_N = \partial D, \quad \Gamma_{in} \cap \Gamma_{out} \cap \Gamma_N = \emptyset,$$

the pressure and velocity fields p and u are given by the solution of the following Partial Differential Equation (PDE)

$$\begin{cases} u = -A\nabla p, & \text{in } D, \\ \nabla \cdot u = f, & \text{in } D, \\ p = p_0, & \text{on } \Gamma_{in}, \\ p = 0, & \text{on } \Gamma_{out}, \\ \nabla p \cdot n = 0, & \text{on } \Gamma_N, \end{cases} \tag{1}$$

where Γ_{in}, Γ_{out} are the inlet and outlet portions of the boundary of D, and an impermeability condition is imposed on Γ_N. The coefficient A is the permeability constant of the material. In practical applications a measure of the composition of the ground in every point is not available, therefore it is necessary to model A using a random field. Problem (1) is then a random PDE, where p, u and A depend on the space variable and on the event ω. Moreover, the smoothness of the random field A is not granted, which makes the problem harder to solve and its solution rough. We are interested in studying the features of the trajectory of a pollutant which is dispersed in the groundwater. For example, an accurate knowledge of the characteristic time it takes for a particle to cover a fixed distance within D could help estimating security zones around extraction wells. The trajectory of pollutant particles in a transport field is intrinsically stochastic and is modeled by a Stochastic Differential Equation (SDE) with the solution of the Darcy's problem as a transport field. Given a starting point X_0 in the domain, the trajectory $X(t)$ is given by the solution of

$$\begin{cases} dX(t) = u(X(t))dt + \sigma dW(t), & 0 \le t \le T, \\ X(0) = X_0 \in D, \end{cases} \tag{2}$$

where $W(t)$ is a standard two-dimensional Wiener process and the scalar value σ represents the stochastic diffusion. Since the transport field u is only defined on the domain D, equation (2) has to be properly equipped with boundary conditions. Both the Darcy problem and the solution of the SDE are stochastic, thus Montecarlo simulations have to be performed to estimate the trajectories of the solution of (2). Therefore, in this project we search an efficient numerical scheme in order to simulate the trajectories of the solution of (2).

This outline of the work is the following.
In Section 2 we investigate the performances of three numerical schemes used to approximate the mean exit time of the solution of a general SDE from a domain. Theoretical results regarding the weak convergence of the schemes are tested and verified.
In Section 3 we present a brief theoretical investigation regarding the convergence of the analytic and numerical solution of an SDE subject to a perturbation in the transport field.
In Section 4 we apply the techniques presented in Section 2 to the uncertain Darcy problem and we present results for the mean exit time in this case.
Finally, in Section 5 we conclude reporting some considerations about possible future developments.

2 Expected exit time from a domain

We aim to estimate the exit time of a particle driven by a deterministic transport field and a stochastic diffusion from a domain $D \subset \mathbb{R}^d$. Given a vector $W(t)$ of m independent Brownian motions and two functions $f\colon \mathbb{R}^d \to \mathbb{R}^d, g\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$, we consider the following stochastic differential equation (SDE)

$$\begin{cases} dX(t) = f(X(t))dt + g(X(t))dW(t), & 0 < t \le T, \\ X(0) = X_0, & X_0 \in D. \end{cases} \tag{3}$$

The problem is equipped with two different types of boundary conditions, namely

i. *killing boundaries*: if the particle exits D passing through this boundary the process is stopped,

ii. *reflecting boundaries*: the particle trajectory is reflected normally inside D when it touches the boundary ∂D.

Our aim is to estimate numerically the first exit time of the solution $X(t)$ from D. Since the solution is computed in a bounded time interval, it could stay in the domain up to the final time T. Hence, we estimate the quantity

$$\tau = \min\{\tau_e, T\}, \text{ where } \tau_e = \min\{t\colon X(t) \notin D\}. \tag{4}$$

Let us remark that estimating the value of τ is meaningful only if there exists a portion of the boundary $\Gamma_k \subset \partial D$ that is endowed with killing boundary conditions. Otherwise, the process $X(t)$ will stay in D for the whole time interval, giving as a result $\tau = T$ for each realisation of $X(t)$. Another quantity of interest is defined as follows

$$\varphi = \varphi(T, X_0, F) = \mathbb{1}_{\{T<\tau_e\}} F(X(T)), \tag{5}$$

where $F\colon \mathbb{R}^d \to \mathbb{R}$. Let us remark that both τ_e and $X(T)$ depend on the initial condition, therefore φ depends on X_0. However, we do not write explicitely the dependence on X_0 in order to simplify the notation. An interesting choice of F could be the function mapping every x of $\mathbb{R}^d$ to one. In this case, the expectation of φ is equal to the probability that $X(t)$ does not exit the domain before the final time T. Therefore, choosing the notation $F = 1$, we can compute the probability of exit before final time as

$$\Phi(T, X_0) \coloneqq \Pr(\tau < T | X(0) = X_0) = 1 - \mathbb{E}(\varphi(T, X_0, 1)). \tag{6}$$

In case of general f, g, F and for a d-dimensional SDE, no closed form for τ and φ is available. Therefore, we approximate the value of τ and φ by means of three numerical schemes, briefly presented in the following.

2.1 Numerical Methods

2.1.1 Discrete Euler-Maruyama

Given $N \in \mathbb{N}$ let us define a partition of $[0, T]$ as $P_h = \{t_i\}_{i=0}^N, t_i = ih, h = T/N$. The Discrete Euler-Maruyama method (DEM) for problem (3) is defined as follows

$$\begin{cases} X^d_{h,i+1} = f(X^d_{h,i})h + g(X^d_{h,i})(W(t_{i+1}) - W(t_i)), \\ X^d_{h,0} = X_0. \end{cases} \tag{7}$$

The exit time τ is approximated with the quantity τ^d_h defined as

$$\tau^d_h = \min\{\tau^d_{h,e}, T\}, \text{ where } \tau^d_{h,e} = h \min\{i\colon X^d_{h,i} \notin D\}. \tag{8}$$

We approximate analogously φ as

$$\varphi^d_h = \mathbb{1}_{\{T<\tau^d_{h,e}\}} F(X^d_{h,N}). \tag{9}$$

It has been shown [2, 4, 7] that under appropriate assumptions on the functions f and g, as well as the domain D and its boundary, the weak error of the method for approximating the values of interest is of order 1/2, *i.e.*,

$$|\mathbb{E}(\tau^d_h) - \mathbb{E}(\tau)| = O(\sqrt{h}), \tag{10}$$

$$|\mathbb{E}(\varphi^d_h) - \mathbb{E}(\varphi)| = O(\sqrt{h}). \tag{11}$$

It is known that the weak error for Euler-Maruyama is of order 1, therefore there is a loss of 1/2 in case of killing and reflecting boundaries [8, Chapter 14]. This is due to the missed exits, *i.e.*, a part of the solution which lays outside the domain connecting two discrete values that are inside the domain. An example of a missed exit is depicted in Figure 1, where the solution computed with a large step size h stays in the domain, while with a smaller step size one can conclude that the solution has actually exited the domain. The methods that we present in the following implement two different strategies to correct this behaviour.

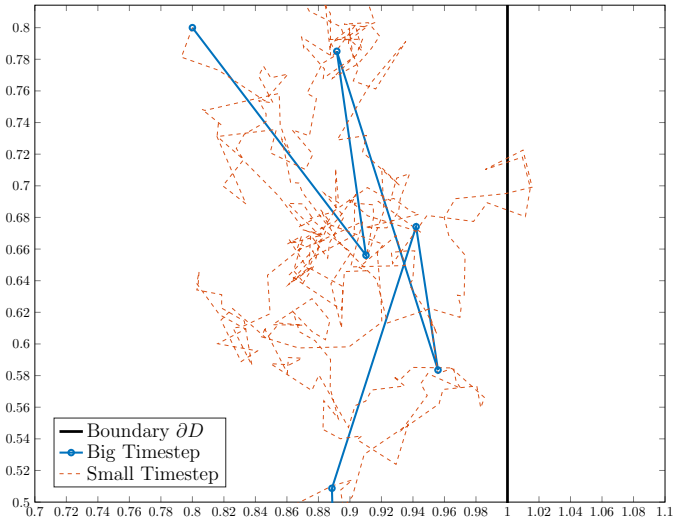


Figure 1: An example of missed exit. The trajectory computed with the small step size exits the domain, while for the big step size the exit is not detected.

2.1.2 Continuous Euler-Maruyama.

Let us consider the partition P_h of $[0,T]$ as above. The Continuous Euler-Maruyama (CEM) method is defined as

$$\begin{cases} X_h^c(t) = f(X_h^c(t_i))(t-t_i) + g(X_h^c(t_i))(W(t)-W(t_i)), \quad t_i < t \leq t_{i+1}, \\ X_h^c(0) = X_0. \end{cases} \tag{12}$$

Let us remark that in case the particle does not exit the domain, $X_h^c(t_i) = X_{h,i}^d$ for all $t_i \in P_h$. It is possible to compute the probability that a particle has exited the domain at a time t between two consecutive timesteps t_i, t_{i+1} when D is an half-space with the following formula [3]

$$\Pr(\exists t \in [t_i, t_{i+1}] \quad X_h^c(t) \notin D | X_h^c(t_i) = x_i, X_h^c(t_{i+1}) = x_{i+1}) = p(x_i, x_{i+1}, h), \tag{13}$$

with $p(x_i, x_{i+1}, h)$ given by

$$p(x_i, x_{i+1}, h) = \exp\Big(-2\frac{[n\cdot(x_i - z_i)][n\cdot(x_{i+1}-z_i)]}{hn\cdot(gg^T(x_i)n)}\Big), \tag{14}$$

where z_i is the projection of x_i on ∂D and n is the vector normal to ∂D in z_i pointing outwards D. At each timestep t_{i+1} we compute the probability $p(x_i, x_{i+1}, h)$, and then simulate a variable U distributed uniformly in the interval $[0,1]$, thus obtaining a realization $\tilde{U}_i$. Hence, we conclude that the particle has left the domain for a time t in (t_i, t_{i+1}) if $\tilde{U}_i$ is smaller than $p(x_i, x_{i+1}, h)$. This procedure is equivalent to simulating a Bernoulli variable of parameter $p(x_i, x_{i+1}, h)$ at each timestep. Therefore, we approximate the exit time as

$$\begin{aligned} \tau_h^c &= \min\{T, \tau_{h,e}^c\}, \\ \text{where } \tau_{h,e}^c &= \min\{h\min\{i\colon X_h^c(t_i) \notin D\}, h\min\{i\colon \tilde{U}_i < p(X_h^c(t_{i-1}), X_h^c(t_i), h)\}\}. \end{aligned} \tag{15}$$

In the same way as in DEM, we can approximate φ as

$$\varphi_h^c = \mathbb{1}_{\{T<\tau_{h,e}^c\}} F(X_h^c(T)). \tag{16}$$

We show the pseudo-code for the implementation of CEM in Algorithm 1. It has been shown [4] that under appropriate assumptions on the functions f and g, as well as the domain D and its boundary, the weak error of this method when approximating the values of interest is 1, *i.e.*,

$$|\mathbb{E}(\tau_h^c) - \mathbb{E}(\tau)| = O(h), \tag{17}$$

$$|\mathbb{E}(\varphi_h^c) - \mathbb{E}(\varphi)| = O(h). \tag{18}$$

Algorithm 1: Continuous Euler-Maruyama

for $t_i \in P_h$ **do**
  $X(t_{i+1}) = f(X(t_i))h + g(X(t_i))(W(t_{i+1}) - W(t_i))$;
  if $X(t_{i+1}) \notin D$ **then**
    $\tau_h^c = t_{i+1}$;
    $\varphi_h^c = 0$;
    return
  else
    compute $p = p(x_i, x_{i+1}, h)$;
    draw $\tilde{U}_i$ from $U \sim \text{Unif}(0,1)$;
    if $\tilde{U}_i < p$ **then**
      $\tau_h^c = t_{i+1}$;
      $\varphi_h^c = 0$;
      return
    end
  end
  $\tau_h^c = T$;
  $\varphi_h^c = F(X_h^c(T))$;
end

This result implies that computing the probability of exit at each timestep it is possible to avoid the phenomenon of the missed exits, restoring the weak order 1 of the standard Euler-Maruyama method. Let us remark that the complexity of CEM for computing a single time step is higher than for DEM, as the probability has to be computed for each part of the boundary of D and compared with $\tilde{U}_i$. On the other side, it is likely that CEM terminates before DEM, as the exit could be detected at an earlier stage.

2.1.3 A space and time adaptive procedure

Let us consider the function g in (3) to be fixed to a real constant σ. The numerical methods we presented can be completed with an adaptive procedure for the step size h [1]. In this frame, adaptivity is performed with respect to the position of the numerical solution $X_h(t_i)$, *i.e.*, the step size is reduced if the solution is near to the boundary of the domain D. If we denote by d the distance between $X_h(t_i)$ and its normal projection on ∂D, by $l \in \mathbb{N}$ a fixed parameter and by h_0 a maximum bound for h, the step size is chosen using the following formula [1]

$$h = \max\left\{2^{-2l}h_0, \min\left\{2^{-l}h_0, \left(\frac{d}{(l+3)\sigma}\right)^2\right\}\right\}. \tag{19}$$

This formula is equivalent to dividing the domain D in three zones, in particular

- an interior zone where $h = h_{int} = 2^{-l}h_0$,
- a boundary zone where $h = h_{bound} = 2^{-2l}h_0$,
- an intermediate zone where $h = \left(\frac{d}{(l+3)\sigma}\right)^2$.

Since the loss of 0.5 in the weak order of convergence of Euler-Maruyama is given by the missed exits, the convergence issue of DEM is due to the values of the solution near the boundary. Therefore, refining the step size only near to the boundary could lead to an error equal to the one obtained using a fine step size on the whole domain.

2.1.4 Reflecting boundaries

The reflecting boundaries are treated in the same way for both DEM and CEM. Let us denote by Γ_k and Γ_r the killing and reflecting subsets of ∂D, *i.e.*,

$$\Gamma_r \cup \Gamma_k = \partial D, \quad \Gamma_r \cap \Gamma_k = \emptyset \tag{20}$$

If for a timestep of $t_i \in P_h$, X_i is not in D and the line connecting X_{i-1} and X_i crosses Γ_r, we update the solution to be the normal reflection inside D of X_i. Let us denote by $\tilde{X}_i$ the new guess for the solution at time $t = t_i$. One has to consider the case in which $\tilde{X}_i$ is outside D as well. In this case, if the line connecting X_{i-1} and X_i crosses Γ_k, the algorithm terminates. On the other hand, if it has crossed another portion of Γ_r, another reflection has to be performed, and the procedure explained above is repeated. This procedure does not spoil the precision of the numerical solution [6].

2.2 A PDE approach

It is possible to express the mean exit time and the probability of exit from a domain in terms of the solution of partial differential equations (PDE's). Let us denote by Γ_k, Γ_r the killing and reflecting subsets of ∂D. We consider then the expectation of the exit time from the domain D for a trajectory that at $t = 0$ is at position x, *i.e.*,

$$\bar{\tau}(x) = \mathbb{E}(\tau | X(0) = x). \tag{21}$$

Let us define the operator $\mathcal{L}$ induced by (3), which is applied to a function $u \colon \mathbb{R}^d \to \mathbb{R}$ as follows

$$\mathcal{L}u = f \cdot \nabla u + \frac{1}{2} g g^T : \nabla \nabla u, \tag{22}$$

where the : operator between two matrices A, B in $\mathbb{R}^{d \times d}$ is defined as follows

$$A : B = \sum_{i,j=1}^{d} \{A\}_{ij} \{B\}_{ij} = \mathrm{tr}(A^T B). \tag{23}$$

The following result allows computing the mean exit time as the solution of an appropriate PDE.

Proposition 2.1. *Let $\mathcal{L}$ be the differential operator defined as (22). Then, if Γ_k and Γ_r are respectively the killing and reflecting subsets of ∂D, such that $\Gamma_k \cup \Gamma_r = \partial D, \Gamma_k \cap \Gamma_r = \emptyset$, the mean exit time $\bar{\tau}(x)$ for the solution $X(t)$ of (3) with $X_0 = x$ is the solution of the following boundary value problem*

$$\begin{cases} \mathcal{L}\bar{\tau}(x) = -1, & \text{in } D, \\ \bar{\tau}(x) = 0, & \text{on } \Gamma_k, \\ \nabla \bar{\tau}(x) \cdot n = 0, & \text{on } \Gamma_r, \end{cases} \tag{24}$$

where n is the normal to Γ_r.

Further analytic treatment of the mean exit time can be found in [9, 11].
We now consider the probability of exit from D for a solution $X(t)$ that is equal to x for at a time s smaller than the final time T. This probability is the solution of a boundary value problem.

Proposition 2.2. *Let $\mathcal{L}$ be the differential operator defined as (22). Then, if Γ_k and Γ_r are respectively the killing and reflecting subsets of ∂D, such that $\Gamma_k \cup \Gamma_r = \partial D, \Gamma_k \cap \Gamma_r = \emptyset$*

$$\Pr(\tau < T | X(s) = x) = \Phi(x, s, T) \tag{25}$$

where $\Phi(x, t, T)$ is the solution of the following backwards PDE

$$\begin{cases} \dfrac{\partial}{\partial t}\Phi(x,t,T) + \mathcal{L}\Phi(x,t,T) = 0 & \text{in } D, s \le t < T, \\ \Phi(x,t,T) = 1 & \text{on } \Gamma_k, s \le t \le T, \\ \nabla\Phi(x,t,T) \cdot n = 0, & \text{on } \Gamma_r, s \le t \le T, \\ \Phi(x,T,T) = 0 & \text{in } D, \end{cases} \tag{26}$$

where n is the normal to Γ_r.

The proof in case $\Gamma_k = \partial D$ of this result can be found in [14]. Further treatment in case of mixed boundary conditions and the closed form of the solution for some particular geometries of $D \subset \mathbb{R}^2$ can be found in [5]. It is therefore possible to approximate the mean exit time and the exit probability by means of classical methods for solving PDE's numerically, such as finite differences or the Finite Elements Method. In Appendix A we present an analytic formula to compute the mean exit time in the one-dimensional case, and in Appendix B we show a finite difference approach to approximate numerically the solution of (26) in the one-dimensional case. In the two-dimensional case, we solve (24) and (26) using linear Finite Elements using the PDE toolbox of `Matlab`.

2.3 Numerical Experiments

In the following we present a series of experiments on test cases in order to verify the properties of the methods presented above. All experiments are performed using Montecarlo simulations to approximate the mean exit time and the exit probability. For each realization (or trajectory) of the Brownian motion we compute τ and φ using DEM or CEM. We denote by M the number of simulated trajectories and approximate the quantities of interest as

$$\begin{aligned}
\mathbb{E}(\tau_h^d) &\approx \frac{1}{M}\sum_{j=1}^{M}(\tau_h^d)_j, \\
\mathbb{E}(\tau_h^c) &\approx \frac{1}{M}\sum_{j=1}^{M}(\tau_h^c)_j, \\
\mathbb{E}(\varphi_h^d) &\approx \frac{1}{M}\sum_{j=1}^{M}(\varphi_h^d)_j, \\
\mathbb{E}(\varphi_h^c) &\approx \frac{1}{M}\sum_{j=1}^{M}(\varphi_h^c)_j,
\end{aligned} \tag{27}$$

where $(\tau_h^d)_j, (\varphi_h^d)_j$ denote the values of τ and φ given by the j-th trajectory of DEM (respectively CEM).

2.3.1 One-Dimensional Case

We consider problem (3) in case $d = 1$. Given f, g defined on an interval $D = [l, r]$ and a Brownian motion $W(t)$, let us consider the following one dimensional SDE

$$\begin{cases} dX(t) = f(X(t))dt + g(X(t))dW(t), & 0 < t \leq T, \\ X(0) = X_0, & X_0 \in D. \end{cases} \tag{28}$$

In this case, the boundary of D consists of the two points $\{l, r\}$. In order for the problem of the determination of τ to be meaningful, at least one of the two points should be endowed with a killing boundary condition.

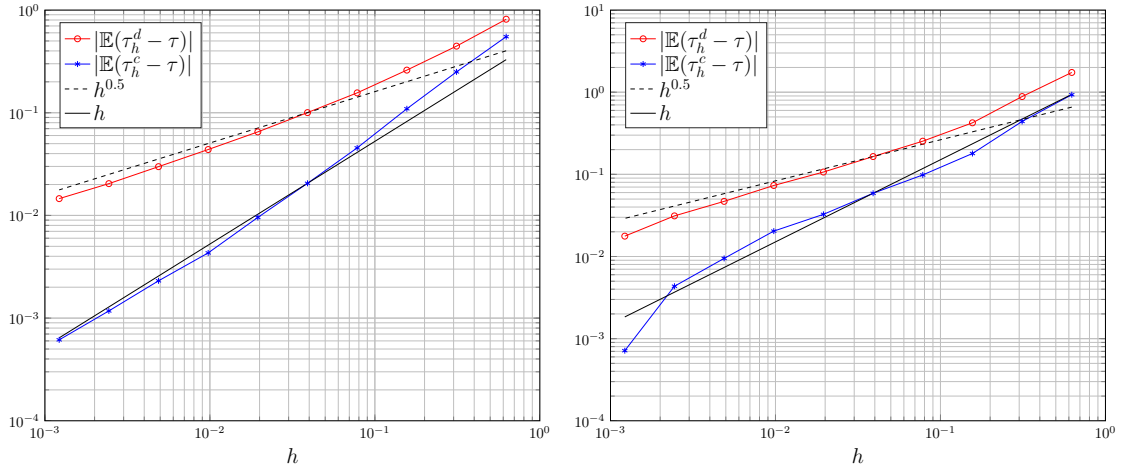
Estimation of the exit time. We consider as a domain for (28) the interval $D = [-1, 1]$, final time $T = 5$ and the following functions

$$\begin{aligned}
&f(x) = -V'(x), \text{ where } V(x) = 0.1(8x^4 - 8x^2 + x + 2), \\
&g(x) = \sigma = 3.
\end{aligned} \tag{29}$$

We approximate the value of τ with a Montecarlo simulation of τ_h^d and τ_h^c computed as in (8) and (15) from the solutions provided by DEM and CEM respectively. In order to verify the order of convergence of the methods, we let N vary in the set $2^i, i = 3, \ldots, 12$ and we fix the number of trajectories M to 10^4. In this way, the error caused by the Montecarlo estimation should not spoil the order of convergence. We compute the error with respect to the analytic value of the mean exit time in this simple case (see Appendix A). In Figure 2 we show the errors obtained fixing $X_0 = 0$ in both the cases of killing and reflecting boundary conditions in $x = 1$. Moreover, in Figure 3 we show an approximation of τ obtained with the two methods with $h = T/128$ and $M = 10^3$ for a set of 10 initial values equispaced along D. It is possible to remark that computing the probability of exit between two consecutive timesteps as in (14) allows correcting the overestimation of τ obtained simply using DEM. As far as the performances are concerned, we remark that the computational time required by CEM is higher than for DEM if the same value of h is employed. On the other hand, fixing the error, CEM is faster than DEM in this case.

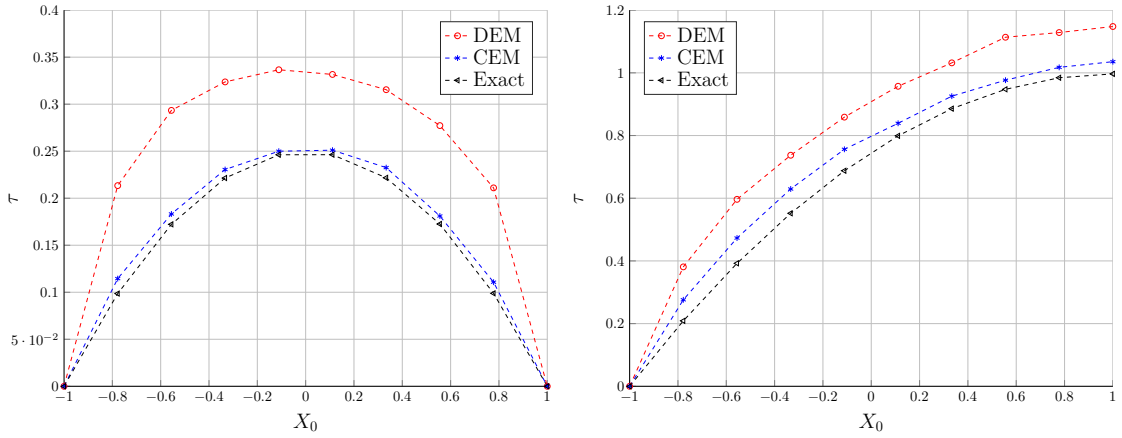
Estimation of the exit probability. We consider (28) with $D = [-1, 1]$, the final time $T = 1$, initial condition $X_0 = 0$ and we define

$$\begin{aligned}
&f(x) = -V'(x), \text{ where } V(x) = 0.1(8x^4 - 8x^2 + x + 2), \\
&g(x) = \sigma = 2.
\end{aligned} \tag{30}$$



(a) Killing boundary in $x = 1$ (b) Reflecting boundary in $x = 1$

Figure 2: Approximation of τ. Orders of convergence for DEM and CEM in the one-dimensional case.



(a) Killing boundary in $x = 1$ (b) Reflecting boundary in $x = 1$

Figure 3: Approximation of τ as a function of the initial value X_0.

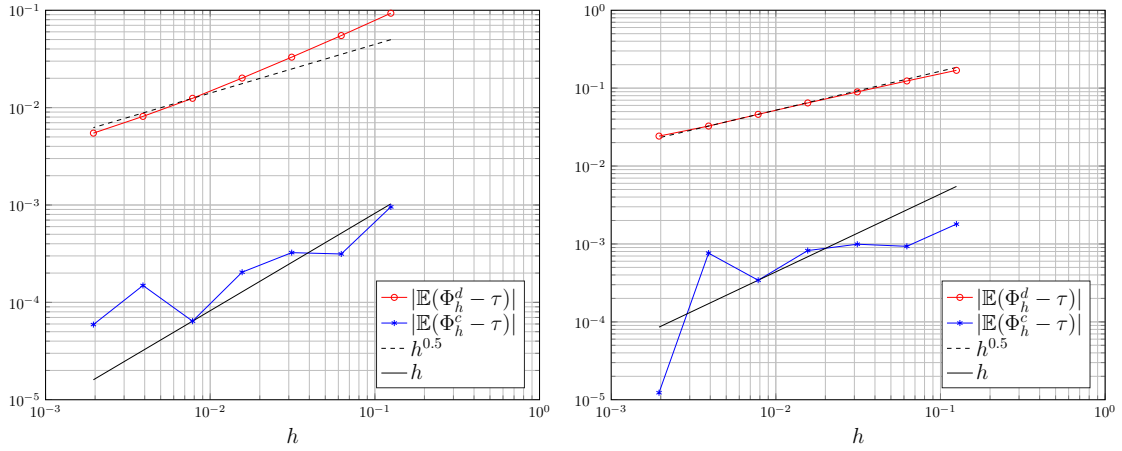
In order to approximate Φ, we perform a Montecarlo simulation using both DEM and CEM, with $M = 10^5$ trajectories. We consider the number of timesteps for the time integration to be $N = 2^i, i = 3, \dots, 9$. The reference solution used to compute errors is obtained with Finite Differences applied to problem (26) (see Appendix B). Numerical results (Figure 4a) confirm that the weak error for DEM is of order 0.5. For CEM it is not trivial to show a clear order of convergence one. This is due to the fact that the results are accurate already with a big step size, therefore the statistical error is not negligible with respect to the error due to time integration. In order to avoid the noise on the order of convergence, it would be necessary to increase the number of trajectories dramatically, leading to unaffordable computational times. Let us remark that CEM approximates the exit probability better than DEM for any initial value X_0, as shown in Figure 5.

2.3.2 Two-dimensional case

We are interested in estimating the exit time of a particle from a domain $D \subset \mathbb{R}^2$. Given $W(t)$ a vector of two independent Brownian motions, we consider the equation (3). In this case, $f \colon \mathbb{R}^2 \to \mathbb{R}^2, g \colon \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$. We compute the mean exit time and the exit probability using DEM and CEM and compare results with the numerical solution of the PDE's presented in 2.2.

Estimation of the exit time. We consider a simple case of (3) in $D = [-1, 1]^2$, where

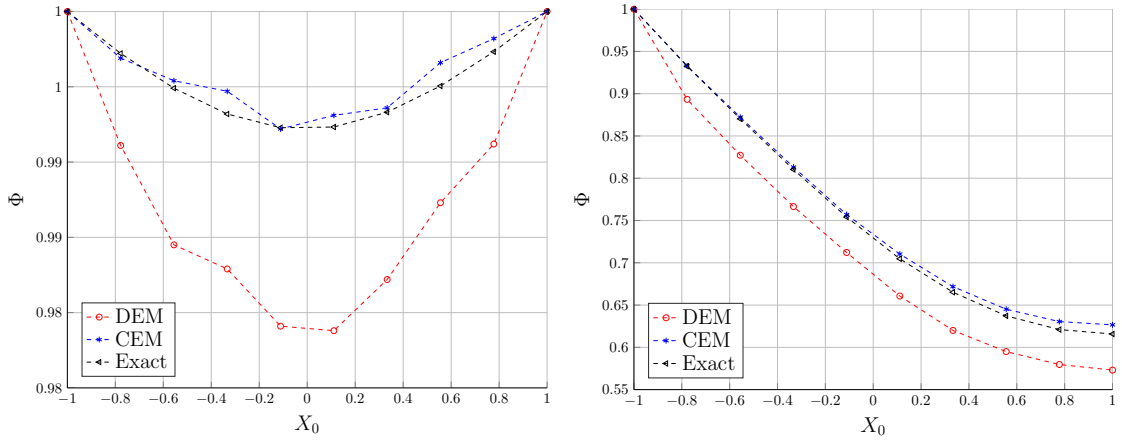
$$f = 0 \in \mathbb{R}^2, \; g = \sigma I \in \mathbb{R}^{2\times 2}, \sigma \in \mathbb{R}.$$



(a) Killing boundary in $x = 1$

(b) Reflecting boundary in $x = 1$

Figure 4: Approximation of Φ. Orders of convergence of DEM and CEM in the one-dimensional case.



(a) Killing boundary in $x = 1$

(b) Reflecting boundary in $x = 1$

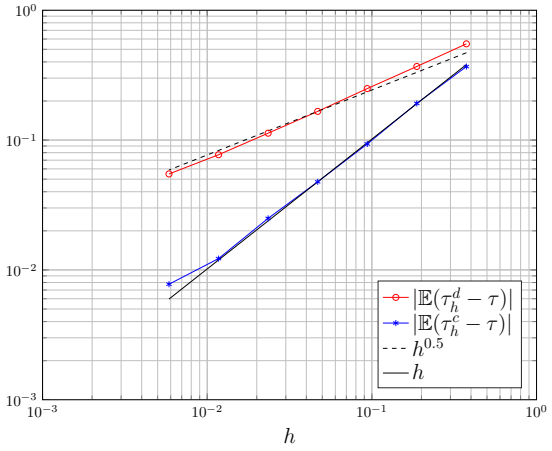
Figure 5: Approximation of Φ as a function of the initial value X_0.

Moreover, we consider ∂D to be a killing boundary. The solution in this case is a Brownian motion. In this case, the partial differential equation (24) reduces to

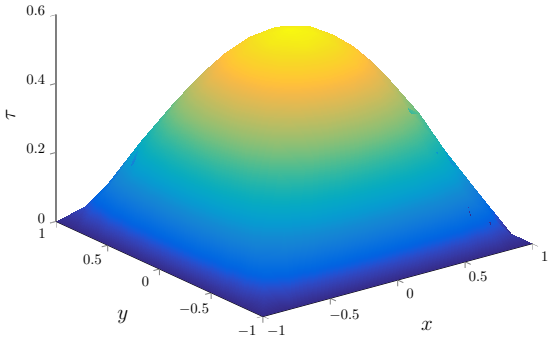
$$\begin{cases} -\sigma^2 \Delta \bar{\tau} = 2, & \text{in } D, \\ \bar{\tau} = 0, & \text{on } \partial D. \end{cases} \tag{31}$$

This is the Poisson equation, hence it is possible to solve it numerically with the Finite Elements Method or the finite differences avoiding a high computational cost. We use the Finite Elements Method adopting a regular mesh with equal constant spacing in the x and y directions, obtaining a solution as in Figure 6b. In order to verify the orders of convergence of DEM and CEM, we set $T = 3$, $\sigma = 1$, $X_0 = (0, 0)^T$, with $M = 10^4$ and $N = 2^i, i = 3, \ldots, 9$. We then compare the Montecarlo estimation with the value of $\bar{\tau}$ in $(0, 0)$, obtained by an evaluation of the Finite Elements solution. The orders of convergence for this numerical experiment are shown in Figure 6a. The results confirm the theoretical orders of convergence for DEM and CEM, with an average order of 0.55 for DEM and 0.93 for CEM, which corrects to 0.98 if the last point is not taken into account.

We consider the same problem as above with mixed killing and reflecting boundary conditions. The functions f and g are the same as above, so the SDE model does not change, but we consider the two left and right boundaries of D, defined by $x = \pm 1$, to be reflecting. We denote this portion

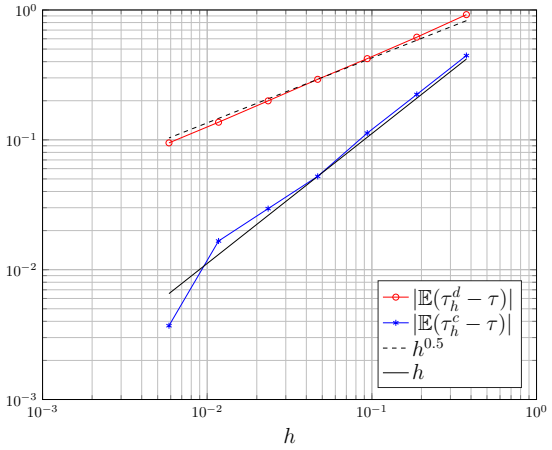


(a) Convergence of CEM and DEM.

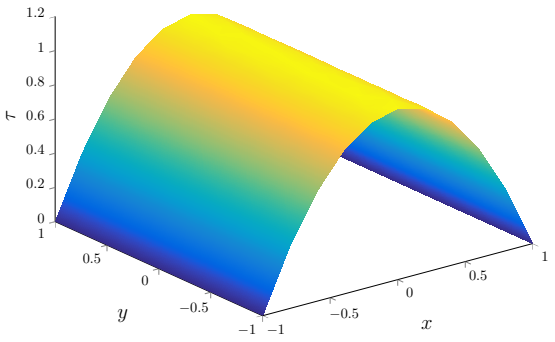


(b) Expectation of exit time.

Figure 6: Summary of the results for τ in the two-dimensional case with pure killing boundary conditions.



(a) Convergence of CEM and DEM.



(b) Expectation of exit time.

Figure 7: Summary of the results for τ in the two-dimensional case with mixed boundary conditions.

of the boundary as Γ_r, and the rest as Γ_k. In this case, the equation for $\bar{\tau}$ becomes

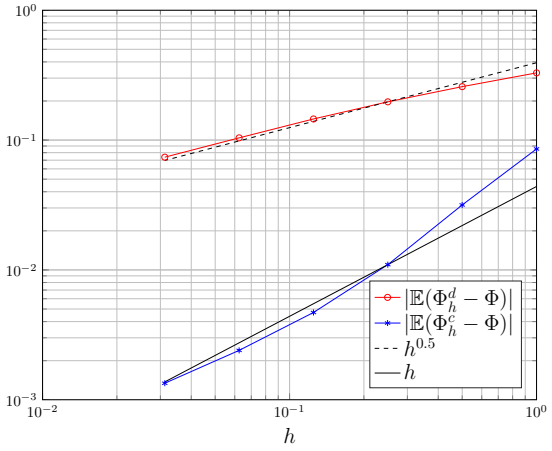
$$\begin{cases} -\sigma^2 \Delta \bar{\tau} = 2, & \text{in } D, \\ \bar{\tau} = 0, & \text{on } \Gamma_k, \\ \partial \bar{\tau} \cdot n = 0, & \text{on } \Gamma_r. \end{cases} \tag{32}$$

The solution of this equation obtained with the Finite Elements Method is shown in Figure 7b. We compute the expectation of τ with DEM and CEM with the same parameters as above. Results (Figure 7a), show that the theoretical orders of convergence are not spoiled by this choice of boundary conditions. The mean order for DEM in this case is 0.55, while for CEM it is 1.15.

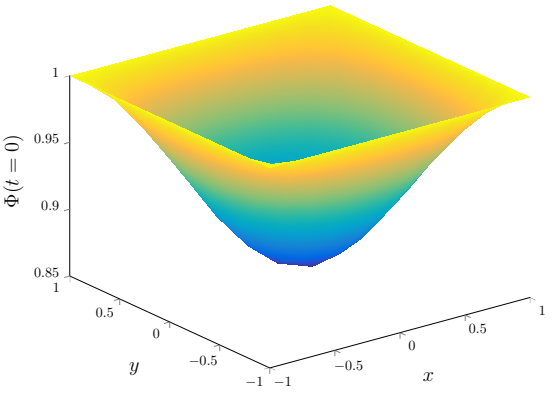
Estimation of the exit probability. We consider the same simple case as for the estimation of the exit time, with ∂D endowed with pure killing boundary conditions. In this case, the partial differential equation (26) reduces to

$$\begin{cases} \dfrac{\partial}{\partial t}\Phi(x,t,T) + \dfrac{1}{2}\sigma^2 \Delta \Phi(x,t,T) = 0, & \text{in } D, 0 \le t < T, \\ \Phi(x,t,T) = 1, & \text{on } \partial D, 0 \le t < T, \\ \Phi(x,T,T) = 0, & \text{in } D. \end{cases} \tag{33}$$

We solve this problem numerically with the Finite Elements Method as for (32). The solution at $t = 0$ is shown in Figure 8b. We verify the orders of convergence of DEM and CEM setting



(a) Convergence of CEM and DEM.



(b) Probability of exit.

Figure 8: Summary of the results for Φ in the two-dimensional case with pure killing boundary conditions.

$X_0 = (0,0)^T, \sigma = 1, T = 1$. We consider $M = 10^6$ trajectories and $N = 2^i, i = 0, \dots, 5$. We then compare the Montecarlo estimation with the value of Φ in $(0,0)$, obtained by interpolation on the Finite Elements solution. The orders of convergence for this numerical experiment are shown in Figure 8a. The theoretical orders of convergence are confirmed in this case as well, with an average order of 0.43 for DEM and 1.19 for CEM.

We consider now mixed boundary conditions. We consider the same values for the parameters, the time integration and the Montecarlo estimation as in the pure killing case. In this case, we set the boundary conditions to be reflecting on the subset of the boundary of D defined by $x = \pm 1$ and killing for the other boundaries.Therefore, in this case the exit probability Φ is the solution of the following PDE

$$\begin{cases} \dfrac{\partial}{\partial t}\Phi(x,t,T) + \dfrac{1}{2}\sigma^2 \Delta\Phi(x,t,T) = 0, & \text{in } D, 0 \le t < T, \\ \Phi(x,t,T) = 1, & \text{on } \Gamma_k, 0 \le t < T, \\ \nabla\Phi(x,t,T)\cdot n = 0, & \text{on } \Gamma_r, 0 \le t < T, \\ \Phi(x,T,T) = 0, & \text{in } D. \end{cases} \tag{34}$$

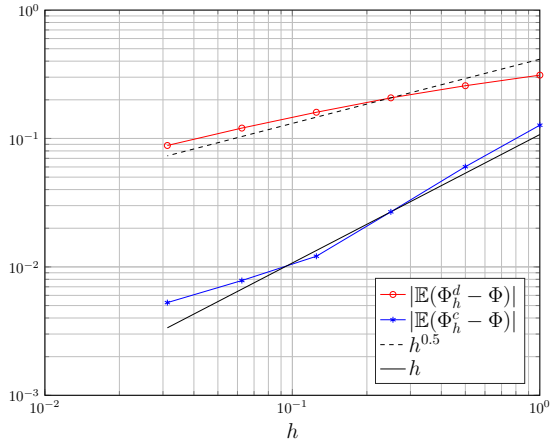
The solution of this equation computed with Finite Elements is shown in Figure 9b. The convergence results for DEM and CEM are shown in Figure 9a. The mean orders in this case are 0.37 for DEM and 0.87 for CEM, which is less than the prediction given by theoretical results. This decrease in the convergence rate is remarkable for small values of h, and it could then be due to the statistical error.

2.3.3 Adaptivity

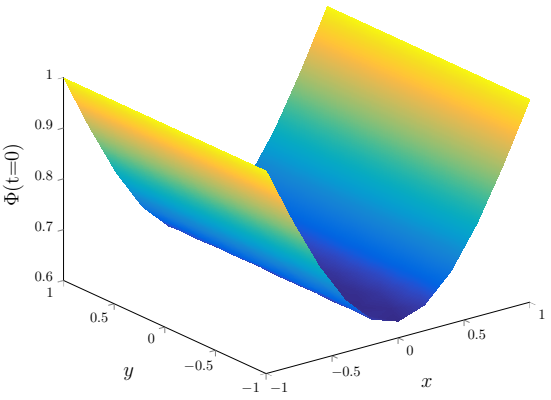
We apply to a test case the adaptivity procedure explained in section 2.1.3. In particular, we fix $f = 0, g = \sigma = 1, T = 3$ and $X_0 = (0,0)^T$ in the square domain $D = [-1,1]^2$ with pure killing boundary conditions. We vary l in (19) in the range $l = 0, \dots, 7$ and we consider three methods

1. Adaptive method with $h_0 = T, h_{int} = 2^{-l}h_0, h_{bound} = 2^{-2l}h_0$,
2. DEM with $h = h_{bound}$ for each timestep,
3. CEM with $h = h_{int}$ for each timestep,

Since $h_{bound} \sim h_{int}^2$, we expect the three methods to have an error of order $O(h_{int})$ when estimating the mean exit time. In Figure 10a it is possible to remark that the error is in fact of order $O(h_{int})$ for the three methods. Moreover, the methods have approximately the same error, *i.e.*, the constant multiplying h_{int} is the same for the three methods. In order to choose which method performs better from the point of view of computational cost, we compute the mean number of timesteps

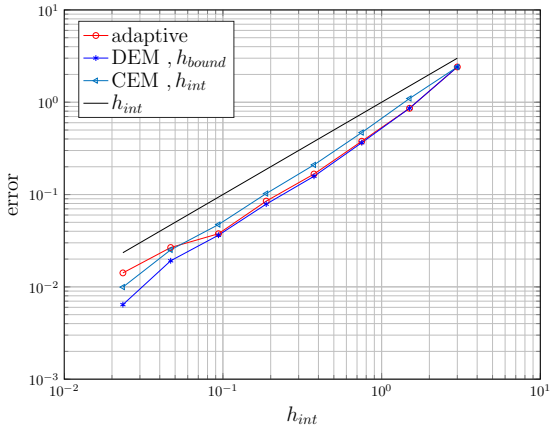


(a) Convergence of CEM and DEM.

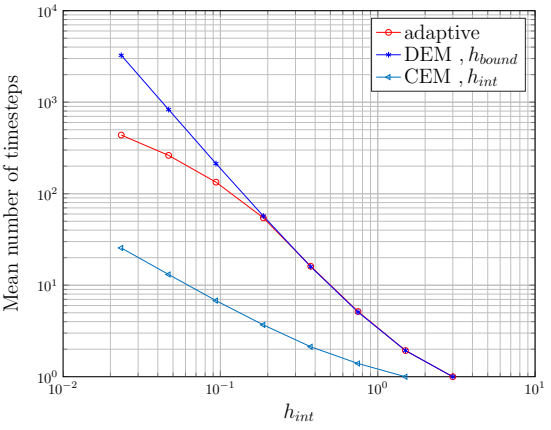


(b) Probability of exit.

Figure 9: Summary of the results for Φ in the two-dimensional case with mixed boundary conditions.



(a) Convergence of the methods.



(b) Computational cost.

Figure 10: Estimation of τ in the square domain with mixed boundary conditions. Comparison between DEM with fixed or adaptive step size.

that they perform in order to estimate the exit time. From Figure 10b it is clear that CEM with h_{int} for every timestep performs better than the other two methods, with the adaptive procedure applied to DEM which implies lower computational time than DEM with constant timestep equal to h_{bound}. We can conclude that the approach based on Brownian bridge proposed in [3] has better performances when estimating the mean exit time from a domain.

3 Theoretical investigation

In this section we analyse the impact on the analytic and numerical solution of an SDE given by a perturbation on the transport field. The theoretical investigation we present is needed in order to give a meaning to the results we find in the Darcy case. Since the velocity field is approximated by means of the Finite Elements method, we hope that the solution of an SDE that has the numerical approximation of the velocity field as transport term would converge to the solution of the SDE that uses the exact velocity field. Moreover, it can be costly to evaluate the Finite Element solution at each timestep of DEM or CEM, but an interpolation procedure can be exploited in order to obtain faster simulations. Hence, another source of perturbation is introduced, and a theoretical background is fundamental to strengthen the basis of our method.

3.1 Analysis of perturbed SDEs

Let us consider $(\Omega, \mathcal{A}, P)$ a complete probability space, $(W_1(t), t \geq 0), (W_2(t), t \geq 0)$ two d-dimensional standard Wiener processes not necessarily independent and a filtration $(\mathcal{F}(t), t \geq 0)$ such that $W_1(t), W_2(t)$ are $\mathcal{F}(t)$-measurable. Moreover, let us consider $\sigma \in \mathbb{R}$, and a function $f\colon \mathbb{R}^d \to \mathbb{R}^d$ and the following SDE

$$\begin{cases} dX(t) = f(X(t))dt + \sigma I dW_1(t), \quad 0 < t \leq T, \\ X(0) = X_0, \end{cases} \tag{35}$$

where I is the identity matrix in $\mathbb{R}^{d\times d}$. Let us consider a perturbation of the transport field $f^\varepsilon\colon \mathbb{R}^d \to \mathbb{R}^d$. Then, we consider the perturbed SDE

$$\begin{cases} dX^\varepsilon(t) = f^\varepsilon(X^\varepsilon(t))dt + \sigma I dW_2(t), \quad 0 < t \leq T, \\ X^\varepsilon(0) = X_0. \end{cases} \tag{36}$$

Finally, let us introduce the following notation. Given a function $F\colon \mathbb{R}^d \to \mathbb{R}$

$$\|F\|_\infty = \sup_{x\in\mathbb{R}^d} |F(x)|.$$

We can state the following preliminary result

Lemma 3.1. *With the notation above, let us consider $d = 1$, $W_1 = W_2 = W$ almost everywhere. If the following assumptions are verified for a constant $K > 0$*

1. $|f(x) - f(y)| \leq K|x - y|$, $\forall x, y \in \mathbb{R}$,
2. $|f(x)| \leq K(1 + |x|)$, $\forall x \in \mathbb{R}$,

and if the solution $X^\varepsilon(t)$ of (36) exists, then $X^\varepsilon(t)$ and $X(t)$ the solution of (35) satisfy

$$\mathbb{E} \sup_{0\leq t\leq T} |X^\varepsilon(t) - X(t)|^2 \leq 2T^2\|f - f^\varepsilon\|_\infty^2 e^{2K^2T^2}.$$

Proof. For almost all $\omega \in \Omega$

$$\begin{aligned} |X^\varepsilon(t) - X(t)|^2 &= \Big| \int_0^t (f^\varepsilon(X^\varepsilon(s)) - f(X(s)))ds\Big|^2 \\ &\leq T \int_0^t |f^\varepsilon(X^\varepsilon(s)) - f(X(s))|^2 ds \\ &\leq 2T \int_0^t |f^\varepsilon(X^\varepsilon(s)) - f(X^\varepsilon(s))|^2 ds + 2T \int_0^t |f(X^\varepsilon(s)) - f(X(s))|^2 ds \\ &\leq 2T^2\|f - f^\varepsilon\|_\infty^2 + 2T^2K^2 \int_0^t |X^\varepsilon(s) - X(s)|^2 ds, \end{aligned}$$

where we applied Cauchy-Schwarz inequality, Young inequality, assumption 1 and the definition of $\|\cdot\|$ respectively. We then apply Gronwall's inequality, which gives

$$|X^\varepsilon(t) - X(t)|^2 \leq 2T^2\|f - f^\varepsilon\|_\infty^2 e^{2K^2T^2}, \quad a.e. \tag{37}$$

Since the right hand side of the inequality is independent of time and ω, we can then take the supremum over time and the expectation at both sides, *i.e.*,

$$\mathbb{E} \sup_{0\leq t\leq T} |X^\varepsilon(t) - X(t)|^2 \leq 2T^2\|f - f^\varepsilon\|_\infty^2 e^{2K^2T^2},$$

which concludes the proof. □

We now consider the case of two independent Wiener processes.

Lemma 3.2. *With the notation above and the assumptions of Lemma 3.1, let us consider W_1 independent of W_2 and $d = 1$. Then*

$$\mathbb{E} \sup_{0\leq t\leq T} |X^\varepsilon(t) - X(t)|^2 \leq 4T(T\|f - f^\varepsilon\|_\infty^2 + 4\sigma^2)e^{2K^2T^2}.$$

Proof. Let us compute the difference between $X^\varepsilon(t)$ and $X(t)$. Applying Young's inequality, we get

$$\mathbb{E}\sup_{0\le t\le T}|X^\varepsilon(t)-X(t)|^2 \le 2\mathbb{E}\sup_{0\le t\le T}\Big|\int_0^t (f^\varepsilon(X^\varepsilon(s))-f(X(s)))ds\Big|^2 + 2\sigma^2\mathbb{E}\sup_{0\le t\le T}\Big|\int_0^t dW_1(s)-\int_0^t dW_2(s)\Big|^2.$$

Let us define $Z(t) := W_1(t) - W_2(t)$. The process $Z(t)$ is a Wiener process with variance $2t$, thus it is a martingale. Hence, we can apply Doob's maximal quadratic inequality (*e.g.*, [12, Page 11]) to the second term, obtaining

$$\mathbb{E}\sup_{0\le t\le T}\Big|\int_0^t dW_1(s)-\int_0^t dW_2(s)\Big|^2 = \mathbb{E}\sup_{0\le t\le T}|Z(t)|^2 \le 4\mathbb{E}|Z(T)|^2 = 8T.$$

For the first term, we can apply the same technique as in Lemma 3.1. Therefore, we get the result

$$\mathbb{E}\sup_{0\le t\le T}|X^\varepsilon(t)-X(t)|^2 \le (4T^2\|f-f^\varepsilon\|_\infty^2 + 16T\sigma^2)e^{2K^2T^2}.$$

□

With those preliminary results, we can now consider the general case of a d-dimensional SDE.

Proposition 3.1. *With the notation above and if there exists a real constant K such that*

1. $\|f(x)-f(y)\| \le K\|x-y\|$, $\forall x, y \in \mathbb{R}^d$,
2. $\|f(x)\| \le K(1+\|x\|)$, $\forall x \in \mathbb{R}^d$,

then it is true for $X(t), X^\varepsilon(t)$ the solutions of (35) *and* (36) *that*

$$\mathbb{E}\sup_{0\le t\le T}\|X^\varepsilon(t)-X(t)\|_2^2 \le 4T\Big(T\sum_{i=1}^d\|f_i(x)-f_i^\varepsilon(x)\|_\infty^2 + 4d\sigma^2\Big)e^{2dK^2T^2}. \tag{38}$$

Proof. The proof follows from Lemma 3.2. Let us denote by $X_i(t), X_i^\varepsilon(t)$ the i-th component of the solution, $i = 1, \dots, d$. Then

$$\mathbb{E}\sup_{0\le t\le T}\|X^\varepsilon(t)-X(t)\|^2 = \mathbb{E}\sup_{0\le t\le T}\sum_{i=1}^d|X_i^\varepsilon(t)-X_i(t)|^2 \le \sum_{i=1}^d\mathbb{E}\sup_{0\le t\le T}|X_i^\varepsilon(t)-X_i(t)|^2$$

Then, applying Lemma 3.2 to each component of the solution, one obtains (38). □

Remark 3.1. (convergence considerations.) The result of Proposition 3.1 shows that in case one uses two different Wiener processes W_1, W_2 for (35) and (36), the strong convergence is not granted. In fact, while the first term in estimation (38) tends to zero due to the uniform convergence of the perturbed transport field towards the non-perturbed one, the error due to the different Wiener processes is independent of ε. However, in underground flow models the variance σ of the Brownian diffusion is often small with respect to the transport field. Therefore, one could consider the solution $X^\varepsilon(t)$ of (36) to be practically converging to the solution $X(t)$ of (35) in case f^ε converges to f with respect to ε.

Remark 3.2. (interpolation results.) Let us consider an interval $D = [l, r] \subset \mathbb{R}$ where f and f^ε are defined and f^ε to be a polynomial interpolation of f on a grid $x_i = l+\varepsilon i, i = 0, \dots, N, r = l+N\varepsilon$. Then, the interpolation error of can be estimated if f is regular enough. In particular, we can state the following results

1. If f is Lipschitz continuous of constant K, and f^ε is its piecewise constant interpolation computed at the midpoint of each subinterval of the grid

$$\sup_{x_i \le x \le x_{i+1}} |f(x) - f^\varepsilon(x)| \le K|x - r| \le \frac{1}{2}K\varepsilon,\ i = 0, \dots, N-1.$$

Since f is continuous

$$\|f - f^\varepsilon\|_\infty \le \frac{1}{2}K\varepsilon.$$

2. If f is of class $\mathcal{C}^1$, f^ε is its piecewise constant interpolation and for a real constant C_1 [13, Chapter 8]

$$\|f - f^\varepsilon\|_\infty \le C_1\varepsilon\|f'(x)\|_\infty.$$

3. If f is of class $\mathcal{C}^2$, f^ε is its piecewise linear interpolation and for a real constant C_2 [13, Chapter 8]

$$\|f - f^\varepsilon\|_\infty \le C_2\varepsilon^2\|f''(x)\|_\infty.$$

In all these cases, f^ε converges uniformly to f with respect to ε, therefore the solution of (36) converges to the solution of (35) with respect to ε.

3.2 Analysis of numerical convergence

We now consider the Euler-Maruyama method applied to the perturbed equation (36). We would like to find a balance between the error due to the numerical integration of the SDE with step size h and the approximation of the transport field f with f^ε. In this way, knowing ε one can choose wisely the step size h in order to avoid extra computational time.
Let us consider $X(t), X^\varepsilon(t)$ the solution of (35) and (36) respectively, and let us denote with X_n, X_n^ε the numerical solution obtained with Euler-Maruyama method applied to the two equations at time $t_n = hn, t_N = T$. It is known that if f satisfies the assumptions of Lemma 3.1, and if σ is a constant, then the strong error for X_n approximating $X(t)$ is given by (*e.g.*, [8, Chapter 10])

$$\sup_{n=0,\dots,N} \mathbb{E}\|X(nh) - X_n\| \le Ch, \tag{39}$$

for a constant C independent of h. It is possible to obtain a similar estimate for X_n^ε estimating $X(t)$.

Proposition 3.2. *Let us consider* (35), (36) *with* $W_1 = W_2 = W$ *almost everywhere. Given* $h > 0$ *such that* $hN = T$ *for* $N \in \mathbb{N}, N > 0$, *let us consider* X_n^ε *the numerical solution given by the Euler-Maruyama method applied to* (36), *i.e.,*

$$\begin{cases} X_{n+1}^\varepsilon = X_n^\varepsilon + f^\varepsilon(X_n^\varepsilon)h + \sigma(W(t_{n+1}) - W(t_n)), & n = 0, \dots, N-1, \\ X_0^\varepsilon = X_0. \end{cases}$$

Then, if f, f^ε *satisfy the assumptions of Lemma 3.1*

$$\sup_{n=0,\dots,N} \mathbb{E}\|X(nh) - X_n^\varepsilon\| \le Ch + \|f^\varepsilon - f\|_\infty \frac{e^{KT} - 1}{K}, \tag{40}$$

with C *a real constant independent of* h *and depending only on the final time* T *and the Lipschitz constant* K *of* f.

Proof. We consider X_n the numerical approximation of $X(t)$ obtained using Euler-Maruyama with the same initial condition X_0. If we add and subtract X_n and apply the triangular inequality we obtain

$$\mathbb{E}\|X_n^\varepsilon - X(nh)\| \le \mathbb{E}\|X_n^\varepsilon - X_n\| + \mathbb{E}\|X_n - X(nh)\|.$$

Since f is regular enough and σ is a constant, for the second term it is known that

$$\sup_{n=0,\dots,N} \mathbb{E}\|X_n - X(nh)\| \le Ch. \tag{41}$$

We can then make a recursive analysis of the first term. For almost all ω in Ω

$$\begin{aligned}
\|X_n^\varepsilon - X_n\| &\leq \|X_{n-1}^\varepsilon - X_{n-1}\| + h\|f^\varepsilon(X_{n-1}^\varepsilon) - f(X_{n-1})\| \\
&\leq \|X_{n-1}^\varepsilon - X_{n-1}\| + h\|f^\varepsilon(X_{n-1}^\varepsilon) - f(X_{n-1}^\varepsilon)\| + h\|f(X_{n-1}^\varepsilon) - f(X_{n-1})\| \\
&\leq h\|f^\varepsilon - f\|_\infty + (1+hK)\|X_{n-1}^\varepsilon - X_{n-1}\| \\
&\leq h\|f^\varepsilon - f\|_\infty + (1+hK)[h\|f^\varepsilon - f\|_\infty + (1+hK)\|X_{n-2}^\varepsilon - X_{n-2}\|] \\
&\overset{(\cdots)}{\leq} h\|f^\varepsilon - f\|_\infty \sum_{i=0}^{n-1}(1+hK)^i + (1+hK)^n\|X_0^\varepsilon - X_0\|
\end{aligned}$$

Since $X_0^\varepsilon = X_0$ and using the geometric series

$$\begin{aligned}
\|X_n^\varepsilon - X_n\| &\leq h\|f^\varepsilon - f\|_\infty \frac{(1+hK)^n - 1}{hK} \\
&\leq \|f^\varepsilon - f\|_\infty \frac{(1+hK)^N - 1}{K} \\
&\leq \|f^\varepsilon - f\|_\infty \frac{e^{KT} - 1}{K},
\end{aligned}$$

where the last inequality is valid since $N = T/h$ and K, T, h are all positive real numbers. Since the bound we found is independent of ω and n, we can take the expectation and the supremum, obtaining

$$\sup_{n=0,\dots,N} \mathbb{E}\|X_n^\varepsilon - X_n\| \leq \|f^\varepsilon - f\|_\infty \frac{e^{KT} - 1}{K}. \tag{42}$$

This result combined with (41) concludes the proof. □

As far as the weak convergence is concerned, it is known that the Euler-Maruyama method is of weak order one [8, Chapter 14]. The term due to the perturbation of the transport field can be treated as for the strong error, therefore we get without any further assumption for a constant C independent of h

$$\sup_{n=0,\dots,N} \|\mathbb{E}(X_n^\varepsilon - X(nh))\| \leq Ch + \|f^\varepsilon - f\|_\infty \frac{e^{KT} - 1}{K}.$$

Remark 3.3. If f^ε is the interpolation of f on a regular grid of size ε, the result of Proposition 3.2 allows to balance the interpolation error and the error due to numerical integration. In fact, we reported in Remark 3.2 some results on interpolation of f with piecewise polynomials f^ε, which we can plug in (40) as follows

1. if f is Lipschitz continuous of constant K or of class $\mathcal{C}^1$ and f^ε is a piecewise constant interpolation of f, then

$$\sup_{n=0,\dots,N} \mathbb{E}\|X_n^\varepsilon - X(nh)\| = O(h) + O(\varepsilon).$$

2. if f is of class $\mathcal{C}^2$, and f^ε is a piecewise linear interpolation of f, then

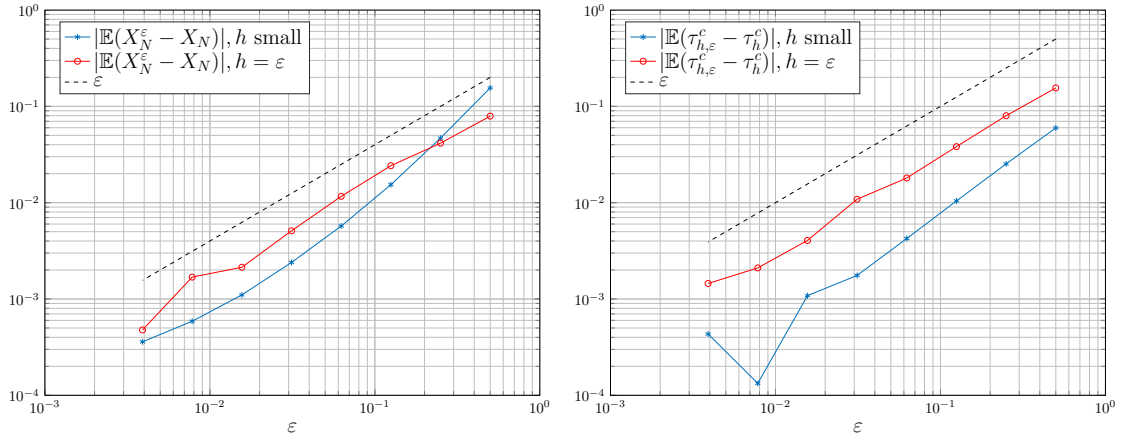
$$\sup_{n=0,\dots,N} \mathbb{E}\|X_n^\varepsilon - X(nh)\| = O(h) + O(\varepsilon^2).$$

Therefore, in the first case one should choose the step size h to be of the same order of magnitude as ε, while in the second case it should scale as ε^2.

3.3 Numerical experiments

We perform a numerical experiment in order to verify the theoretical bounds presented above. Let us consider the domain $D = [-1, 1]^2$ and the deterministic transport field given by

$$f(x, y) = \frac{1}{2}\left(x^2 + y^2, \quad |x - y|\right)^T.$$



(a) Error on the solution.

(b) Error on the exit time.

Figure 11: Convergence of the numerical solution with respect to the interpolation characteristic size ε.

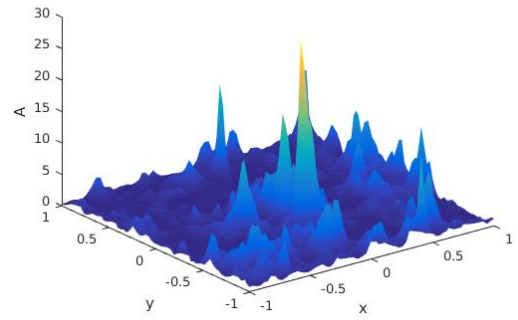
In the domain, the transport field is Lipschitz continuous. Our aim is verifying whether the numerical solution computed using a piecewise constant interpolation of f leads to convergence with respect to the characteristic size of the grid used for interpolation. In the following, f^ε denotes the piecewise constant interpolation over D of f on a structured grid of equal size ε in both directions. Since the theoretical results interest the value of the solution itself and not the mean exit time from a domain, we start by considering the convergence of the numerical solution. Then, we verify if the results are practically valid for the exit time as well.

Numerical solution. We consider the boundary conditions to be reflecting on all the boundary of D. We fix the parameters to be $T = 1, \sigma = 1, X_0 = (0, 0)^T$ and perform a Montecarlo simulation over $M = 10^4$ realizations of Euler-Maruyama. Let us remark that since the boundary is completely reflecting, there is no distinction between CEM and DEM. Maintaining the notation of Proposition 3.2, we compute the weak error of X_n^ε with respect to X_n, which should be of order $O(\varepsilon)$. The reference solution X_n is computed over $N = 2^{10}$ timesteps over the time span. Then, we vary ε in the range $\varepsilon_i = 2^{-i}, i = 1, \dots, 8$ and compute the numerical solution X_n^ε either fixing the number of timesteps $N = 2^9$ or varying it so that $h_i = \varepsilon_i$, *i.e.*, the time and space discretizations have the same order of magnitude. In the first case, we wish to overkill the error due to numerical integration, while in the second case we wish to maintain an acceptable computational cost. Moreover, the bound presented in Remark 3.3 implies that theoretically the error should be independent of h, so we would expect similar results for both the approaches. Results (Figure 11a) confirm the theoretical results, with a clear rate of convergence equal to one, and the errors which are similar for the two approaches.

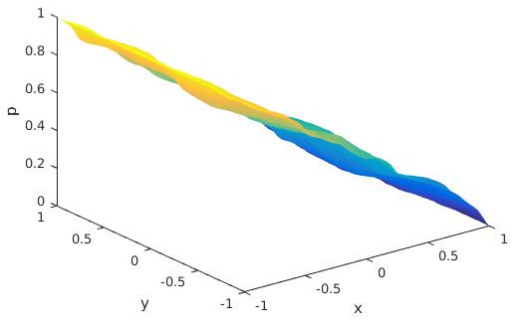
Mean exit time. We perform an experiment with the same values for all parameters as above, but we consider mixed killing and reflecting boundary conditions and we estimate the exit time τ. We wish that the theoretical results obtained for the solution apply practically to the exit time itself. We use the same strategies as above, either fixing a small step size h for any value of ε, or maintaining h equal to ε. Results (Figure 11b) show that the error is $O(\varepsilon)$ in both cases, with a smaller constant in case a small value of h is chosen. On the other side, we notice that fixing, *e.g.*, the error to 0.01, the computational time is in this case approximately sixteen times smaller in case h and ε are balanced.

4 The uncertain Darcy problem

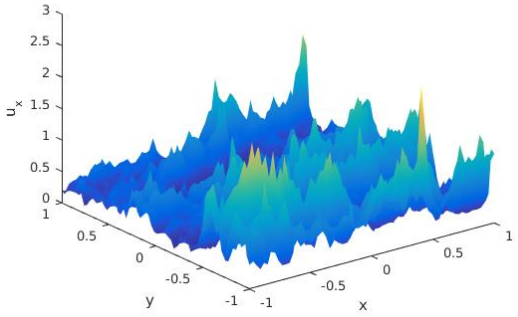
The methods for approximating the mean exit time have been investigated in a general frame. In the following we will consider (3) with $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by the solution of the uncertain Darcy problem.



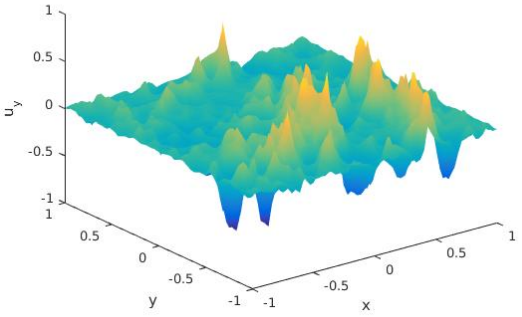
(a) Random field.



(b) Pressure field.



(c) x component of velocity field.



(d) y component of velocity field.

Figure 12: Approximate solution of the uncertain Darcy problem.

4.1 Problem statement

Let us consider a domain $D \subset \mathbb{R}^2$. Let us define the Neumann boundaries of D as Γ_N, its inlet boundary as Γ_{in} and its outlet boundary as Γ_{out}. Then, we search the solution of the following problem

$$\begin{cases} u = -A\nabla p, & \text{in } D, \\ \nabla \cdot u = 0, & \text{in } D, \\ p = p_0, & \text{on } \Gamma_{in}, \\ p = 0, & \text{on } \Gamma_{out}, \\ \nabla p = 0, & \text{on } \Gamma_N, \end{cases} \tag{43}$$

where A is a random field. The solution u of this equation is used as transport field in equation (3), which can be therefore written as

$$\begin{cases} dX(t) = u(X)dt + \sigma dW(t), & 0 < t \leq T, \\ X(0) = X_0, & X_0 \in D, \end{cases} \tag{44}$$

where we set the boundary conditions to be reflecting on Γ_N and killing on both $\Gamma_{in}, \Gamma_{out}$.

4.2 Finite Elements solution of the Darcy problem

Let us consider the domain $D = [-1, 1]^2$. The random field A in (43) is chosen to be lognormal, *i.e.*,

$$A = e^{\gamma}, \tag{45}$$

where γ is a normal random field defined by its covariance function $\mathrm{cov}_\gamma(x_1, x_2)$ for any couple of points x_1, x_2 in the domain D. The covariance function is of the Matern family [10], thus having

the following form

$$\mathrm{cov}_\gamma(x_1,x_2)=\frac{\sigma_A^2}{\Gamma(\nu)2^{\nu-1}}\Big(\sqrt{2\nu}\frac{|x_1-x_2|}{L_c}\Big)^\nu K_\nu\Big(\sqrt{2\nu}\frac{|x_1-x_2|}{L_c}\Big),\quad \nu\geq 0.5, \tag{46}$$

where σ_A^2 is the variance, L_c is the correlation length, Γ is the gamma function, K_ν is the modified Bessel function of the second kind and ν is a parameter. Let us remark that the covariance function does not depend on x_1, x_2 but only on their euclidean distance $|x_1 - x_2|$. The regularity of the covariance function and of the realizations of A depend on ν. In particular each realization of the field A is α-Hölder continuous for $0 < \alpha < \nu$. Therefore, for $\nu = 1 + \varepsilon, \varepsilon > 0$, the random field is Lipschitz continuous. Results concerning further regularity properties of A can be found in [10]. The realizations of A are computed using a discrete Fourier transformation on the vertices of a grid of D, equispaced on both the x and y directions with the same spacing Δ_A. Then, the numerical solution $\hat{p}$ of (43) is obtained with linear Finite Elements on a regular mesh T_p with maximum element size Δ_p. Since the vertices of the grid on which we compute A do not coincide with the vertices of T_p, we interpolate A on T_p. Then, the velocity field $\hat{u}$ is retrived computing the gradient of $\hat{p}$ as in equation (43). We choose to perform this procedure using the scientific software `FreeFem++`. The results for a realization of A are shown in Figure 12, where the value of the inlet pressure p_0 is equal to 1, and the parameters for the random field are $\nu = 0.5, L_c = 0.05$.

4.3 Solution of the SDE

Once the Finite Element approximation $\hat{u}$ of the velocity field is available, it is possible to approximate by means of DEM and CEM the solution of (44). The values of the numerical solution X_h can take any value in D, therefore it is necessary that the velocity field is defined in any point in D. If an interpolation of $\hat{u}$ is performed at each step, both DEM and CEM lose in computational efficiency. Hence, an interpolation of $\hat{u}$ has to be performed before the numerical integration of the SDE. Therefore, we define a grid with spacing Δ_u, interpolating the values of $\hat{u}$ in the center of each square defined by the grid (Figure 13). Let us denote by Q the set of the interpolation points, whose elements are defined by

$$\{Q\}_{ij}=\big(-1+(i-0.5)\Delta_u,\quad -1+(j-0.5)\Delta_u\big)^T,\quad i,j=1,\dots,\frac{2}{\Delta_u}=:N_u. \tag{47}$$

We compute two matrices U_x, U_y of $\mathbb{R}^{N_u\times N_u}$ containing the values of the two components of $\hat{u}$ interpolated on the points of Q. Then, the velocity field is considered to be piecewise constant in each square of the grid defined by Δ_u. Therefore, if we denote by $\tilde{u}$ the transport field for the SDE, at the i-th step of the integration $\tilde{u}$ is evaluated as follows

$$\tilde{u}(X_h(t_i))=\begin{pmatrix}U_x(\lceil (X_{h,1}(t_i)+1)/\Delta_u\rceil,\lceil (X_{h,2}(t_i)+1)/\Delta_u\rceil)\\ U_y(\lceil (X_{h,1}(t_i)+1)/\Delta_u\rceil,\lceil (X_{h,2}(t_i)+1)/\Delta_u\rceil)\end{pmatrix}, \tag{48}$$

where $X_{h,1}, X_{h,2}$ denote the first and second components of X_h and $U_x(i,j)$ represents the element (i,j) of the matrix U_x (respectively U_y). Let us remark that this operation involves only an evaluation of a matrix at each timestep, which is an operation of negligible computational cost. This implies a relevant improvement with respect to interpolating the solution at each timestep, which is on the other side a costly operation. Then, given the step size h, one step of DEM will be defined as

$$X^d_{h,i+1}=\tilde{u}(X^d_{h,i})h+\sigma(W(t_{i+1})-W(t_i)). \tag{49}$$

Given an input initial value X_0 for (44), we approximate the solution using DEM and CEM using the strategy above. The theoretical investigation presented in Section 3 guarantees that if the interpolation and the element size chosen for the Finite Element approximation tend to zero, the solution will tend to the exact solution of (43) for each realization of A. Unfortunately, the requirements of smoothness that we included in Section 3 are not satisfied by the solution of (43), unless ν is strictly bigger than one, as the random field and therefore the solution u would be Lipschitz continuous. We hope that numerically the solution will in practice provide reasonable results even for a rough transport field. In Figure 14 we display fifteen trajectories for $X_0 = (-0.8, -0.8)^T$ with two different step sizes. The choice of the initial point is made in order to observe reflections on the lower boundary of the domain D on which we compute the solution, as well as the killing boundary at the right side.

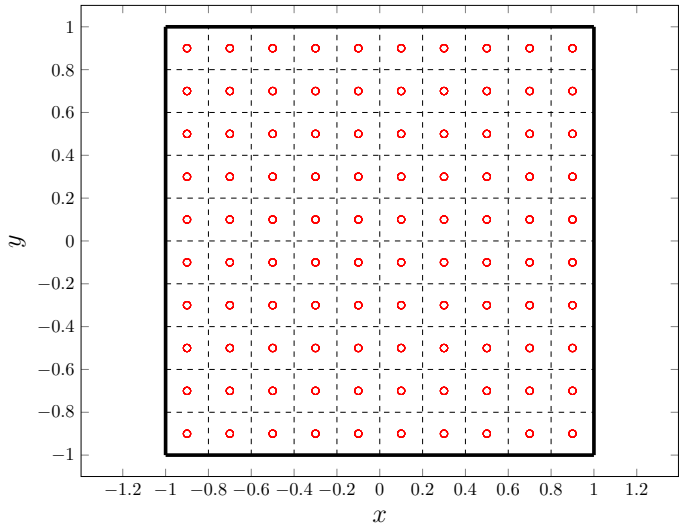
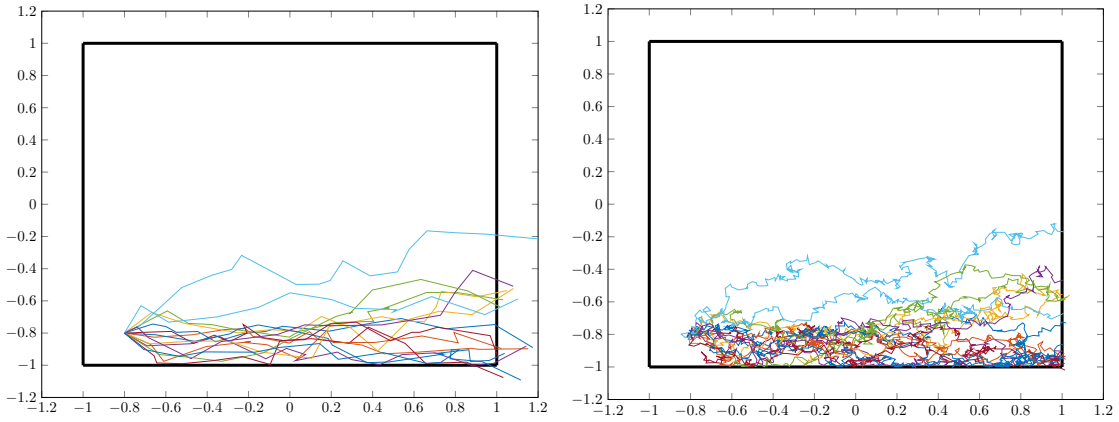


Figure 13: Grid used for interpolation of $\hat{u}$. Dots represent the interpolation points.



(a) Big timestep.

(b) Small timestep .

Figure 14: Trajectories of the numerical solution of (44) with DEM.

4.4 Summary

Let us summarize the procedure we use for approximating the solution of (44).

1. Generate A on a grid with spacing Δ_A with a Fourier transform method.
2. Approximate the solution of (43) with the Finite Element methods on a fine triangulation T_p with maximum element size Δ_p.
3. Compute the velocity field from the FEM solution and interpolate it on a grid with spacing Δ_u to obtain a piece-wise constant field $\tilde{u}$.
4. Solve (44) with DEM or CEM to evaluate the mean exit time $\bar{\tau}$ using $\tilde{u}$ as the transport field.

4.5 Estimation of the exit time

The numerical and theoretical path we followed in the previous paragraphs enables us to build a Montecarlo simulation that we exploit to estimate the mean exit time from a domain with the solution of the uncertain Darcy problem as a transport field. Thanks to all the theoretical and numerical considerations we reported above, we can perform efficiently and accurately the estimation of the exit time *for each realization* of the Darcy problem. Therefore, it is sufficient to average over different realizations of the Darcy problem in order to obtain an estimation of the exit time, as in Algorithm 2. We consider the square domain $D = [-1, 1]^2$ and the parameters listed in

Algorithm 2: Estimation of the exit time

Data: number of realizations M_r, number of trajectories M_t
Result: Estimate of the exit time $\bar{\tau}$
for $i = 1, \dots, M_r$ **do**
    Generate the random field A ;
    Find the solution of the Darcy problem ;
    Interpolate the velocity field on a structured grid of size Δ_u ;
    Estimate τ_i using CEM with step size Δ_u over M_t trajectories ;
end
$\bar{\tau} = 1/M_r \sum_{i=1}^{M_r} \tau_i$;

Table 1. We vary the value of σ and generate 100 solutions of the Darcy problem for each value of σ, estimating the exit time with 10^4 trajectories for each realization. The results we obtain are in Table 2. We remark that if the value of σ is negligible with respect to the magnitude of the transport field the exit time stabilizes on a value of approximately four seconds.

Random field				Darcy	FEM	Interpolation	Trajectories	
ν	L_c	σ_A	Δ_A	p_0	Δ_p	Δ_u	X_0	T
0.5	0.05	1	0.0039	1	$5 \cdot 10^{-3}$	0.0625	$\begin{pmatrix} -0.8, & 0 \end{pmatrix}^T$	20

Table 1: Parameters for the estimation of the exit time in the Darcy case

σ	1	0.7	0.3	0.1	10^{-2}	10^{-3}	10^{-4}
τ	0.4718	0.9906	3.1318	3.9437	4.2634	3.9767	3.9906

Table 2: Results of exit time estimations in the Darcy case for different values of σ.

5 Conclusion

We investigated the stochastic particle transport problem in the framework of underground flow. In particular, we analysed three numerical schemes that allow estimating the exit time of the solution of a general SDE from a bounded domain. Numerical experiments show that CEM seems to be the most appropriate choice in order to fulfill this purpose. In fact, estimating the probability of exit at each timestep using the Brownian bridge approach implies a small computational cost and improves dramatically the precision of DEM. On the other side, an adaptivity procedure we analysed based on the position of the trajectory in the domain succeeds in restoring the weak order of the Euler-Maruyama scheme in an unbounded domain but does not imply a considerable advantage in computational time. Then, we managed to prove the properties of convergence of the analytic and numerical solution of a perturbed SDE to its non-perturbed solution. This is fundamental when an interpolation of the transport field is needed to increase computational speed. The convergence is confirmed by numerical results on a test case, where we succeeded in tuning the parameters of interpolation and time integration in order to have good performances. Finally, we applied the numerical schemes to the Darcy case, providing the details of a double Montecarlo simulation that can be exploited to estimate the exit time of a pollutant particle in the frame of the underground flow problem. Future improvements of the procedure explained in this work could be done employing Multi Level Montecarlo techniques for estimating the exit time in the Darcy case. Moreover, the modeling of extraction wells has not been taken into account, since we feel that the integration of the SDE in presence of singularities in the velocity field would have implied a deeper theoretical investigation and different numerical techniques.

Appendix A Analytic expression of the mean exit time in the one-dimensional case

In the one dimensional case, it is possible to deduce an analytic solution of (24). Let us consider the domain $D = [l, r]$, the boundary condition at $x = l$ fixed as *killing* and vary the boundary condition at $x = r$. Since the scope is deducing the exit time of a particle from D, this assumption is plausible. In this frame, it is possible to rewrite (24) as

$$\begin{cases} f(x)\bar{\tau}'(x) + \dfrac{1}{2}g^2(x)\bar{\tau}''(x) = -1, & l < x < r, \\ \bar{\tau}(l) = 0, & \\ \bar{\tau}(r) = 0, & \text{if for } x = r \text{ the boundary is } \textit{killing}, \\ \bar{\tau}'(r) = 0, & \text{if for } x = r \text{ the boundary is } \textit{reflecting}. \end{cases} \tag{50}$$

It is possible to show [9, 11] that $\bar{\tau}$ is in the one-dimensional case given by

$$\bar{\tau}(x) = -2\int_l^x \exp(-\psi(z)) \int_l^z \frac{\exp(\psi(y))}{g^2(y)} dy + c_1 \int_l^x \exp(-\psi(y)) dy + c_2, \tag{51}$$

where the function ψ is defined as

$$\psi(x) = \int_l^x \frac{2f(y)}{g^2(y)} dy, \tag{52}$$

and the constants $c_1, c_2 \in \mathbb{R}$ depend on the boundary conditions as follows

$$\begin{aligned} c_1 &= 2\frac{\int_l^r \exp(-\psi(z)) \int_l^z \frac{\exp(\psi(y))}{g^2(y)} dy}{\int_l^r \exp(-\psi(y)) dy}, && \text{if for } x = r \text{ the boundary is } \textit{killing}, \\ c_1 &= 2\int_l^r \frac{\exp(-\psi(y))}{g(y)^2} dy, && \text{if for } x = r \text{ the boundary is } \textit{reflecting}, \\ c_2 &= 0. \end{aligned} \tag{53}$$

Let us remark that in case $f = -V'$ for some smooth function V and $g = \sigma \in \mathbb{R}$, the expression of ψ simplifies to

$$\psi(x) = 2\frac{V(l) - V(x)}{\sigma^2}. \tag{54}$$

Appendix B Numerical approximation of the exit probability with Finite Differences in the one-dimensional case

Let us consider D as the interval $[l, r]$, the boundary condition in l to be fixed to killing and in r to be either killing or reflecting. In this case, since f is in our case independent of t and $g = \sigma \in \mathbb{R}$ (26) can be written as the following initial value PDE

$$\begin{cases} -\dfrac{\partial}{\partial t}\Phi(t,x) + f\dfrac{\partial}{\partial x}\Phi(t,x) + \dfrac{1}{2}\sigma^2\dfrac{\partial^2}{\partial x^2}\Phi(t,x) = 0, & l < x < r \\ \Phi(t,l) = 1, & \\ \Phi(t,r) = 1, & \text{if for } x = r \text{ the boundary is } \textit{killing} \\ \dfrac{\partial}{\partial x}\Phi(t,r) = 0, & \text{if for } x = r \text{ the boundary is } \textit{reflecting} \\ \Phi(0,x) = 0. & \end{cases} \tag{55}$$

The solution of this equation can be approximated using finite differences. In particular, we employ the theta method. Let us consider the case in which r is a killing boundary, *i.e.*, the PDE is endowed with pure Dirichlet boundary conditions. Given a step size Δ_t for time integration and an uniform grid $x_i = l + i\Delta_x, i = 0, \dots, N+1, x_{N+1} = r$, at each timestep k one has to find the solution of the linear system

$$(I - \Delta_t\theta A)u^{k+1} = (I + \Delta_t(1-\theta)A)u^k + hF,\ 0 \le \theta \le 1, \tag{56}$$

where I is the identity matrix of $\mathbb{R}^{N\times N}$. The matrix A of $\mathbb{R}^{N\times N}$ and the vector F of $\mathbb{R}^N$ define the space discretization and the boundary conditions and are defined by

$$A = \frac{1}{2\Delta_x}\begin{pmatrix} \alpha_1 & \beta_1 & & \\ \gamma_1 & \alpha_2 & \beta_2 & \\ & \ddots & \ddots & \ddots \end{pmatrix}, \quad F = \frac{1}{2\Delta_x}\begin{pmatrix} F_1 & 0 & \cdots & 0 & F_N \end{pmatrix}^T \tag{57}$$

and the coefficients are given by

$$\begin{aligned} \alpha_i &= -\frac{2\sigma^2}{\Delta_x},\ i = 1, \dots, N, \\ \beta_i &= \frac{\sigma^2}{\Delta_x} + f(x_i),\ i = 1, \dots, N-1, \\ \gamma_i &= \frac{\sigma^2}{\Delta_x} - f(x_{i+1}),\ i = 1, \dots, N-1, \\ F_1 &= \frac{\sigma^2}{\Delta_x} - f(x_1), \\ F_N &= \frac{\sigma^2}{\Delta_x} - f(x_{N-1}). \end{aligned} \tag{58}$$

The case of reflecting boundary condition in $x = r$ is similar and affects only the computation of the matrix A and the vector F. In particular, we introduce a *ghost node* at position $x = r + \Delta_x$, compute the derivative using a central approximation and impose that it is equal to 0, which leads to the condition that the value in the node in $x = r$ is equal to the value in $x = r - \Delta_x$. Since the matrix defining the system (56) is tridiagonal, one can choose Δ_t, Δ_x to be small and obtain a precise solution of (55) in a reasonable computational time. Let us remark that for $\theta = 0.5$ the error of the numerical solution is of order 2 with respect to the time discretization.

References

[1] M. Giles and K. Ramanan, *MLMC for Multi-Dimensional Reflected Diffusions*, in SIAM Conference on Uncertainty Quantification, 2016.

[2] E. Gobet, *Weak approximation of killed di usion using Euler schemes*, Stochastic Processes and their Applications, 87 (2000), pp. 167–197.

[3] E. Gobet, *Euler schemes and half-space approximation for the simulation of diffusion in a domain*, ESAIM: Probability and Statistics, 5 (2001), pp. 261–297.

[4] E. Gobet and S. Menozzi, *Stopped diffusion processes: boundary correction and overshoot*, Stochastic Processes and their Applications, 120 (2010), pp. 130–162.

[5] D. S. Grebenkov and R. Voituriez, *Exit time distribution in spherically symmetric two-dimensional domains*, (2014), pp. 1–34.

[6] J. Helmuth, *Stochastic Simulation of Reacting Particles*, ETHZ internal report.

[7] D. J. Higham, X. Mao, M. Roj, Q. Song, and G. Yin, *Mean Exit Times and the Multilevel Monte Carlo Method*, SIAM/ASA Journal on Uncertainty Quantification, 1 (2013), pp. 2–18.

[8] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, 1992.

[9] S. Krumscheid, M. Pradas, G. A. Pavliotis, and S. Kalliadasis, *Data-driven coarse graining in action: Modeling and prediction of complex systems*, Physical Review E, 92 (2015), p. 042139.

[10] F. Nobile and F. Tesei, *A Multi Level Monte Carlo method with control variate for elliptic PDEs with log-normal coefficients*, Stochastic Partial Differential Equations: Analysis and Computations, (2015), pp. 1–47.

[11] G. A. Pavliotis, *Stochastic Processes and Applications*, Springer, 2014.

[12] P. E. Protter, *Stochastic Integration and Differential Equations*, Springer, second ed., 2004.

[13] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics*, Springer, 2007.

[14] L. Sirovich, J. E. Marsden, and S. S. Antman, *Theory and Applications of Stochastic Processes: An Analytical Approach*, vol. 170, Springer, 2010.