REVERSE-TIME DIFFUSION IN ENVIRONMENTAL MODELS

Dar'ya Spivakovs'ka
(Daria Spivakovskaya)

Reverse-time diffusion in environmental models

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op maandag 24 september 2007 om 15.00 uur
door

Dar'ya Yakivna SPIVAKOVS'KA,

Master of Science in Mathematics
de Dnipropetrovsk Nationale Universiteit, Oekraine,

geboren te Dnipropetrovsk, Oekraine (Sovjet-Unie).

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr.ir. A.W. Heemink

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr.ir. A.W. Heemink,	Technische Universiteit Delft, promotor
Prof. dr.ir. G.S. Stelling,	Technische Universiteit Delft
Prof. dr. F.M. Dekking	Technische Universiteit Delft
Prof. dr. ir. A.E. Mynett	Technische Universiteit Delft/ UNESCO-IHE/ WL | DELFT HYDRAULICS
Prof. dr. E. Deleersnijder	Université catholique de Louvain, Belgium
Prof. dr. A. Voytishek	Novosibirsk State University, Russia
Dr. J.G.M. Schoenmakers	Weierstrass Institute for Applied Analysis and Stochastics, Germany

ISBN 978-90-8559-587-8

This research was carried out in the section of Mathematical Physics at the Department of Applied Mathematics, Delft Institute of Applied Mathematics, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, The Netherlands.

Typesetting system: LaTeX 2_ε
Printed in The Netherlands by: Optima Grafische Communicatie, Rotterdam.

Mathematics has to do with everything.
Any imaginable world is governed by the same rules as the real world.
A. Pérez-Reverte in *The Flanders Panel*

To my parents

Contents

Chapter 1

Introduction

1.1 Environmental impact of oil spills

On March 24, 1989, the Exxon Valdez oil tanker grounded on a reef in Prince William Sound, 40 miles (65 km) off the Alaskan coast. According to official reports, the ship carried 53 million gallons of crude oil, of which around 11 million gallons were spilled into the waters and contaminated about 1,300 miles (2,080 km) of coastline. This figure has been accepted by the State of Alaska's Exxon Valdez Oil Spill Trustee Council, NOAA (National Oceanic and Atmospheric Administration) and environmental groups such as Greenpeace. However, some groups, for example Defenders of Wildlife believe that the spill was much larger than reported and that about 30 million gallons spilled into the ocean, pointing out that oil reclaimed from the damaged tanker (which was the basis for Exxon's calculations) was later discovered to have a large amount of seawater in it.

The effects on the local economy and wildlife were devastating. Thousands of animals died immediately; the best estimates include 250,000 seabirds, 2,800 sea otters, 300 harbour seals, 250 bald eagles, up to 22 orcas, and billions of salmon and herring eggs. Thanks to a thorough cleanup, little visual evidence of the event remained in areas frequented by humans just one year later, but the effects of the spill continue to be felt today. In the long term, reductions in population have been seen in various ocean animals, including stunted growth in pink salmon populations. Sea otters and ducks also displayed higher death rates in the following years, partly because they ingested contaminated creatures. Aanimals were also exposed to oil when they dug up their prey in dirty soil. In the fishing town of Cordova, three of five fish processors were closed in the years following the spill. In 1988 there were 270 salmon-seining boats working out of Cordova; ten years later that number had dwindled to 80.

The Exxon Valdez spill was not the largest in the world, but it still affects the environment. Researchers stated that some shoreline habitats, such as contaminated mussel beds, could take up to 30 years to recover. While it will take years for a reliable long-term study, some interim effects have already been noted; it would also have an effect on fishing as most fish would have perished because of the spillage. There are several possible causes of the accident, according to the investigation by the National Transport Safety Board. The captain of the vessel, Joseph Hazelwood, admitted drinking vodka before boarding the vessel, and failed to provide a proper navigation watch. The third mate failed to

properly manoeuvre the vessel, possibly due to fatigue and an excessive workload. The U.S. Coast Guard failed to provide an effective vessel traffic system. The list of possible reasons can be continued. For instance, the British journalist Greg Palast claimed that the main cause of the Exxon Valdez grounding was not human error but was, instead, due to an Exxon decision not to fix the ship's radar in order to save money. In any case, this accident demonstrated that there is always a risk of accident as soon as the human factor is taken into account.

Not a year goes by without the newspapers, radio and TV news reporting on yet another oil spill. Last year was no exception. On March 2, 2006 an oil spill was discovered at a pipeline owned by British Petroleum in Prudhoe Bay, Alaska. The Alaska Department of Environmental Conservation estimated that a minimum of 201,000 gallons (4,785 barrels) was spilt over 1.9 acres, although other estimates ranged up to 267,000 gallons (6,357 barrels). On July 14 and July 15, 2006 the storage tanks at the thermal power station in Jiyeh, Lebanon were damaged and leaked 20,000 to 30,000 tones of oil into the eastern Mediterranean Sea. On August 11, 2006 the oil tanker Solar I sank in the Guimaras Strait off the coast of the Guimaras and Negros Occidental provinces, causing some 500,000 liters of oil to pour into the strait. This calamity is designated as the worst oil spill the Philippines has ever seen.

Oil spills due to large ship or plants accidents are not the only source of pollution in coastal zones. Much smaller accidents also harm the ecosystem at the same level. The water is also polluted by industrial discharge of chemical wastes and byproducts, discharge of poorly-treated or untreated sewage, surface runoff containing pesticides and so on. Nowadays, many measures are being taken in order to prevent possible accidents. For instance, ships are designed in a way that will minimise the chances of a spill. However, the increasing demand for oil products and chemicals in modern society also increases the risk of water pollution. To prevent possible damage as a result of an accident, it is extremely important to predict movement of the pollutant. Advanced computer models have been used in many locations worldwide to examine the fate of pollutants in aquatic systems.

1.2 Transport modelling in the North Sea

The motivations for the study of diffusion in the North Sea are to track and monitor of pollutants. In the late -1980s, a collaborative European study (the North Sea Programme) brought together researchers from countries that border the North Sea with the object of improving our understanding of the underlying processes governing the behaviour of the sea itself, the sediments and the life within the sea. It was widely recognized that within this ambitious programme, one of the most important processes to get to grips with was diffusion [37].

The overall circulation pattern in the North Sea is reasonably well known and is free of controversy. There is an anticlockwise flow; southwards down the east coast of the UK, eastwards along the northern coasts of Belgium, the Netherlands and Germany, and then northwards along the Danish coast and finally along the west Norwegian coast to exit the North Sea. (Figure 1.2). The North Sea, being a semi-enclosed basin, is particularly vulnerable to environmental stress. Hundreds of institutions and research

(a) (b)

Figure 1.1: (a) The Exxon Valdez, aground and leaking oil (b) Wildlife was severely impacted by the oil spill. Source: http://www.evostc.state.ak.us

groups in Europe have responsibility for building numerical models to simulate pollutant spreading.

The coastal zone along the entire western and northern half of the Netherlands can be considered as one of the most densely populated areas in Europe. As a consequence, the Dutch coast is very sensitive to disturbances in its ecosystem. There are several areas in the North Sea for which a special conservation regime has been determined. One of these areas is the Wadden Sea. The Wadden Sea is famous for its rich fauna, avifauna and flora and is a popular place for recreation and wadlopen. In order to save the unique ecosystem of this and other regions it is very important to predict the outcome of the pollution. An example of numerical modelling based on the Lagrangian approach is SIMPAR developed by the Dutch National Institute for Marine and Coastal Management, the RIKZ.

1.3 Motivation

Numerical models of contaminant transport, as descriptive and predictive tools for pollutant spreading, are very important to support policy decision regarding the regulation and remediation of contamination in seas and oceans The usefulness of these models depends upon their ability to provide highly accurate information about the location and concentrations of contaminants within the subsurface. However, numerical modelling is always an approximation of reality and many phenomena are neglected in order to reduce the size of the problem. With the increased number of applications of contaminant outcome, researchers are finding that accurate models of transport of pollutants must include more physical processes. For instance, a spatial correlation of the turbulence is usually not taken into account; however, it can significantly influence contaminant spread, and pollutant concentration in certain areas may be much greater than the average value in a

Figure 1.2: The North Sea Source: http://maps.google.com

larger domain.

Given the increasing complexity of modern transport models and the increasing size of model domains it is necessary to find ways to alleviate the computational burden associated with transport modelling and to make the models more accurate. There are multiple ways to approach this problem:

- Rely upon the increasing power of computational hardware
- Increase the computational efficiency of current numerical transport simulators, for example, by parallelizing them or reducing the size of the problem
- Apply new approaches and computational methods for certain problems
- Include more physical phenomena in the current transport models

The last two approaches are the strategies adopted in this dissertation. By constructing numerical models, we are not able to consider the whole variety of physical phenomena that influence the spread of the pollutant. Usually, only the main processes are taken into account and the model consists of a system of equations. Further, the accuracy is improved by increasing the resolution of the model and making the parameters more accurate. This approach has its own limitations when the additional modifications cannot improve the accuracy of the model significantly. In order to make the prediction of pollutant spreading by numerical modelling more accurate, more complex models are required.

Traditionally, transport models are solved by discretising the domain into cells (for finite difference models) or elements (in finite element models). The governing equations

are solved for each cell or element node in the grid or mesh. This method is connected with the Eulerian approach. An alternative is using the particle tracking method and this method is associated with the Lagrangian approach. The governing advection-diffusion equation can be interpreted as a Fokker-Planck equation and the corresponding stochastic system can be derived. By numerical integration of this system particle tracks are simulated. Then, by averaging the results of such simulations, the pollutant spreading is described. Both of the methods have advantages and disadvantages. A more detailed discussion about both methods can be found in Chapter 3. In this dissertation, particle models have been used.

By using Lagrangian methods, many additional phenomena can easily be incorporated into the numerical model. Unfortunately, complicated numerical models are computationally very expensive and are not suitable for real-life applications. To handle this problem, new computational methods are required. In the present work, a new concept of reverse-time diffusion was applied to model the contaminant transport in the Dutch coastal zone. The reverse-time diffusion was firstly introduced by Thomson [125] and then used and analyzed by many researchers. The main idea is, by introducing the reverse-time variable, to interpret the Fokker-Planck equation as a backward Kolmogorov equation and derive a reverse-time stochastic model. This model is used to simulate the particle tracks in reverse time. This new method has several features:

- The reverse-time concept is very attractive for the inverse problems. It is an effective and natural way to find the possible source of pollutant releases and to construct a risk map.
- The combination of the forward and reverse-time diffusions is an effective method for solving advection-diffusion problems compared with pure forward or reverse methods. This method has been recently introduced by Milstein, Schoenmakers and Spokoiny [85].
- The reverse-time stochastic system is very close to the forward system and, as a result, the methods based on the reverse-time modelling may be easily implemented for real-life applications.

The reverse-time approach will be used for different models and its benefits will be analyzed in this dissertation. Also, new Lagrangian models for simulating transport processes in the ocean will be introduced.

1.4 Outline of the dissertation

The dissertation is structured as follows.

Chapter 2 (Stochastic differential equations) contains the background information about the mathematical theories which are used in the following chapters. The stochastic calculus is introduced and the main properties of the solution of a stochastic differential equation are analyzed (Section 2.2). The different numerical schemes for a solving a stochastic differential equation (Section 2.3) are examined, as well as finding the probability law of its solution (Section 2.4). Section 2.5 contains the analysis of the numerical error of the numerical integration.

Chapter 3 (Particle models in diffusion processes) contains the information about methods used for solving the advection-diffusion model, with particular attention paid to transport modelling of the two-dimensional vertically-integrated advection-diffusion system. The random walk models, that are consistent with three- and two-dimensional advection-diffusion equations are introduced. The literature review surveys many of the different methods (Eulerian and Lagrangian) used for transport modelling and discusses the impact of the methods upon the stability, accuracy and computational cost of a given model.

Chapter 4 (Random walk model for space-varying diffusivities) is devoted to the Lagrangian methods for transport modelling for non-diagonal space-varying diffusivity. This random walk model is applied for idealized test cases for which the analytical solutions are known. The linear modelling of the advection-diffusion process near pycnocline (Sections 4.3 and 4.4) is considered. The random walk method is also tested for the settling and diffusion model (Section 4.5). The results of this chapter are published in [107, 108].

Chapter 5 (Simulation of the transport of particles in coastal waters using forward and reverse-time diffusion) contains the results of applying the concept of the reverse-time diffusion to the Dutch coastal zone. The combination of the reverse-time and forward diffusion is used to calculate the concentrations of the pollutant in several locations along the Dutch coast.The results show that the computation time may be reduced by orders of magnitude. The application based on pure reverse-time process is given in Section 5.6 and some aspects of parallelization of the new method is considered in Section 5.7. The results of this chapter are published in [110, 111, 113, 56].

In **Chapter 6** (Two-particle models for the estimation of the mean and standard deviation of concentrations in coastal waters) a new model that takes into account the spatial correlation of the turbulence is introduced and analyzed. The new model is applied to a relatively simple test case (Section 6.3) and then used for finding the mean and standard deviations of the concentrations along the Dutch coast (Section 6.4). It is also shown that using the concept of the reverse-time diffusion leads to a reduction of the computation time. The results of this chapter are published in [112, 114].

Chapter 7 (The backward Îto method for the Lagrangian simulation of transport processes with large space variations of the diffusivity). The random walk model based on backward Îto calculation is applied in the case of the discontinuous diffusivity term. This model is tested for two problems with known analytical solutions. The results of this chapter are published in [109].

Chapter 8 (Conclusion) briefly summarizes the research and results contained within this dissertation.

Chapter 2

Stochastic differential equations

2.1 Introduction

The theory of stochastic differential equations was originally developed by mathematicians as a tool for explicit construction of the trajectories of diffusion processes for given coefficients of drift and diffusion. Since the pioneering work of Gihman, Îto and others in the early fifties, the theory of stochastic differential equations has been in the center of attention of mathematicians worked both in theoretical and in applied fields. Such popularity is explained by the wide use of the stochastic differential equations as a tool for modeling of a great number of phenomena in physics, control theory, economics and other areas.

In the first section the stochastic differential equations are introduced and the conditions of the existence and uniqueness of the solution are studied. In the next section we investigate numerical methods for stochastic differential equations. Finally, some properties of the transition density function, such as Kolmogorov backward and forward equations and Chapman-Kolmogorov equation are considered.

In this chapter only the basic definitions and properties of the stochastic differential equations and the transition density function are summarized. The theoretical background can be found in [2, 36, 46, 91, 62]. The aspects of the numerical integration of the stochastic differential equations are considered in [68, 69, 84, 86]. The kernel estimator for the density functions is discussed in [106, 144].

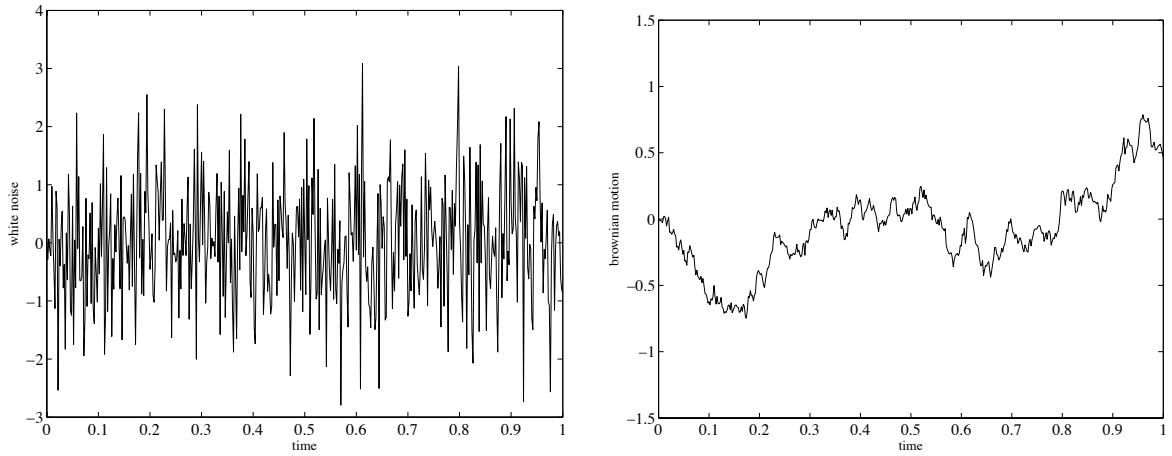
2.2 Stochastic differential equations

A number of processes in physics, economics, biology, chemistry and others can be described with the help of differential equations. The theory of differential equations is well developed and modern computers allow to simulate the complicated processes accurately. However, a lot of natural phenomena are influenced by random disturbances. For example, we can consider the following problem.

Suppose we study the growth of the population of species

$$dX(t) = a(t)X(t)dt, \quad X(t_0) = x_0 \tag{2.1}$$

where $X(t)$ is the size of the population at time t and $a(t)$ is the relative rate of growth



(a) White noise (b) Brownian motion

Figure 2.1: In this figure an example of a white noise process is shown. (a) This functions is discontinuous in any point and it can be not integrated in the sense of Lebesgue or Riemann integrals. The Brownian motion (or Wiener process) (b) can be considered as a formal integral of the white noise process.

at time t. It might happen that $a(t)$ is not completely known, because of the influence of some random environmental factors, so we have

$$a(t) = r(t) + \text{"noise"}$$

For solving these kind of problems we need to introduce the proper mathematical instrument called "stochastic differential equations".

2.2.1 The definition of stochastic differential equation

Suppose we need to solve Equation (2.1) or more generally the equation

$$\frac{dX(t)}{dt} = a(t, X(t)) + \sigma(t, X(t))\xi(t) \tag{2.2}$$

where the function $\xi(t)$ represents the noise. For any value of the parameter $t \in [t_0, T]$, $\xi(t)$ is a random variable. Such a parameterized collection of random variables is called a stochastic process. Further we assume that the stochastic process $\xi(t)$ is Gaussian, i.e. for all $t_0 \leq t_1 \leq \ldots \leq t_n \leq T$ the random vector $\boldsymbol{Z} = (\xi(t_1), \ldots, \xi(t_n)) \in \mathbb{R}^n$ has a normal distribution. That means that there exists a vector $\boldsymbol{m} = (m_1, \ldots, m_n) \in \mathbb{R}^n$ and a symmetric positive definite matrix $\boldsymbol{A} = [a_{ij}] \in \mathbb{R}^n \times \mathbb{R}^n$, $1 \leq i, j \leq n$ that the random vector $\boldsymbol{Z}$ has a density of the form

$$p_{\boldsymbol{Z}}(x_1, \ldots, x_n) = \frac{\sqrt{\det A}}{(2\pi)^{n/2}} \exp\left(\frac{1}{2}\sum_{i,j=1}^{n}(x_i - m_i)a_{ij}(x_j - m_j)\right)$$

Furthermore assume that $\xi(t)$ is a "white noise" process, i.e. $\xi(t)$ satisfies the following conditions

- $E(\xi(t)) = 0$
- $E(\xi(t_1)\xi(t_2)) = \delta(t_2 - t_1), \quad t_2 \geq t_1$

where $\delta(t)$ is the Dirac delta function. The name 'white noise' comes from the fact that such a process has a spectrum in which all frequencies participate with the same intensity, which is characteristic of white light.

The formal integration of the equation (2.2) leads us to the equation

$$X(t) = X(t_0) + \int_{t_0}^{t} a(\tau, X(\tau))d\tau + \int_{t_0}^{t} \sigma(\tau, X(\tau))\xi(\tau)d\tau \tag{2.3}$$

However, it is impossible to find the last integral in (2.3) using only the standard mathematical instruments known from the real analysis. The new mathematical theory that allows to solve the equation (2.3) was developed in the middle of the last century by Îto, K., Kolmogorov, A.N., Stratonovich, R.L. Before we turn to the proper mathematical interpretation of the integrals (2.3) we need to introduce Brownian motion. The Gaussian stochastic process $W(t)$ is called Brownian motion (or Wiener process) if

- $E(W(t_2) - W(t_1)) = 0$
- $E((W(t_4) - W(t_3))(W(t_2) - W(t_1))) = 0, \quad t_4 \geq t_3 \geq t_2 \geq t_1$
- $E(W(t_2) - W(t_1))^2 = (t_2 - t_1), \quad t_2 \geq t_1$

Formally we can consider the white noise process $\xi(t)$ as a derivative of Brownian motion $W(t)$, however, the derivative of Brownian motion does not exist in the ordinary sense (see, for instance, [62])

$$dW(t) = \xi(t)dt \tag{2.4}$$

Equation (2.3) can be written in the form

$$X(t) = \int_{t_0}^{t} a(\tau, X(\tau))d\tau + \int_{t_0}^{t} \sigma(\tau, X(\tau))dW(\tau) \tag{2.5}$$

or in differential form

$$dX(t) = a(t, X(t))dt + \sigma(t, X(t))dW(t) \tag{2.6}$$

Equation (2.6) is called the stochastic differential equation.

To make the definition of the stochastic differential equation (2.5) complete we need first to define the stochastic integral

$$F(t) = \int_{t_0}^{t} f(\tau)dW(\tau)$$

from some function $f(\tau)$. The integral is constructed in similar to the construction of the Riemann-Stieltjes integral: we approximate the integral $F(t)$ by the sums

$$\sum_{k=0}^{L-1} f(t_k^*)(W_{k+1} - W_k) \tag{2.7}$$

where $t_0 < t_1 < \dots t_L = T$, t_k^* is an arbitrary point on $[t_k, t_{k+1})$ and $W_k = W(t_k), \quad k = 1, \dots L$. Here are, however, at least two major differences between the Riemann-Stieltjes integral and the stochastic integral. The first difference is the type of convergence: the Riemann-Stieltjes integral converges in $\mathbb{R}$, while the stochastic integral converges in the space L_2. Furthermore, the value of the Riemann-Stieltjes integral does not depend on the choice of the points t_k^* , while it does make a difference for the stochastic integral. The following two choices have turned out to be the most useful

- $t_k^* = t_k$ (the left end point), which leads to the Îto integral, denoted by

$$\int_{t_0}^{t} f(\tau) dW(\tau)$$

- $t_k^* = (t_k + t_{k+1})/2$ (the mid point), which leads to the Stratonovich integral, denoted by

$$\int_{t_0}^{t} f(\tau) \circ dW(\tau)$$

Further, in Chapter 7, we consider also the backward Îto integral that can be obtained by choosing $t_k^* = t_{k+1}$.

Of course, the following natural question arises: which interpretation provides us the proper mathematical model of the differential equation with "white noise" (2.2). Actually, both calculi may be used as mathematical model of equation (2.2). The choice of integral depends on the real situation being modelled. It is possible to convert an Îto stochastic differential equation to Stratonovich form and vice versa. Suppose that the Îto stochastic differential equation is given by equation (2.6). Then the related Stratonovich stochastic differential equation can be written in the form

$$dX(t) = \overline{a}(t, X(t))dt + \sigma(t, X(t)) \circ dW(t) \tag{2.8}$$

where

$$\overline{a}(t, x) = a(t, x) - \frac{1}{2}\frac{\partial \sigma}{\partial x}(t, x)\sigma(t, x) \tag{2.9}$$

From the last equation it is clear that in the case of so called additive noise, i.e. $\sigma(t, x) \equiv \sigma(t)$ depends only on the time, the Îto and Stratonovich stochastic differential equations have the same form.

Both the Îto and Stratonovich interpretations have some advantages and disadvantages. The stochastic differential equation in the Stratonovich sense satisfies the usual rule of derivation, while for the Îto integral we have to use more complicated rule (2.10).

On the other hand, the feature of the Îto model of "not being into the future" [91] seems to be a reason for choosing an Îto interpretation in many cases. More information about the comparison of Îto and Stratonovich can be found in [91, 140] (see also Chapter 7).

Further, speaking about the stochastic integral, we will usually mean the Îto integral. Of course, for the complete definition of the Îto integral we should prove the existence of the limit of the sums (2.7) when $\Delta t_k \to 0$. However, this is beyond the scope of this thesis and can be found, for instance in [2, 46, 91, 62, 16].

Here we consider only some properties of the Îto integral. Let f and g are two functions, that can be integrated in the Îto sense (for instance, continuous functions on $[t_0, T)$, then

- $\int\limits_{t_0}^{T} f(\tau)dW(\tau) = \int\limits_{t_0}^{t^*} f(\tau)dW(\tau) + \int\limits_{t^*}^{T} f(\tau)dW(\tau)$
- $\int\limits_{t_0}^{T} (\alpha f(\tau) + \beta g(\tau))dW(\tau) = \alpha \int\limits_{t_0}^{T} f(\tau)dW(\tau) + \beta \int\limits_{t_0}^{T} g(\tau)dW(\tau)$, where $\alpha, \beta \in \mathbb{R}$
- $E\left[\int\limits_{t_0}^{T} f(\tau)dW(\tau)\right] = 0$
- $E\left(\left|\int\limits_{t_0}^{T} f(\tau)dW(\tau)\right|^2\right) = E\left(\int\limits_{t_0}^{T} |f(\tau)|^2 d\tau\right)$

One of the most important properties of the process $X(t)$ is the Markov property: the future behavior of the process given what has happened up to time t is the same as the behavior obtained when starting the process at the point $X(t)$. The precise mathematical formulation of the Markov property can be found, for instance, in [16, 46, 91].

Another important result was discovered by K.Îto, in 1951 [61] and it is known as the Îto rule.

Theorem. (Îto rule)
Suppose that the stochastic process $X(t)$ is the solution of Equation (2.4). Let $g(t, x)$ be a twice continuously differentiable on $[0, \infty) \times \mathbb{R}$ function. Then the stochastic process $Y(t) = g(t, X(t))$ satisfies the stochastic differential equation in the form

$$dY(t) = \left[\frac{\partial g}{\partial t}(t, X(t)) + \frac{1}{2}\sigma^2 \frac{\partial^2 g}{\partial x^2}(t, X(t))\right] dt + \frac{\partial g}{\partial x}(t, X(t))dX(t) \tag{2.10}$$

The definition of the one-dimensional stochastic differential equation can be extended into the case of d-dimensional stochastic process $\boldsymbol{X} = (X_1, \ldots, X_n)^T$. The d-dimensional stochastic differential equation has the form

$$d\boldsymbol{X}(t) = \boldsymbol{a}(t, \boldsymbol{X}(t))dt + \boldsymbol{\sigma}(t, \boldsymbol{X}(t))d\boldsymbol{W}(t) \tag{2.11}$$

where $\boldsymbol{a} = (a_1, \ldots, a_n)^T$ is a d-dimensional function, $\boldsymbol{\sigma} = \{\sigma_{ij}\}$, $1 \le i \le d$, $1 \le j \le m$ is a $d \times m$ matrix function and m-dimensional Brownian motion $\boldsymbol{W} = (W_1, \ldots, W_m)^T$ has the following statistics

- $E(\boldsymbol{W}(t_2) - \boldsymbol{W}(t_1)) = 0$
- $E((\boldsymbol{W}(t_4) - \boldsymbol{W}(t_3))(\boldsymbol{W}(t_2) - \boldsymbol{W}(t_1))) = 0, \quad t_4 \geq t_3 \geq t_2 \geq t_1$
- $E((\boldsymbol{W}(t_2) - \boldsymbol{W}(t_1))(\boldsymbol{W}(t_2) - \boldsymbol{W}(t_1))) = (t_2 - t_1)\boldsymbol{I}_m, \quad t_2 \geq t_1$

where $\boldsymbol{I}_m$ is a $m \times m$ identity matrix.

Similarly to the one-dimensional case there is an explicit multidimensional formula which expresses the Îto interpretation (2.11) in terms of Stratonovich integrals as follows

$$\begin{aligned} d\boldsymbol{X}(t) &= \overline{\boldsymbol{a}}(t, \boldsymbol{X}(t))dt + \boldsymbol{\sigma}(t, \boldsymbol{X}(t)) \circ d\boldsymbol{W}(t) \\ \overline{a}_i(t, \boldsymbol{x}) &= \overline{a}_i(t, \boldsymbol{x}) - \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{d}\frac{\partial \sigma_{ij}(t, \boldsymbol{x})}{\partial x_k}\sigma_{kj}(t, \boldsymbol{x}), \quad 1 \leq i \leq n \end{aligned} \tag{2.12}$$

The following theorem shows the general Îto formula in the multi-dimensional case

Theorem. (multi-dimensional Îto rule)
Suppose that the d-dimensional stochastic process $\boldsymbol{X}(t)$ is the solution of the stochastic differential equation (2.11). Let $\boldsymbol{g}(t, \boldsymbol{x}) = (g_1(t, \boldsymbol{x}), \ldots, g_p(t, \boldsymbol{x}))$ be twice continuously differentiable map with respect to all variables (map from $[0, \infty) \times \mathbb{R}^d$ into $\mathbb{R}^p$). Then the kth component ($k = 1, \ldots, p$) of the stochastic process $\boldsymbol{Y} = g(t, \boldsymbol{X}(t))$ satisfies the stochastic differential equation in the form ($\boldsymbol{b} = \boldsymbol{\sigma}^T\boldsymbol{\sigma}$)

$$dY_k = \left[\frac{\partial g_k}{\partial t}(t, \boldsymbol{X}(t)) + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2 g_k}{\partial x_i \partial x_j}(t, \boldsymbol{X}(t))b_{ij}(t, \boldsymbol{X}(t))\right]dt + \sum_{i=1}^{d}\frac{\partial g_k}{\partial x_i}(t, \boldsymbol{X}(t))dX_i(t) \tag{2.13}$$

Example 1. (The linear stochastic differential equation in the narrow sense)
Suppose that the function $\boldsymbol{a}(t, \boldsymbol{x})$ in (2.11) is linear in the variable $\boldsymbol{x} \in \mathbb{R}^d$ on $[t_0, T] \times \mathbb{R}^d$, in other words,

$$\boldsymbol{a}(t, \boldsymbol{x}) \equiv \boldsymbol{C}(t)\boldsymbol{x} + \boldsymbol{c}(t)$$

where $\boldsymbol{C}(t)$ is $d \times d$-matrix and $\boldsymbol{c}(t)$ is $\mathbb{R}^d$-valued and the diffusion term $\boldsymbol{\sigma}(t, \boldsymbol{x})$ depends only on the time variable t

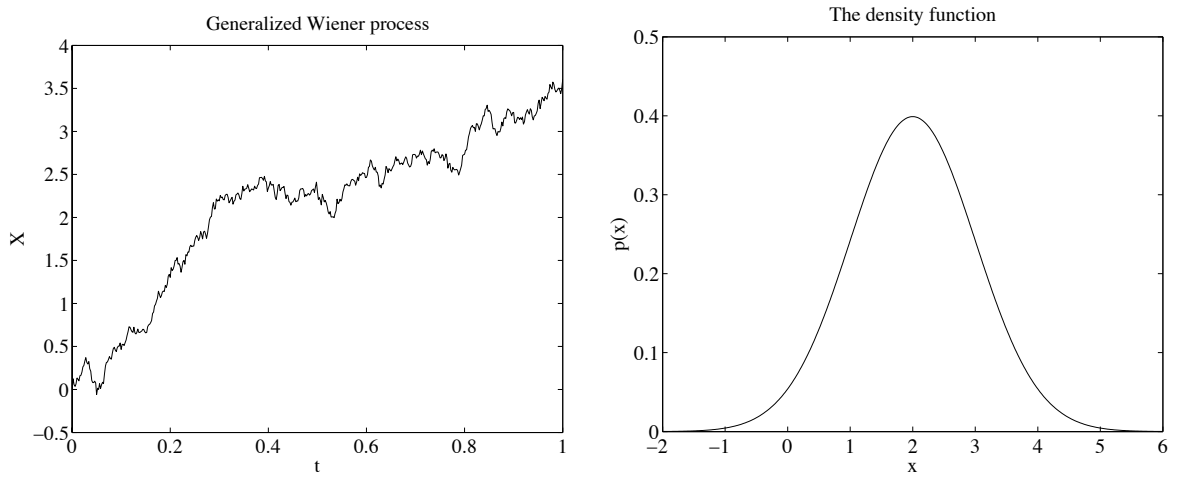
$$\boldsymbol{\sigma}(t, \boldsymbol{x}) \equiv \boldsymbol{\sigma}(t)$$

In this case Equation (2.11) is called the linear equation (in the narrow sense). In the more general situation the diffusion term $\boldsymbol{\sigma}$ is also linear, but we restrict ourselves only to the case of the linear equation in the narrow sense.

It can be shown (see, for instance, [2]) that the stochastic process $X(t)$ has on $[t_0, T]$ the solution

$$\boldsymbol{X}(t) = \boldsymbol{\Phi}(t)\left(\boldsymbol{x_0} + \int_{t_0}^{t}\boldsymbol{\Phi}(\tau)^{-1}\boldsymbol{c}(\tau)d\tau + \int_{t_0}^{t}\boldsymbol{\Phi}(\tau)^{-1}\boldsymbol{\sigma}(\tau)d\boldsymbol{W}(\tau)\right) \tag{2.14}$$

where $\boldsymbol{\Phi}(\tau)$ is the fundamental matrix of the deterministic equation $\dot{\boldsymbol{X}} = \boldsymbol{C}(t)\boldsymbol{X}(t)$. In this formula the solution $\boldsymbol{X}(t)$ is represented as the sum of three statistically independent



(a) Generalized Wiener process (b) The density function

Figure 2.2: (a) On this figure a realization of the one-dimensional generalized Wiener process given by (2.16) is shown. The parameters were chosen as follows $a = 1$, $\sigma = 1$, $t_0 = 0$, $T = 1$, $x_0 = 0$. (b) The density function of $X(T)$, that, in this case, is the density function of one-dimensional normally distributed random variable with parameters $(2, 1)$ is shown.

terms and it is always normally distributed

$$\boldsymbol{X}(t) \sim \phi\left(\boldsymbol{\Phi}\boldsymbol{x}_0 + \boldsymbol{\Phi}(t)\int_{t_0}^{t}\Phi(\tau)^{-1}\boldsymbol{c}(\tau)d\tau, \int_{t_0}^{t}\boldsymbol{\Phi}(t)\boldsymbol{\Phi}(\tau)^{-1}\boldsymbol{\sigma}(\tau)\boldsymbol{\sigma}(\tau)^T(\boldsymbol{\Phi}(\tau)^{-1})^T\boldsymbol{\Phi}(t)^T d\tau\right) \tag{2.15}$$

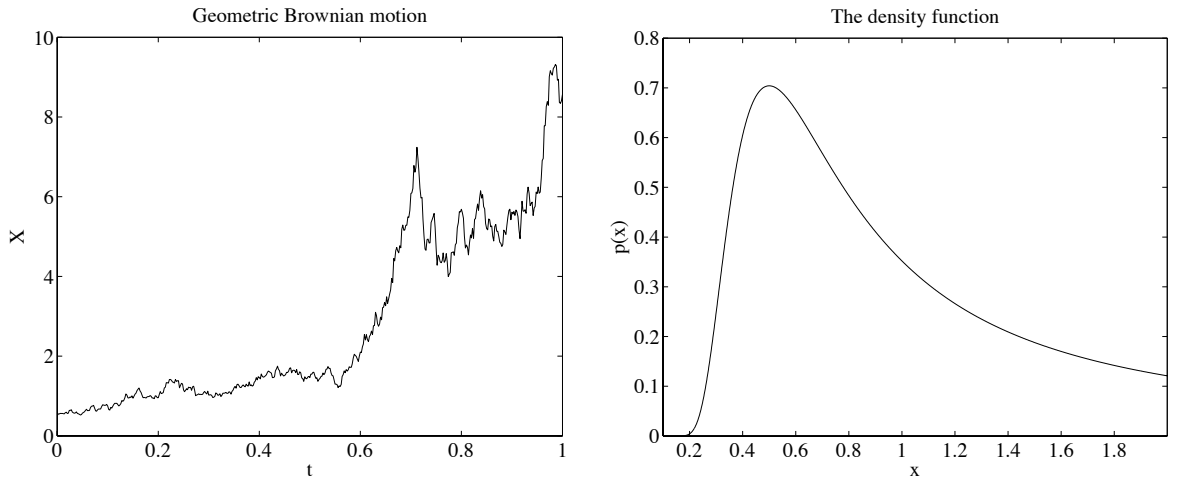
The detailed discussion of the linear differential equations and the proof of formulae (2.14) and (2.15) can be found in [2, 46]. In the simplest case, when the functions $\boldsymbol{a}(t, \boldsymbol{x}) \equiv \boldsymbol{a}$ and $\boldsymbol{\sigma}(t, \boldsymbol{x}) \equiv \boldsymbol{\sigma}$ are the constants, the solution $\boldsymbol{X}(t)$ is the generalized Wiener process

$$\boldsymbol{X}(t) = \boldsymbol{x}_0 + \boldsymbol{a}(t - t_0) + \boldsymbol{\sigma}\boldsymbol{W}(t) \tag{2.16}$$

and the density function of the random variable $\boldsymbol{X}(t)$ has the form ($\boldsymbol{b} := \boldsymbol{\sigma}^T\boldsymbol{\sigma}$)

$$p(\boldsymbol{x}) = ((2\pi)^d \det(\boldsymbol{b}))^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0 - \boldsymbol{a}(t - t_0))^T\boldsymbol{b}^{-1}(\boldsymbol{x} - \boldsymbol{x}_0 - \boldsymbol{a}(t - t_0))\right) \tag{2.17}$$

Example 2. (The Geometric Brownian motion)
Let us consider the stochastic process that is used as a model of the stock price behavior in financial mathematics. Suppose that $X(t)$ is the stock price at time $t \in [t_0, T]$ ($X(t_0) = x_0$). The expected drift rate in X should be assumed to be aX for some constant parameter a. A reasonable assumption is that the standard deviation of the change in a short period



(a) Geometric Brownian motion (b) The density function

Figure 2.3: (a) On this figure the realization of the Geometric Brownian motion given by (2.18) is shown. The parameters were chosen as follows $a = 1$, $\sigma = 1$, $t_0 = 0$, $T = 1$, $x_0 = 0.5$. (b) The density function of the random walk $X(T)$ given by (2.19) is shown

of time Δt should be proportional to the stock price. These assumptions lead to the following model

$$\begin{aligned} dX(t) &= aX(t)dt + \sigma X(t)dW(t) \\ X(t_0) &= x_0 \end{aligned} \tag{2.18}$$

where σ is the volatility of the stock price. This process is called geometric Brownian motion and can be used as model of the population growth (2.1).

Let consider the stochastic process $\ln X(t)$ and use the Îto formula (2.10). In this case

$$g(t,x) = \ln x, \quad \frac{\partial g}{\partial x} = \frac{1}{x}, \quad \frac{\partial^2 g}{\partial^2 x} = -\frac{1}{x^2}$$

and

$$d\ln X(t) = \left(a - \frac{\sigma^2}{2}\right) dt + \sigma dW(t)$$

Hence, the random variable $\ln X(t)$ is normally distributed (we use the previous example)

$$\ln X(t) \sim \phi\left(\ln x_0 \left(a - \frac{\sigma^2}{2}\right)(t - t_0), \sigma^2(t - t_0)\right)$$

and the stochastic process $X(t)$ has the lognormal distribution with the following density function

$$p(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{x\sqrt{2\pi\sigma^2(t-t_0)}} \exp\left(-\frac{(1/x-(a-\sigma^2/2)(t-t_0))^2}{2\sigma^2(t-t_0)}\right), & x > 0 \end{cases} \tag{2.19}$$

The geometric Brownian motions are very important as models for stochastic prices in financial mathematics [57].

2.2.2 Existence and uniqueness theorem of stochastic differential equation

The linear equation in Example 1 has the explicit solution (2.6). However, in the most cases we have no explicit solutions and, therefore, we need somehow to ensure the existence and uniqueness of a process

$$\boldsymbol{X}(t) = \{\boldsymbol{X}(t) : t \in [t_0, T],\, \boldsymbol{X}(t_0) = \boldsymbol{x}_0\}$$

which satisfies (2.6).

We shall say that the solution of (2.6) is unique in the strong sense if any two such solutions

$$\begin{aligned}\boldsymbol{X} &= \{\boldsymbol{X}(t) : t \in [t_0, T],\, \boldsymbol{X}(t_0) = \boldsymbol{x}_0\} \text{ and}\\ \tilde{\boldsymbol{X}} &= \{\tilde{\boldsymbol{X}}(t) : t \in [t_0, T],\, \tilde{\boldsymbol{X}}(t_0) = \boldsymbol{x}_0\}\end{aligned}$$

have, almost surely, the same paths on $[t_0, T]$, that is if

$$P\left(\sup_{t_0 \le t \le T} |\boldsymbol{X}(t) - \tilde{\boldsymbol{X}}(t)| > 0\right) = 0$$

It is natural to ask what conditions should satisfy the vector function $\boldsymbol{a}(t, \boldsymbol{x})$ and the matrix function $\boldsymbol{\sigma}(t, \boldsymbol{x})$ to provide the existence and uniqueness of the solution of (2.11). The answer gives as the following theorem [91, 62, 46, 2].

Theorem. (Existence and uniqueness theorem for stochastic differential equations) *Suppose that functions* $\boldsymbol{a}(\cdot, \cdot) : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\boldsymbol{\sigma}(\cdot, \cdot) : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ *are measurable functions satisfying*

$$|\boldsymbol{a}(t, \boldsymbol{x})| + |\boldsymbol{\sigma}(t, \boldsymbol{x})| \le \mathcal{K}_1(1 + |\boldsymbol{x}|), \quad \boldsymbol{x} \in \mathbb{R}^d, t \in [t_0, T] \tag{2.20}$$

for some constant $\mathcal{K}_1$, *(where* $|\boldsymbol{\sigma}| = \sum |\sigma_{ij}|^2$*) and such that*

$$|\boldsymbol{a}(t, \boldsymbol{x}) - \boldsymbol{a}(t, \boldsymbol{y})| + |\boldsymbol{\sigma}(t, \boldsymbol{x}) - \boldsymbol{\sigma}(t, \boldsymbol{y})| \le \mathcal{K}_2|\boldsymbol{x} - \boldsymbol{y}|; \quad \boldsymbol{x} \in \mathbb{R}^d, t \in [t_0, T] \tag{2.21}$$

for some constant $\mathcal{K}_2$. *Then, stochastic differential equation (2.11) with the initial condition* $\boldsymbol{X}(t_0) = \boldsymbol{x}_0$ *has a unique t-continuous solution* $\boldsymbol{X}(t)$ *with the property that*

$$E\left[\int_{t_0}^{T} |\boldsymbol{X}(t)| dt\right] < \infty \tag{2.22}$$

In some situations Equation (2.11) may have solutions which are unique in the weaker sense: only their probability laws coincide, but not necessary their sample paths. In this case we shall say that we have a unique weak solution.

A strong solution is of course also a weak solution, but the converse is not true. The example of the stochastic differential equation, that has only weak solutions is the Tanaka equation

$$\begin{aligned}dX(t) &= \text{sign}(X(t))dW(t)\\ X(t_0) &= x_0\end{aligned}$$

Further, unless something else is said, speaking about the solution of the stochastic differential equation we will mean the weak solution. Also, we will assume that the $d \times d$ matrix $\boldsymbol{b} := \boldsymbol{\sigma}^T \boldsymbol{\sigma}$ is of full rank for every $(t, \boldsymbol{x}) \in [t_0, T] \times \mathbb{R}^d$ and the functions $a_i(t, \boldsymbol{x})$ and $\sigma_{ij}(t, \boldsymbol{x})$ and their first derivatives are continuous and bounded. This particularly implies the implementation of the conditions (2.20) and (2.21) and, as a result, the existence and uniqueness of the solutions of (2.11) in the strong and weak senses, smoothness of the transition density function $p(t, \boldsymbol{x}, s, \boldsymbol{y})$ $(t_0 \le t \le s \le T)$ of the stochastic process $\boldsymbol{X}(t)$ and existence of all the moments $p(\cdot, \cdot, \cdot, \boldsymbol{y})$.

2.3 Numerical solution of the stochastic differential equations

For the most of the stochastic differential equations it is impossible to find the explicit solution, so we have to use some numerical scheme to obtain the approximation of the solution of (2.11). A lot of methods for numerical solution of ordinary differential equations can be extended for the stochastic differential equations, for example, methods based on the Taylor type expansion of the solution. However, the high order schemes for the stochastic differential equations are very complicated and require too much CPU time for their implementation, especially for high dimensional systems. There are many books and papers about the numerical integration of stochastic differential equations that have been recently published (see, for instance, [3, 68, 69, 84, 86]).

2.3.1 The explicit Euler scheme

The simplest and more often used method is the Euler scheme, also called the Euler-Maruyama scheme [82]. This scheme is one-step approximation method and has the following form

$$\begin{aligned} \overline{\boldsymbol{X}}_{k+1}^{\Delta t} &= \overline{\boldsymbol{X}}_{k}^{\Delta t} + \boldsymbol{a}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta t + \boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta \boldsymbol{W}_k \\ \overline{\boldsymbol{X}}_{0}^{\Delta t} &= \boldsymbol{x}_0 \end{aligned} \tag{2.23}$$

where $k = 0, \ldots, L-1$, $\overline{\boldsymbol{X}}_{k}^{\Delta t} := \overline{\boldsymbol{X}}^{\Delta t}(t_k)$ is the numerical approximation of the position $\boldsymbol{X}(t_k)$, $t_k = t_0 + k\Delta t$, $\Delta t = (T - t_0)/L$ is the time step of numerical integration and $\Delta \boldsymbol{W}_k$ are mutually independent Gaussian variable with zero mean and covariance matrix $\Delta t \boldsymbol{I}_m$.

In the previous section we examined two solutions of stochastic differential equation (2.11): in the strong sense and in the weak sense. If we consider the numerical scheme (2.23) as the approximation of the strong solution, we speak about strong converge and say that the approximation $\overline{\boldsymbol{X}}^{\Delta t}$ converges strongly with order $\vartheta > 0$ at time T if there exists a positive constant $\mathcal{K}$, which does not depend on Δt, such that

$$E\left(|\boldsymbol{X}(T) - \overline{\boldsymbol{X}}^{\Delta t}(T)|\right) \le \mathcal{K}\Delta t^{\vartheta} \tag{2.24}$$

The weak convergence concerns the consideration of (2.23) as the approximation of the weak solution and we shall say that an approximation $\overline{\boldsymbol{X}}^{\Delta t}$ converges to the solution

$\boldsymbol{X}$ of (2.11) at time T with weak order $\chi > 0$ if for each positive polynomial g there exists a positive constant $\mathcal{K}$, which does not depend on Δt such that

$$\epsilon_{\text{app}} = \left| E(g(\boldsymbol{X}(T))) - E(g(\overline{\boldsymbol{X}}^{\Delta t}(T))) \right| \leq \mathcal{K}\Delta t^{\chi} \tag{2.25}$$

In general, the Euler scheme converges with weak order $\chi = 1$ in contrast with the strong order $\vartheta = 0.5$. We shall say that the strong and weak criteria lead to the development of different approximation schemes which are only efficient with respect to one of the two criteria. Because in our application the approximation of some functional of the stochastic process $\boldsymbol{X}$ is more important than the pathwise approximation, we will interest only the order of weak convergence.

2.3.2 The semi-implicit Euler scheme

The simplest implicit weak scheme is the implicit Euler method, which has the form

$$\begin{aligned} \tilde{\boldsymbol{X}}_{k+1}^{\Delta t} &= \overline{\boldsymbol{X}}_{k}^{\Delta t} + \boldsymbol{a}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta t + \boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta \boldsymbol{W}_k \\ \overline{\boldsymbol{X}}_{k+1}^{\Delta t} &= \overline{\boldsymbol{X}}_{k}^{\Delta t} + \boldsymbol{a}(t_{k+1}, \tilde{\boldsymbol{X}}_{k+1}^{\Delta t})\Delta t + \boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta \boldsymbol{W}_k \\ \overline{\boldsymbol{X}}_{0}^{\Delta t} &= \boldsymbol{x}_0 \end{aligned} \tag{2.26}$$

where $k = 0, \ldots, L-1$, $\overline{\boldsymbol{X}}_{k}^{\Delta t}$, $\Delta \overline{\boldsymbol{W}}_k$ and Δt means the same as in the explicit Euler scheme (2.23).

The implicit Euler method (2.26) has the same order of accuracy as the explicit method (2.23), however it is more stable [68]. On the other hand the implicit method is more computationally expensive than the explicit one.

We can also form a family of implicit Euler scheme

$$\begin{aligned} \tilde{\boldsymbol{X}}_{k+1}^{\Delta t} &= \overline{\boldsymbol{X}}_{k}^{\Delta t} + \boldsymbol{a}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta t + \boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta \boldsymbol{W}_k \\ \overline{\boldsymbol{X}}_{k+1}^{\Delta t} &= \overline{\boldsymbol{X}}_{k}^{\Delta t} + \left\{ (1-\alpha)\boldsymbol{a}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t}) + \alpha \boldsymbol{a}(t_{k+1}, \tilde{\boldsymbol{X}}_{k+1}^{\Delta t}) \right\} \Delta t + \boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta \boldsymbol{W}_k \\ \overline{\boldsymbol{X}}_{0}^{\Delta t} &= \boldsymbol{x}_0 \end{aligned} \tag{2.27}$$

The parameter α can be interpreted as the degree of implicity. With $\alpha = 0.0$ the scheme (2.27) reduces to the explicit Euler scheme (2.23), while the choice $\alpha = 1.0$ leads to the implicit method (2.26).

2.3.3 Milstein scheme

The next numerical scheme was proposed by Milstein and has the strong order $\gamma = 1.0$ and weak order order $\beta = 1.0$. For simplicity, we consider only the case of diagonal noise

$$\begin{aligned} \overline{\boldsymbol{X}}_{k+1}^{\Delta t} &= \overline{\boldsymbol{X}}_{k}^{\Delta t} + \boldsymbol{a}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta t + \boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t})\Delta \boldsymbol{W}_k + \\ &\quad \frac{1}{2}\boldsymbol{\sigma}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t}) \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{x}}(t_k, \overline{\boldsymbol{X}}_{k}^{\Delta t}) \left\{ (\Delta \boldsymbol{W}_k)^2 - \Delta t \right\} \\ \overline{\boldsymbol{X}}_{0}^{\Delta t} &= \boldsymbol{x}_0 \end{aligned} \tag{2.28}$$

The additional term in Milstein scheme is connected with the corresponding term of the Taylor expansion of the Îto integral and marks the point of divergence of stochastic

numerical analysis from the deterministic. In this sense the Milstein scheme can be considered as a generalization of the deterministic Euler scheme, because it gives the same order of strong convergence as for the deterministic case [68].

2.3.4 Extrapolation method

Sometimes we can obtain higher order accuracy from an one-step scheme by the extrapolation method. We introduce now the extrapolation method for the simulation of functional of the stochastic process $\boldsymbol{X}$ based on Euler approximation assuming in what follows that the function g and its derivatives are smooth. First one should use Euler approximation $\overline{\boldsymbol{X}}^{\Delta t}$ generated by (2.23) with step size Δt and simulated the functional

$$U^{\Delta t} = E(g(\overline{\boldsymbol{X}}^{\Delta t}(T)))$$

then find the functional for the simulation with double time step

$$U^{2\Delta t} = E(g(\overline{\boldsymbol{X}}^{2\Delta t}(T)))$$

Finally, the two results are combined to yield the approximation

$$\overline{U}^{\Delta t} = 2U^{\Delta t} - U^{2\Delta t} \tag{2.29}$$

In this way we obtain from the first weak order $\chi = 1.0$ Euler scheme a method of weak order $\chi = 2.0$. This method is the extension of Romberg or Richardson method for ordinary differential equations. It was proposed by Talay and Tubaro [121].

2.4 Transition density function

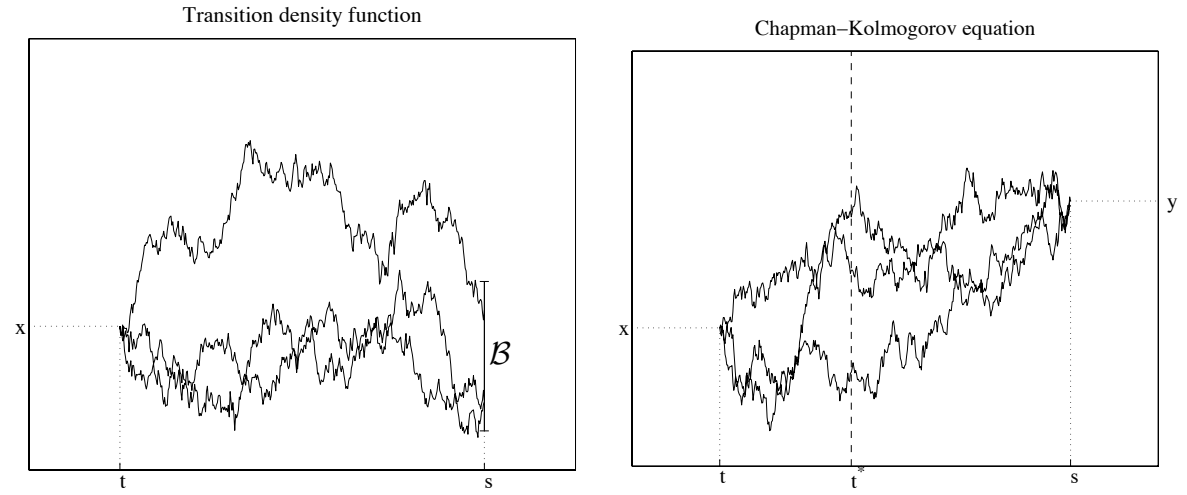
Let $\boldsymbol{X}(t)$ be the solution of (2.11). In this section we study some properties of the transition density function $p(t, \boldsymbol{x}; s, \boldsymbol{y})$, $t_0 \leq t \leq s \leq T$. This function is a function of four arguments and, roughly speaking, can be considered as a density function of random variable $\boldsymbol{X}(s)$ (for fixed s) under the condition $\boldsymbol{X}(t) = \boldsymbol{x}$. The formal definition can be found, for instance, in [2].

2.4.1 Properties of the transition density function

First of all, the transition density function should satisfy the Kolmogorov-Chapman equations, for any $t^* \in [t_0, T]$

$$p(t, \boldsymbol{x}; s, \boldsymbol{y}) = \int\limits_{\mathbb{R}^d} p(t, \boldsymbol{x}; t^*, \boldsymbol{u}) p(t^*, \boldsymbol{u}; s, \boldsymbol{y}) d\boldsymbol{u} \tag{2.30}$$

Now we will explore the intimate connection, which exists between stochastic differential equations and certain partial differential equations. The detailed discussion can be found in [13]. For simplicity we consider the one-dimensional case. Let us consider the following so called Cauchy problem.



(a) The transition probability $P(t, \boldsymbol{x}; s, \mathcal{B})$ (b) Chapman-Kolmogorov equation

Figure 2.4: Transition density function. (a) Suppose we want to find the probability $P(t, \boldsymbol{x}; s, \mathcal{B}) = P(\boldsymbol{X}(s) \in \mathcal{B} | \boldsymbol{X}(t) = \boldsymbol{x}) = \int_{\mathcal{B}} p(t, \boldsymbol{x}; s, \boldsymbol{y}) d\boldsymbol{y}$. *(b) The Chapman-Kolmogorov equation means, in imprecise language, that the probability of the transition from* $\boldsymbol{x}$ *at time* t *to* $\boldsymbol{y}$ *at time* s *is equal to the probability of the transition to* $\boldsymbol{u}$ *at an intermediate time* t^* *multiplied by the probability of the transition from* $\boldsymbol{u}$ *at the time* t^* *to* $\boldsymbol{y}$ *at the time* s*, summed over all intermediate values* $\boldsymbol{u}$

We are given three scalar functions $a(t, x)$, $\sigma(t, x)$ and $g(x)$. Our task is to find a function $U(t, x)$ which satisfies the following boundary value problem on $[t_0, s] \times \mathbb{R}$

$$\begin{aligned} &\frac{\partial U}{\partial t}(t, x) + a(t, x)\frac{\partial U}{\partial x}(t, x) + \frac{1}{2}\sigma^2(t, x)\frac{\partial U^2}{\partial x^2}(t, x) = 0 \\ &U(s, y) = g(y) \end{aligned} \tag{2.31}$$

Instead of attacking the boundary value problem (2.31) directly, we will produce a so called stochastic representation formula, which gives the solution of (2.31) in terms of the solution to a stochastic differential equation. We assume that there actually exists a solution $g(t, x)$ to (2.31). Let us now fix a time point t and a point in space x. Having fixed these we define the stochastic process X on time interval $[t, s]$ as the solution to the stochastic differential equation

$$\begin{aligned} &dX(\tau) = a(\tau, X(\tau))d\tau + \sigma(\tau, X(\tau))dW(\tau) \\ &X(t) = x \end{aligned} \tag{2.32}$$

Applying the Îto formula (2.10) to the stochastic process $g(\tau, X(\tau))$

$$\begin{aligned} &U(s, X(s)) = U(t, X(t)) + \\ &\int_t^s \left[\frac{\partial U}{\partial \tau}(\tau, X(\tau)) + a \frac{\partial U}{\partial x}(\tau, X(\tau)) + \frac{1}{2}\sigma^2(\tau, X(\tau)) \frac{\partial^2 U}{\partial x^2}(\tau, X(\tau)) \right] d\tau + \\ &\int_t^s \sigma(\tau, X(\tau)) \frac{\partial U}{\partial x}(\tau, X(\tau)) dW(\tau) \end{aligned} \tag{2.33}$$

The integral

$$\int_t^s \left[\frac{\partial U}{\partial \tau}(\tau, X(\tau)) + a \frac{\partial U}{\partial x}(\tau, X(\tau)) + \frac{1}{2}\sigma^2(\tau, X(\tau)) \frac{\partial^2 U}{\partial x^2}(\tau, X(\tau)) \right] d\tau$$

vanishes, because, by assumption the function g satisfies Equation (2.31). If, furthermore, the process $\sigma(\tau, X(\tau)) \frac{\partial U}{\partial x}(\tau, X(\tau))$ is sufficient integrable, then

$$E \int_t^s \sigma(\tau, X(\tau)) \frac{\partial U}{\partial x}(\tau, X(\tau)) dW(\tau) = 0$$

Taking into account the initial value $X(t) = x$ and the boundary condition $g(s, y) = \Phi(y)$ we obtain the formula

$$U(t, x) = E_{t,x}[g(X(T))] \tag{2.34}$$

where we have indexed the expectation operator in order to emphasize that the expected value is to be taken given the initial value $X(t) = x$. This result is known as the Feynman-Kac stochastic representation formula.

Let us consider the particular case when the function g is the indicator function $I_{\mathcal{B}}(y)$ of some Borel set $\mathcal{B}$. Then

$$U(t, x) = E_{t,x}[I_{\mathcal{B}}(X(s))] = P(X(s) \in \mathcal{B} | X(t) = x) = P(t, x; s, \mathcal{B})$$

This argument can also be turned around and we have thus (more or less) proved the following result

Proposition. (Kolmogorov backward equation)
Let X be a solution to Equation (2.32). Then the transition probabilities $P(t, x; s, \mathcal{B})$ are given as the solution to the equation

$$\begin{aligned} &\frac{\partial P}{\partial t}(t, x; s, \mathcal{B}) + a(t, x) \frac{\partial P}{\partial x}(t, x; s, \mathcal{B}) + \frac{1}{2}\sigma^2(t, x) \frac{\partial^2 P}{\partial x^2}(t, x; s, \mathcal{B}) = 0 \\ &P(s, x; s, \mathcal{B}) = I_{\mathcal{B}}(x) \end{aligned} \tag{2.35}$$

Using basically the same reasoning one can also prove the following corresponding result for transition densities.

Proposition. (Kolmogorov backward equation)

Let X be a solution to Equation (2.32) and suppose the function $P(t,x;s,\mathcal{B})$ has a density $p(t,x;s,y)$. Then we have

$$\begin{array}{l}\dfrac{\partial p}{\partial t}(t,x;s,y)+a(t,x)\dfrac{\partial p}{\partial x}(t,x;s,y)+\dfrac{1}{2}\sigma^2(t,x)\dfrac{\partial^2 p}{\partial x^2}(t,x;s,y)=0\\ p(s,x;s,y)=\delta(x-y)\end{array} \tag{2.36}$$

The reason that the equations (2.35) and (2.36) are called backward equations is that the differential operator is working on the "backward variables" (t,x). We will now derive the corresponding "forward" equation, where the action of the differential operator is on the 'forward variables' (s,y).

We assume that the stochastic process X has a transition density. Let us then fix a point in time t $(t<T)$. Now we consider an arbitrary "test function", i.e. an infinite differentiable function $h(s,y)$ with compact support in the set $(t,T)\times\mathbb{R}$. From the Îto rule (2.10) we have

$$h(T,X(T))=h(t,X(t))+\int_t^T\left[\frac{\partial h}{\partial s}+a\frac{\partial h}{\partial y}+\frac{1}{2}\sigma^2\frac{\partial^2 h}{\partial y^2}\right]ds+\int_t^T\sigma\frac{\partial h}{\partial y}dW(s) \tag{2.37}$$

Applying the expectation operator $E_{t,x}[\cdot]$ we obtain

$$\begin{array}{l}\int_{-\infty}^{+\infty}\int_t^T\dfrac{\partial h}{\partial s}(s,y)p(t,x;s,y)dsdy+\int_{-\infty}^{+\infty}\int_t^T a\dfrac{\partial h}{\partial y}p(t,x;s,y)dsdy+\\ \dfrac{1}{2}\int_{-\infty}^{+\infty}\int_t^T\sigma^2\dfrac{\partial^2 h}{\partial y^2}p(t,x;s,y)dsdy=0\end{array}$$

Partial integration with respect to t (we use the fact that, because of the compact support, $h(t,y)=h(T,y)=0$) gives us

$$\int_{-\infty}^{+\infty}\int_t^T\frac{\partial h}{\partial s}pdsdy=\int_{-\infty}^{+\infty}\left[hp\Big|_t^T-\int_t^T h\frac{\partial p}{\partial s}ds\right]dy=\int_{-\infty}^{+\infty}\int_t^T h\frac{\partial p}{\partial s}dsdy$$

Partial integration with respect to y and using the fact that the density function and its derivative vanish when $x\to\pm\infty$ gives us

$$\int_{-\infty}^{+\infty}\int_t^T a\frac{\partial h}{\partial y}pdsdy=\int_t^T\left[hap\Big|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}\frac{\partial(ap)}{\partial y}hdy\right]ds=-\int_{-\infty}^{+\infty}\int_t^T\frac{\partial(ap)}{\partial y}hdsdy$$

and

$$\begin{array}{l}\dfrac{1}{2}\int_{-\infty}^{+\infty}\int_t^T\sigma^2\dfrac{\partial^2 h}{\partial y^2}pdsdy=\dfrac{1}{2}\int_t^T\left[\dfrac{\partial h}{\partial y}\sigma^2 p\Big|_{-\infty}^{+\infty}-\dfrac{1}{2}\int_{-\infty}^{+\infty}\dfrac{\partial h}{\partial y}\dfrac{\partial(\sigma^2 p)}{\partial y}dy\right]ds=\\ \dfrac{1}{2}\int_t^T\left[-h\dfrac{\partial(\sigma^2 p)}{\partial y}\Big|_{-\infty}^{+\infty}+\dfrac{1}{2}\int_{-\infty}^{+\infty}h\dfrac{\partial^2(\sigma^2 p)}{\partial y^2}dy\right]ds=\dfrac{1}{2}\int_{-\infty}^{+\infty}\int_t^T h\dfrac{\partial^2(\sigma^2 p)}{\partial y^2}dsdy\end{array}$$

As a result we have

$$\int_{-\infty}^{+\infty}\int_{t}^{T} h(s,y)\left(-\frac{\partial p}{\partial s}-\frac{\partial}{\partial y}(ap)+\frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2 p)\right)dsdy=0$$

Since this equation holds for all test functions we have shown the following result.

Proposition. (Kolmogorov forward equation)
Assume that the solution X of (2.32) has the transition density function $p(t,x;s,y)$. Then p will satisfy the Kolmogorov forward equation

$$\begin{aligned}&\frac{\partial p}{\partial s}(t,x;s,y)+\frac{\partial}{\partial y}(a(s,y)p)-\frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(s,y)p)=0\\&p(t,x;t,y)=\delta(x-y)\end{aligned}\tag{2.38}$$

The decisive property of the stochastic process $\boldsymbol{X}(t)$ is that its transitions density function $p(t,\boldsymbol{x};s,\boldsymbol{y})$ is uniquely determined merely by the drift vector function $\boldsymbol{a}(t,\boldsymbol{x})$ and diffusion matrix $\boldsymbol{b}(t,\boldsymbol{x}) := \boldsymbol{\sigma}^T(t,\boldsymbol{x})\boldsymbol{\sigma}(t,\boldsymbol{x})$. The next theorem gives the d-dimensional version of Kolmogorov forward equation.

Theorem.
Suppose that the transition density function $p(t,\boldsymbol{x};s,\boldsymbol{y})$ is continuous with respect to t and s, and all first and second derivatives respect to $\boldsymbol{x}$ and $\boldsymbol{y}$ exist and are continuous. Then $p(t,\boldsymbol{x};s,\boldsymbol{y})$ is the solution of the Kolmogorov forward (Fokker-Planck) equation ($t_o \le t \le s \le T$)

$$\begin{aligned}&\frac{\partial p}{\partial s}+\sum_{i=1}^{d}\frac{\partial}{y_i}(a_i(s,\boldsymbol{y})p)-\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial y_i\partial y_j}(b_{ij}(s,\boldsymbol{y})p)=0\\&p(t,\boldsymbol{x};t,\boldsymbol{y})=\delta(\boldsymbol{x}-\boldsymbol{y})\end{aligned}\tag{2.39}$$

and Kolmogorov backward equation

$$\begin{aligned}&\frac{\partial p}{\partial t}+\sum_{i=1}^{d}a_i(t,\boldsymbol{x})\frac{\partial p}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^{d}b_{ij}(t,\boldsymbol{x})\frac{\partial^2 p}{\partial x_i\partial x_j}=0\\&p(s,\boldsymbol{x};s,\boldsymbol{y})=\delta(\boldsymbol{x}-\boldsymbol{y})\end{aligned}\tag{2.40}$$

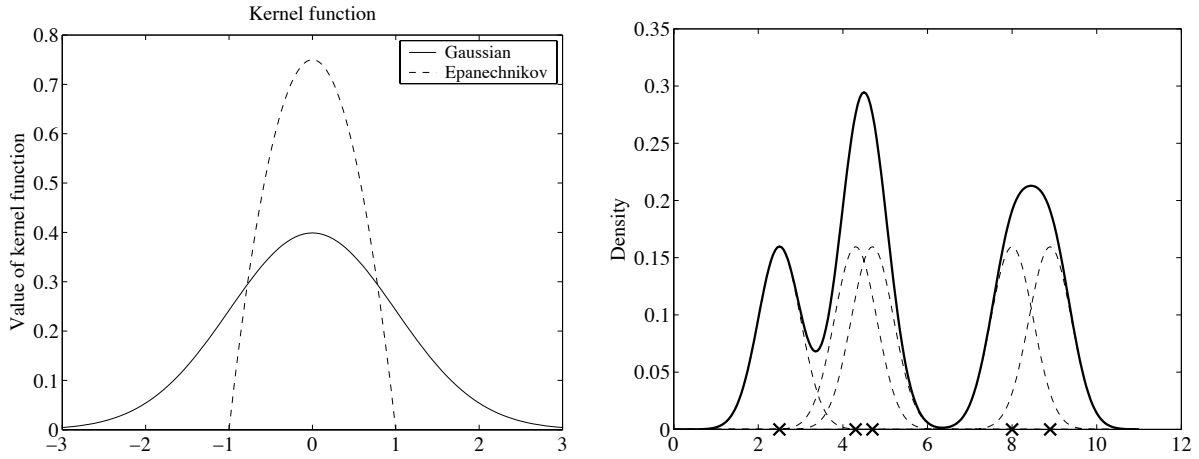
The proof of this theorem is given in [46].

Using the same arguments it is easy to show that the backward Kolmogorov equation (2.40) is also valid for the functional

$$U(t,\boldsymbol{x})=E_{t,\boldsymbol{x}}g(\boldsymbol{X}(T))\tag{2.41}$$

Theorem.
Suppose that $g(\boldsymbol{x})$ is a continuous bounded function with continuous bounded first and second partial derivatives, then the function $U(t,\boldsymbol{x})$ defined by (2.41) satisfies the Kolmogorov backward equation

$$\begin{aligned}&\frac{\partial U}{\partial t}+\sum_{i=1}^{d}a_i(t,\boldsymbol{x})\frac{\partial U}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^{d}b_{ij}(t,\boldsymbol{x})\frac{\partial^2 U}{\partial x_i\partial x_j}=0\\&U(s,\boldsymbol{x})=g(\boldsymbol{x})\end{aligned}\tag{2.42}$$



(a) Kernel functions

(b) Kernel estimator based on five observations

Figure 2.5: The kernel estimator. (a) Usually a kernel function is chosen to be a probability density function that is symmetric about zero, for instance Gaussian or Epanechnikov functions. This ensures that $\widehat{p}(\boldsymbol{x})$ is itself also a density. (b) This figure shows a kernel estimator constructed using five observations with the Gaussian kernel function.

2.4.2 The kernel estimator for the density function

In order to find the density function of he random variable $\boldsymbol{X}$ we can use on of the standard method of non-parametric statistics called kernel estimator

$$\widehat{p}(\boldsymbol{x}) = \frac{1}{N\delta^d} \sum_{n=1}^{N} K\left(\frac{\boldsymbol{x} - \boldsymbol{X}^{(n)}}{\lambda}\right) \tag{2.43}$$

Here $\boldsymbol{X}^{(n)}$, $n = 1, \ldots, N$ is the sampling from $\boldsymbol{X}$, K is kernel function and λ is a positive number, usually called bandwidth. One can think of the kernel estimator as spreading of a "probability mass" of size $1/N$ associated with each data point about its neighborhood. Combining contributions from each data point means that in regions where there are many observations the density has a relatively large value and opposite in regions with only few observations.

Usually kernel functions are chosen to satisfy the conditions

1. $\int K(\boldsymbol{x}) d\boldsymbol{x} = 1$
2. $K(\boldsymbol{x})$ is a symmetric function ($K(\boldsymbol{x}) = K(-\boldsymbol{x})$)
3. $\int \boldsymbol{x}^2 K(\boldsymbol{x}) d\boldsymbol{x} = a^2 < \infty$

4. $K(\boldsymbol{x}) \geq 0 \quad \text{for all} \quad \boldsymbol{x}$

However, the kernel can be any function with the property $\int K(\boldsymbol{x})d\boldsymbol{x} = 1$. The examples of very often used kernel functions are Gaussian function

$$K(\boldsymbol{x}) = (2\pi)^{-d/2} exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \tag{2.44}$$

and Epanechnikov symmetric multivariate kernel

$$K(\boldsymbol{x}) = \frac{1}{2}\nu_d^{-1}(d+2)(1-\boldsymbol{x}^T\boldsymbol{x})1_{\boldsymbol{x}^T\boldsymbol{x}\leq 1} \tag{2.45}$$

where $\nu_d = 2\pi^{d/2}/\{d\Gamma(d/2)\}$ is the volume of the unit d-dimensional sphere and $\Gamma(x)$ is a Gamma function (see Figure 2.5).

The kernel and, especially, bandwidth should be chosen in the way to minimize the error of the estimator

$$\epsilon_{\text{kernel}}(p) := \sqrt{E(\widehat{p}-p)^2} = \sqrt{E(\widehat{p}-E\widehat{p})^2 + (E\widehat{p}-p)^2} = \sqrt{Var(\widehat{p}) + Bias^2(\widehat{p})} \tag{2.46}$$

It is well known [106, 144] that the optimal bandwidth is given by

$$\lambda_{opt} \sim N^{-\frac{1}{d+4}}$$

providing the error (2.46) of the following order

$$\epsilon(p) \sim N^{-\frac{2}{d+4}} \tag{2.47}$$

From this formula it is seen that even the optimal choice of the bandwidth leads to quite poor estimation properties. The error of the estimation $\epsilon(p)$ converges to zero as N increases, but does so extremely slow for large values of d. Hence, reasonable results for high-dimensional systems require a huge size of the sample. In statistical literature this problem is referred as the "curse of dimensionality".

The situation can be a little bit improved by choosing the kernel function and the bandwidth in a more sophisticated way. In the general form, the kernel estimator (2.43) can be written as

$$\widehat{p}(\boldsymbol{x}) = \frac{1}{N}\sum_{n=1}^{N} K_{\boldsymbol{\mathcal{H}}}(\boldsymbol{x} - \boldsymbol{X}^{(n)}) \tag{2.48}$$

where $\boldsymbol{\mathcal{H}}$ is a symmetric positive definite $d \times d$ matrix called the bandwidth matrix, and

$$K_{\boldsymbol{\mathcal{H}}}(\boldsymbol{x}) = \boldsymbol{\mathcal{H}}^{-\frac{1}{2}} K(\boldsymbol{\mathcal{H}}^{\frac{1}{2}}\boldsymbol{x}) \tag{2.49}$$

There are several level of complexification when specifying the bandwidth matrix $\boldsymbol{\mathcal{H}}$. The simplest corresponds to the restriction $\boldsymbol{\mathcal{H}} = \lambda^2\boldsymbol{I}_d$ for some $\lambda > 0$. At the next level $\boldsymbol{\mathcal{H}} = diag(\lambda_1^2, \ldots, \lambda_d^2)$ is a diagonal matrix. So, by introducing $d-1$ additional parameters, one has the flexibility to smooth by different amounts of d coordinate directions. If we wish to smooth in directions different from coordinates, the full bandwidth matrix will be appreciate. In particular, the bandwidth matrix can be taken in the form

$$\boldsymbol{\mathcal{H}} = \lambda^2\boldsymbol{S} \tag{2.50}$$

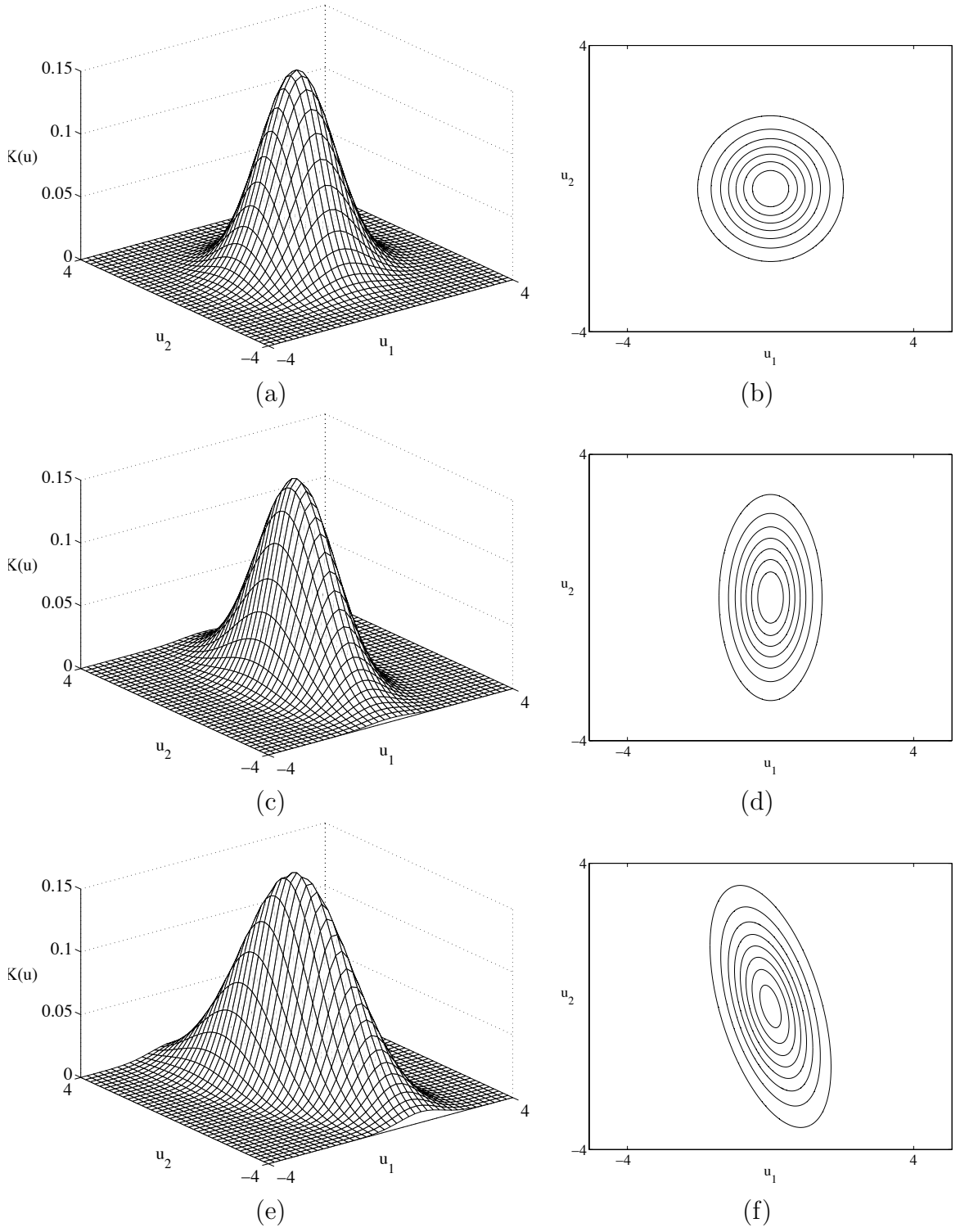


Figure 2.6: The Gaussian kernel function and contour plots parameterized by (a), (b) $\boldsymbol{\mathcal{H}} = \delta^2 \boldsymbol{I}_2$*, (c), (d)* $\boldsymbol{\mathcal{H}} = diag(\delta_1^2, \delta_2^2)$*, (e), (f)* $\boldsymbol{\mathcal{H}} = \delta^2 \boldsymbol{S}$

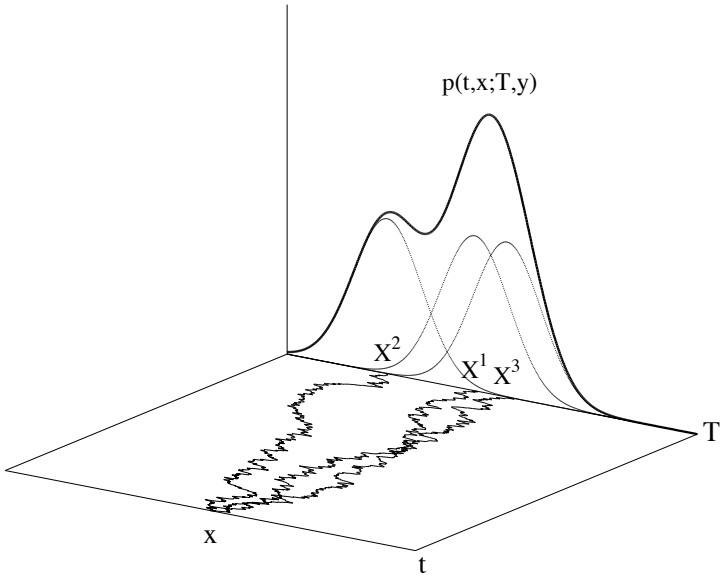


Figure 2.7: In order to estimate the transition density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ we start the simulation of the stochastic process $\boldsymbol{X}$ at the time t from the location $\boldsymbol{x}$ many times. The different realizations of the stochastic process $\overline{\boldsymbol{x}}$ at time T allow to estimate the function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ using a kernel estimator

where $\boldsymbol{S}$ is a sample covariance matrix [44].

Combining the kernel estimator (2.43) and the Euler scheme yields to the estimator for the transition density function

$$\widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \frac{1}{N\delta^d} \sum_{n=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{x}}(T) - \boldsymbol{y}}{\lambda} \right) \tag{2.51}$$

or in another form

$$\widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \frac{|\boldsymbol{\mathcal{H}}|^{-\frac{1}{2}}}{N} \sum_{n=1}^{N} K\left(\boldsymbol{\mathcal{H}}^{-1/2} (\overline{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{x}}(T) - \boldsymbol{y}) \right) \tag{2.52}$$

The estimator $\widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y})$can be written in shorter form

$$\widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \frac{1}{N} \sum_{n=1}^{N} K_{\boldsymbol{\mathcal{H}}} (\overline{\boldsymbol{X}}^{(n)}_{t,x}(T) - \boldsymbol{y}) \tag{2.53}$$

2.4.3 The kernel estimator near boundaries

Throughout of the previous section we supposed that the probability density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ satisfies certain smoothing criteria. In practice, however, most densities are

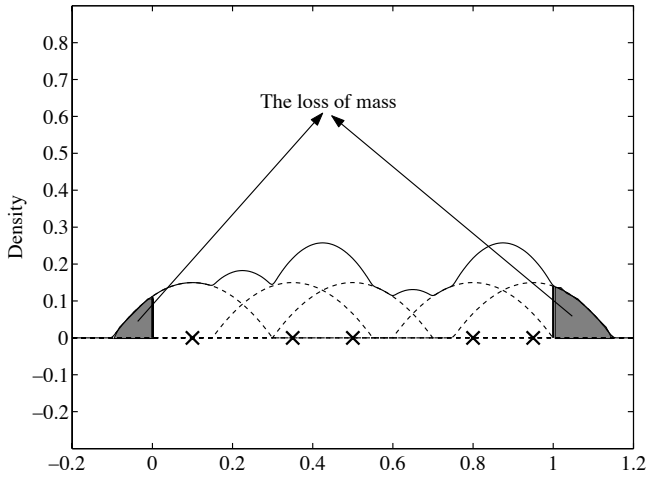


Figure 2.8: The loss of mass near boundaries due to kernel estimators

discontinuous. In this section we study the ways of kernel estimator corrections near boundary. For the sake of simplicity, we consider an one-dimensional case.

Let us assume that $p(x)$ is a density function of the one-dimensional random variable X. If the density $p(\boldsymbol{x})$ is a twice continuously differentiable function, the kernel estimator (2.43) is consistent with the true value of the density function $p(\boldsymbol{x})$. Indeed,

$$E\widehat{p}(x) = \frac{1}{N\lambda}\sum_{n=1}^{N} K\left(\frac{X^{(n)}-x}{\lambda}\right) = \frac{1}{\lambda}EK\left(\frac{X-x}{\lambda}\right) = \frac{1}{\lambda}\int K\left(\frac{y-x}{\lambda}\right)p(y)dy = \int K(\boldsymbol{z})p(x-\lambda z)dz$$

Expanding $p(x-\lambda z)$ in a Taylor series about x we obtain

$$p(x-\lambda z) = p(x) - \lambda z p'(x) + \frac{1}{2}\lambda^2 z^2 p''(x) + \mathcal{O}(\lambda^3)$$

uniformly in z. This leads to

$$E\hat{p}(x) = p(x) + \frac{1}{2}\lambda^2 p''(x)\int z^2 K(z)dz + o(\lambda^2) = p(x) + \mathcal{O}(\lambda^2) \tag{2.54}$$

Here the properties of the kernel function have been used.

Now, we suppose that density $p(x)$ is a function such that $p(t,x) = 0$, $x \notin [0,1]$. Using the kernel estimator (2.43) will lead to the loss of total mass. In this case the kernel estimator overspills the boundaries as shown in Figure 2.8. Moreover, the estimator (2.43) is not consistent with the true value of the density function at and near boundaries and the equation (2.54) is not valid anymore.

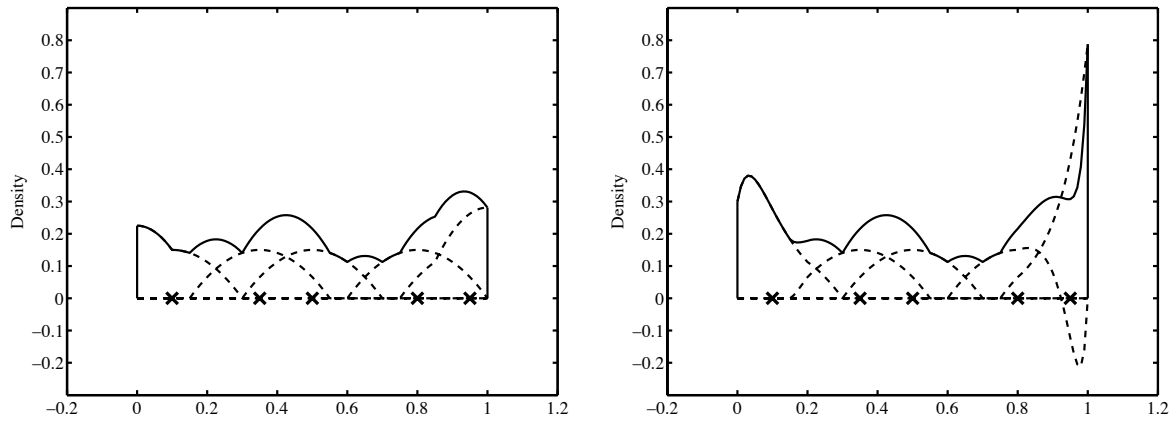


Figure 2.9: The transformation of kernel (a) using a reflection kernel (b) using a boundary kernel

Indeed, if $x = \alpha\lambda$, $0 < \alpha < 1$

$$E\widehat{p}(x) = \frac{1}{\lambda}\int\limits_{y\in\mathbb{R}} K\left(\frac{y-x}{\lambda}\right)p(y)dy = \frac{1}{\lambda}\int\limits_{0}^{1} K\left(\frac{y-x}{\lambda}\right)p(y)dy = \frac{1}{\lambda}\int\limits_{0}^{1} K\left(\frac{x-y}{\lambda}\right)p(y)dy = \int\limits_{-1/\lambda+\alpha}^{\alpha} K(z)p(x-\lambda z)dz$$

Using a Taylor expansion we can write the following expression

$$E\widehat{p}(x) = p(x)\nu_{0,\alpha} - \lambda p'(x)\nu_{1,\alpha} + \frac{\lambda^2}{2}p''(x)\nu_{2,\alpha} + \mathcal{O}(\lambda^3)$$

where $\nu_{i,\alpha} = \int\limits_{-1/\lambda+\alpha}^{\alpha} K(z)z^i dz$. If $\alpha = 0$ than $\nu_{0,0} = \int\limits_{-1/\lambda}^{0} K(z)dz \leq \frac{1}{2}$ ($K(z)$ is a symmetric function) and

$$E\widehat{p}(x) \leq \frac{1}{2}p(x) + \mathcal{O}(\lambda)$$

and the kernel estimator is not consistent with the true value of the probability function at $x = 0$.

Since the locations of the boundaries are usually known the kernel estimator may be adapted to achieve better performance near the boundary. One of the most natural and simplest way is to reinstate the "missing mass" by reflecting the estimate in the boundary (e.g. [106]). The kernel function is reflected with regard to the boundary as it is shown on Figure 2.9. In this way we can achieve consistency of the estimator, however, this method still results in a large bias. Let us assume that the kernel function $K(u)$ has a compact

support on $[-1,1]$ (for instance, Epanechnikov function), than (K_r is the reflection kernel)

$$E\widehat{p}(x) = \frac{1}{\lambda}\int_0^1 K_r\left(\frac{y-x}{\lambda}\right)p(y)dy =$$
$$\frac{1}{\lambda}\int_0^\lambda K_r\left(\frac{y-x}{\lambda}\right)p(y)dy + \frac{1}{\lambda}\int_\lambda^{1-\lambda} K_r\left(\frac{y-x}{\lambda}\right)p(y)dy + \frac{1}{\lambda}\int_{1-\lambda}^1 K_r\left(\frac{y-x}{\lambda}\right)p(y)dy =$$
$$\frac{1}{\lambda}\int_0^\lambda \left[K\left(\frac{y-x}{\lambda}\right) + K\left(\frac{-y-x}{\lambda}\right)\right]p(y)dy + \frac{1}{\lambda}\int_\lambda^{1-\lambda} K\left(\frac{y-x}{\lambda}\right)p(y)dy +$$
$$\frac{1}{\lambda}\int_{1-\lambda}^1 \left[K\left(\frac{y-x}{\lambda}\right) + K\left(\frac{2-y-x}{\lambda}\right)\right]p(y)dy =$$
$$\frac{1}{\lambda}\int_0^1 K\left(\frac{y-x}{\lambda}\right)p(y)dy + \frac{1}{\lambda}\int_0^\lambda K\left(\frac{-y-x}{\lambda}\right)p(y)dy + \frac{1}{\lambda}\int_{1-\lambda}^1 K\left(\frac{2-y-x}{\lambda}\right)p(y)dy =$$
$$\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3$$

$$\mathcal{I}_1 = \int_0^1 K\left(\frac{y-x}{\lambda}\right)p(y)dy = \int_0^1 K\left(\frac{x-y}{\lambda}\right)p(y)dy = \int_{-1/\lambda+x/\lambda}^{x/\lambda} K(z)p(x-\lambda z)dz$$
$$\mathcal{I}_2 = \int_0^\lambda K\left(\frac{-y-x}{\lambda}\right)p(y)dy = \int_0^\lambda K\left(\frac{x+y}{\lambda}\right)p(y)dy = \int_{x/\lambda}^{x/\lambda+1} K(z)p(x-(2x-\lambda z))dz$$
$$\mathcal{I}_3 = \int_{1-\lambda}^1 K\left(\frac{2-y-x}{\lambda}\right)p(y)dy = \int_{-1-1/\lambda+x/\lambda}^{-1+x/\lambda} K(z)p(2-x+\lambda z)dz$$

Let us consider three cases

(I): $x = \alpha\lambda, \qquad 0 < \alpha < 1$

(II): $x = 1-\alpha\lambda, \qquad 0 < \alpha < 1$

(III): $\lambda \le x \le 1-\lambda,$

We suppose that $\lambda < 1/2$. Let us start with case (III).

(III):

$$\mathcal{I}_1 = \int_{-1/\lambda+x/\lambda}^{x/\lambda} K(z)p(x-\lambda z)dz = \int_{-1}^1 K(z)p(x-\lambda z)dz$$
$$\mathcal{I}_2 = \mathcal{I}_3 = 0$$
$$\mathcal{I} = p(x)\int_{-1}^1 K(z)dz - \lambda p'(x)\int_{-1}^1 zK(z)dz + \frac{\lambda^2 p''(x)}{2}\int_{-1}^1 z^2K(z)dz + \mathcal{O}(\lambda^3) = p(x) + \mathcal{O}(\lambda^2)$$
$$E\widehat{p}(x) = p(x) + \mathcal{O}(\lambda^2)$$

(I):

$$\mathcal{I}_1 = \int_{-1/\lambda+x/\lambda}^{x/\lambda} K(z)p(x-\lambda z)dz = \int_{-1}^{\alpha} K(z)p(x-\lambda z)dz =$$
$$p(x)\int_{-1}^{\alpha} K(z)dz - \lambda p'(x)\int_{-1}^{\alpha} zK(z)dz + \frac{\lambda^2 p''(x)}{2}\int_{-1}^{\alpha} z^2K(z)dz + \mathcal{O}(\lambda^3)$$
$$\mathcal{I}_2 = \int_{x/\lambda}^{x/\lambda+1} K(z)p(x-(2x-\lambda z))dz = \int_{\alpha}^1 K(z)dz = p(x-(2x-\lambda z))dz =$$
$$p(x)\int_{\alpha}^1 K(z)dz - p'(x)\int_{\alpha}^1 (2x-\lambda z)K(z)dz + \frac{\lambda^2 p''(x)}{2}\int_{\alpha}^1 (2\alpha - z)^2K(z)dz + \mathcal{O}(\lambda^3)$$

$$\mathcal{I}_3 = 0$$

$$\begin{aligned} E\widehat{p}(x) = \quad & \mathcal{I}_1 + \mathcal{I}_2 = \\ & = p(x) + p'(x)\left(\lambda \int_{\alpha}^{\sigma} zK(z)dz - \lambda \int_{-\sigma}^{\alpha} zK(z)dz - 2x \int_{\alpha}^{\sigma} K(z)dz\right) + \mathcal{O}(\lambda^2) = \\ & = p(x) - 2p'(x)\lambda\left(\alpha \int_{\alpha}^{\sigma} K(z)dz - \int_{\alpha}^{\sigma} zK(z)dz\right) + \mathcal{O}(\lambda^2) \end{aligned}$$

This formula may be found, for instance, in [63]. Case (II) is analogous to case (I).

A variety of further modifications is possible to achieve smaller bias. One can think of these boundary modifications in terms of special "boundary kernels" that are different for every α. The example of the boundary kernel based on Epanechnikov function is

$$\begin{aligned} K(x) = \quad & 6(1+x)(\alpha - x)\tfrac{1}{(1+\alpha)^3} \\ & \left(1 + 5\left(\tfrac{1-\alpha}{1+\alpha}\right)^2 + 10\tfrac{1-\alpha}{(1+\alpha)^2}x\right) \end{aligned} \tag{2.55}$$

This boundary kernel was proposed by [88]. The disadvantage of the boundary kernels is that they do not guarantee the positiveness of the numerical solution (see Figure 2.9)(b). Besides the described methods, there are many techniques of the kernel transformation in the boundary. An overview of boundary kernels can be obtained from [63].

2.5 The error of the kernel estimator

In this chapter we analyze the error of the kernel estimator (2.51)

$$\epsilon_{\text{total}}(\Delta t, N) = p(t, \boldsymbol{x}; T, \boldsymbol{y}) - \widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) \tag{2.56}$$

We can decompose this error into a systematic error ϵ_{sys} and a statistical error ϵ_{stat}, with

$$\epsilon_{\text{total}}(\Delta t, N) = E\epsilon_{\text{total}} + (\epsilon_{\text{total}} - E\epsilon_{\text{total}}) = \epsilon_{\text{sys}}(\Delta t, N) + \epsilon_{\text{stat}}(\Delta t, N)$$

where

$$\epsilon_{\text{sys}} = E(\epsilon_{\text{total}}) = E\left(p(t, \boldsymbol{x}; T, \boldsymbol{y}) - \widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y})\right) = Ep(t, \boldsymbol{x}; T, \boldsymbol{y}) - E\widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) \tag{2.57}$$

and

$$\begin{aligned} \epsilon_{\text{stat}} = \quad & \epsilon_{\text{total}} - E\epsilon_{\text{total}} p(t, \boldsymbol{x}; T, \boldsymbol{y}) - \widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) - Ep(t, \boldsymbol{x}; T, \boldsymbol{y}) + E\widehat{p}_{\Delta t}(t, \boldsymbol{x}; T, \boldsymbol{y}) \\ & = E\frac{1}{N}\sum_{n=1}^{N} K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T) - \boldsymbol{y}) - \frac{1}{N}\sum_{n=1}^{N} K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T) - \boldsymbol{y}) = \\ & = \frac{1}{N}\sum_{n=1}^{N}\left(EK_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T) - \boldsymbol{y}) - K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T) - \boldsymbol{y})\right) \end{aligned} \tag{2.58}$$

For a large number N of independent simulation we conclude from the Central Limit Theorem that the statistical error ϵ_{stat} becomes asymptotically Gaussian with mean zero

and variance

$$\begin{aligned} Var(\epsilon_{\text{stat}}) &= E\left(\frac{1}{N}\sum_{n=1}^{N}\left(EK_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T)-\boldsymbol{y})-K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T)-\boldsymbol{y})\right)^2\right) \\ \frac{1}{N^2}\sum_{n=1}^{N} E\left(K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T)-\boldsymbol{y})-EK_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}^{(n)}(T)-\boldsymbol{y})\right)^2 &= \frac{1}{N}Var(K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}(T))) \end{aligned} \tag{2.59}$$

In the next two section we discuss the statistical and systematic errors in detail.

2.5.1 The variance reduction method

From Equation (2.59) it is clear that the statistical error depends on the variance of the random variable $K_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{X}}(T))$ which is close to the variance of $K_{\boldsymbol{\mathcal{H}}}(\boldsymbol{X}(T))$. This variance which depends totally on the choice of kernel function $K_{\boldsymbol{\mathcal{H}}}$ and the solution $\boldsymbol{X}(T)$ of the stochastic differential equation (2.11) may be extremely large. This problem leads to the question of whether it is possible to construct other estimators which have nearly the same expectation, but smaller variance. Now let us consider along with the stochastic differential equation (2.11) the system of stochastic differential equations for the process $(\tilde{\boldsymbol{X}}, \Theta(t))$ given by

$$\begin{aligned} d\tilde{\boldsymbol{X}}(t) &= (a(t,\tilde{\boldsymbol{X}}(t)) - \boldsymbol{\sigma}(t,\tilde{\boldsymbol{X}}(t))\boldsymbol{d}(t,\tilde{\boldsymbol{X}}(t))dt + \boldsymbol{\sigma}(t,\tilde{\boldsymbol{X}}(t))d\boldsymbol{W}(t) \\ d\Theta(t) &= \sum_{j=1}^{m} d_j(t,\tilde{\boldsymbol{X}}(t))\Theta(t)dW_j(t) \\ \tilde{\boldsymbol{X}}(t_0)) &= \tilde{\boldsymbol{x}}_0, \quad \Theta(t_0) = \Theta_0 \end{aligned} \tag{2.60}$$

Here $\boldsymbol{d} = (d_1, \ldots, d_m)^T \in \mathbb{R}^m$ and $\Theta_0 \neq 0$. Note that $\tilde{\boldsymbol{X}}(t)$ is d-dimensional, whereas the correlation process $\Theta(t)$ is only one-dimensional.

By a measure transformation, the Girsanov transformation [68, 91], it follows that the following expectations are equal

$$E(K_H(\boldsymbol{X}(T))) = E(K_{\boldsymbol{\mathcal{H}}}(\tilde{\boldsymbol{X}}(T))\Theta(T)/\Theta(t_0)) \tag{2.61}$$

Hence, instead to estimate the value of $E(K_{\boldsymbol{\mathcal{H}}}(\boldsymbol{X}(T)))$ using Monte-Carlo method we can estimate the expectation of the random variable

$$E(K_{\boldsymbol{\mathcal{H}}}(\tilde{\boldsymbol{X}}(T))\Theta(T)/\Theta(t_0)) \tag{2.62}$$

As this result does not depend on the choice of the functions d_j, $j = 1, \ldots m$, we can use them as adjustable parameters to reduce the variance of the random variable (2.62).

Let choose the parameter function d_i as

$$d_k(t,\boldsymbol{x}) = -\frac{1}{U(t,\boldsymbol{x})}\sum_{i=1}^{d}\sigma_{ik}(t,\boldsymbol{x})\frac{\partial U}{\partial x_i}(t,\boldsymbol{x}), \tag{2.63}$$

where

$$U(t,\boldsymbol{x}) = EK_{\boldsymbol{\mathcal{H}}}(X_{t,\boldsymbol{x}}(T))$$

Let us consider the stochastic process

$$U(t,\tilde{\boldsymbol{X}}(t,\boldsymbol{x})\Theta(t)$$

Using the Îto rule (2.13) and the backward Kolmogorov equation (2.42) we can write

$$d(U(t,\tilde{\boldsymbol{X}}(T))\Theta(t)) = \Theta(t)dU(t,\tilde{\boldsymbol{X}}(T)) + U(t,\tilde{\boldsymbol{X}}(T))d\Theta(t) =$$
$$\sum_{i=1}^{d}\sum_{k=1}^{m}\Theta(t)\frac{\partial u}{\partial x_i}\sigma_{ik}dW_k(t) + U(t,\tilde{\boldsymbol{X}}(T))\sum_{k=1}^{m}d_k\Theta(t)dW_k(t) =$$
$$\Theta(t)\sum_{k=1}^{m}\left(\sum_{i=1}^{d}\frac{\partial U(t,\tilde{\boldsymbol{X}})}{\partial x_i}\sigma_{ik} + U(t,\tilde{\boldsymbol{X}})d_k\right)dW_k(t)$$

Taking into account the choice of the parameter functions d_k, $k = 1,\dots,m$ we can conclude that

$$U(t,\tilde{\boldsymbol{X}}(T))\Theta(t) = U(t_0,\boldsymbol{x}_0)\Theta_0$$

is nonrandom, so its variance is zero. Unfortunately, we can not construct the functions d_k as proposed in (2.63), because we need to know the function U, but this what we are trying to determine by this simulation.

It is often possible to find or guess a function $\tilde{u}$ which is similar to u and can be used instead of u in (2.63) to define the parameter functions

$$d_k(t,\boldsymbol{x}) = -\frac{1}{\tilde{u}(t,\boldsymbol{x})}\sum_{i=1}^{d}\sigma_{ik}(t,\boldsymbol{x})\frac{\partial\tilde{u}}{\partial x_i}(t,\boldsymbol{x})$$

for all $(t,\boldsymbol{x}) \in [t_0,T]\times\mathbb{R}^d$ and $k = 1,\dots m$. Then the expression

$$K_H(\tilde{\boldsymbol{X}}(t))\Theta(T)/\Theta_0$$

will still be random, but with small variance if $\tilde{u}$ is chosen sufficiently closed to u.

In this section we described only the main idea of the variance reduction method. Different Monte-Carlo variance reduction techniques can be found in [68, 84, 99, 100, 143]. This method is often used in different applications: in stochastic environmental models [24, 72, 81, 103, 117], in finance applications [47, 51]

2.5.2 The analysis of the systematic error

The estimation lost of the kernel estimator (2.51) can be split into two parts: the error ϵ_{app} due to the approximation $\overline{\boldsymbol{X}}(T)$ of the stochastic process $\boldsymbol{X}(T)$ (see Section 2.3.1) and the error ϵ_{kernel} due to the kernel estimator (see Section 2.4.2). In other words, the systematic error of the estimator (2.51) can be written in the form

$$\begin{aligned}&|\epsilon_{\text{sys}}(\Delta t,N)| := |Ep - E\widehat{p}_{\Delta t}| \le |Ep - E\widehat{p}| + |E\widehat{p} - E\widehat{p}_{\Delta t}| \le \epsilon_{\text{kernel}} + \epsilon_{\text{app}}\\ &= \mathcal{O}\left(N^{-\frac{2}{4+d}}\right) + \mathcal{O}(\Delta t^{\chi}) = \left(\alpha_{\text{a}}\Delta t^{\chi} + \alpha_{\text{k}}N^{-\frac{2}{4+d}}\right)(1+o(1))\end{aligned} \tag{2.64}$$

Here we use the following inequality

$$|Ep - E\widehat{p}| \le E|p-\widehat{p}| \le \sqrt{E(p-\widehat{p})^2}$$

When we choose the parameters Δt and N our main goal is to minimize the systematic error lost and from the equation (2.64) it is clear that for any fixed real number ϵ_0 we can

find the parameters $(\Delta t_0, N_0)$ such that $\epsilon_{\text{sys}}(\Delta t_0, N_0) \leq \epsilon_0$. However it can dramatically increase the computational effort of the estimator (2.64). Suppose that the computational effort for the simulation of one realization of the stochastic process $\boldsymbol{X}(T)$ is given by

$$\mathcal{E}(\Delta t) = \mathcal{O}\left(\frac{1}{\Delta t^q}\right) = \frac{\mathcal{K}}{\Delta t^q}(1 + o(1))$$

Then the total computational lost of the estimator (2.64) is given by

$$\mathcal{E}_{\text{total}}(\Delta t, N) := \frac{N\mathcal{K}}{\Delta t^q}(1 + o(1))$$

As a result we can face the following two optimization problems

$$\mathcal{E}^*(\epsilon) = \min_{\{\Delta t, N\}} \{\mathcal{E}_{\text{total}}(\Delta t, N) || \epsilon_{\text{sys}}(\Delta t, N)| \leq \epsilon\} \tag{2.65}$$

and

$$\epsilon^*(\mathcal{E}) = \min_{\{\Delta t, N\}} \{|\epsilon_{\text{sys}}(\Delta t, N)|| \mathcal{E}_{\text{total}}(\Delta t, N) \leq \mathcal{E}\} \tag{2.66}$$

The discussion about these optimization problems and some practical advises can be found in [102]. Let us consider, for example, the optimization problem (2.65). For some fixed value ϵ_0 we have

$$|\epsilon_{\text{sys}}| = \left(\alpha_{\text{a}}\Delta t^{\chi} + \alpha_{\text{k}} N^{-\frac{2}{4+d}}\right)(1 + o(1)) \leq \epsilon_0 \Longrightarrow$$

$$\epsilon_{\text{app}} = \alpha_{\text{a}}\Delta t^{\chi}(1 + o(1)) \leq \omega\epsilon_0 \quad \epsilon_{\text{kernel}} = \alpha_{\text{k}} N^{-\frac{2}{4+d}}(1 + o(1)) \leq (1 - \omega)\epsilon_0$$

for some fixed $0 < \omega < 1$. From this it follows that

$$\Delta t = \left(\frac{\omega\epsilon_0}{\alpha_{\text{a}}}\right)^{\frac{1}{\chi}}(1 + o(1)) \quad \text{and} \quad N = \left(\frac{\alpha_{\text{k}}}{(1-\omega)\epsilon_0}\right)^{2+\frac{d}{2}}(1 + o(1)) \tag{2.67}$$

Now we choose the value of the parameter ω in order to to find $\mathcal{E}^*(\epsilon)$

$$\mathcal{E}_{\text{total}} = \frac{N\mathcal{K}}{\Delta t^q}(1 + o(1)) = \frac{\alpha_{\text{k}}^{2+\frac{d}{2}}\alpha_{\text{a}}^{\frac{q}{\chi}}}{(1-\omega)^{2+\frac{d}{2}}\omega^{\frac{q}{\chi}}\epsilon_0^{\frac{q}{\chi}}}(1 + o(1))$$

This function reaches the minimum when the function $f(\omega) = (1-\omega)^{2+\frac{d}{2}}\omega^{\frac{q}{\chi}}$ reaches the maximum. Using the standard method it is easy to show that the function $f(\omega)$ reaches the maximum at the point $\omega = \dfrac{q/\chi}{2 + d/2 + q/\chi}$.

Summarizing, we can say that for a given computational error ϵ_0 the minimum computational effort can be achieved by choosing the following parameters

$$\Delta t = \epsilon_0^{\frac{1}{\chi}}\left(\frac{q/\chi}{2 + d/2 + q/\chi}\right)^{\frac{1}{\chi}} \frac{1}{\alpha_{\text{a}}^{\frac{1}{\chi}}}(1 + o(1)) \tag{2.68}$$

and

$$N = \left(\frac{\chi}{\epsilon_0}\right)^{2+\frac{d}{2}}\left(\frac{q/\chi}{2 + d/2 + q/\chi}\right)^{2+\frac{d}{2}}(1 + o(1)) \tag{2.69}$$

For practical applications the results (2.68) and (2.69) are best interpreted in the following way. Let $\mathcal{K}_0$ and N_0 be positive constants. Then, by choosing

$$\Delta t(\epsilon) = \mathcal{K}_0 \epsilon^{\frac{1}{\chi}} (1 + o(1)) \quad \text{and} \quad N(\epsilon) = \frac{N_0}{\epsilon^{2+\frac{d}{2}}} (1 + o(1))$$

we can guarantee that the estimation error has the order ϵ and the computational effort $\mathcal{E}(\epsilon)$ has the order

$$\mathcal{E}_{\text{total}}(\epsilon) = \mathcal{O}\left(\frac{1}{\epsilon^{\frac{q}{\chi}}}\right)$$

Chapter 3

Particle models in diffusion processes

3.1 Diffusion and dispersion

3.1.1 Molecular diffusion

In this section some equations and concepts underlying molecular diffusion are presented. Molecular diffusion itself is not of great direct consequence in the environmental problems, it should be taken into account only on a microscopic scale. However, in many cases environmental dispersion problems can be described by processes that are strongly analogous to molecular diffusion, but on larger scale. Also, as we will see later, the molecular diffusion plays a significant role in multi-particle models.

The law of molecular diffusion was first formulated by a German physiologist Adolph Fick in his paper [42] in 1855. Fick's law says that the flux of solute mass, that is, the mass of a solute crossing a unit per unit time in a given direction, is proportional to the gradient of solute concentration in that direction..

For a one-dimensional diffusion process, Fick's law can be stated mathematically as

$$q = -D_m \frac{\partial C}{\partial x} \tag{3.1}$$

where q is the solute mass flux, C is the concentration, D_m is the coefficient of proportionality, called diffusion coefficient. The minus sign indicates transport is from high to low concentration. For diffusion in three dimensions Fick's law can be written as

$$q = -D_m \left(\frac{\partial C}{\partial x_1} + \frac{\partial C}{\partial x_2} + \frac{\partial C}{\partial x_3} \right) \tag{3.2}$$

The equations (3.1) and (3.2) can be written in another form if we use the law of mass conservation [43]

$$\frac{\partial C}{\partial t} = D_m \frac{\partial^2 C}{\partial^2 x} \tag{3.3}$$

and

$$\frac{\partial C}{\partial t} = -D_m \left(\frac{\partial^2 C}{\partial (x_1)^2} + \frac{\partial^2 C}{\partial (x_2)^2} + \frac{\partial^2 C}{\partial (x_3)^2} \right) \tag{3.4}$$

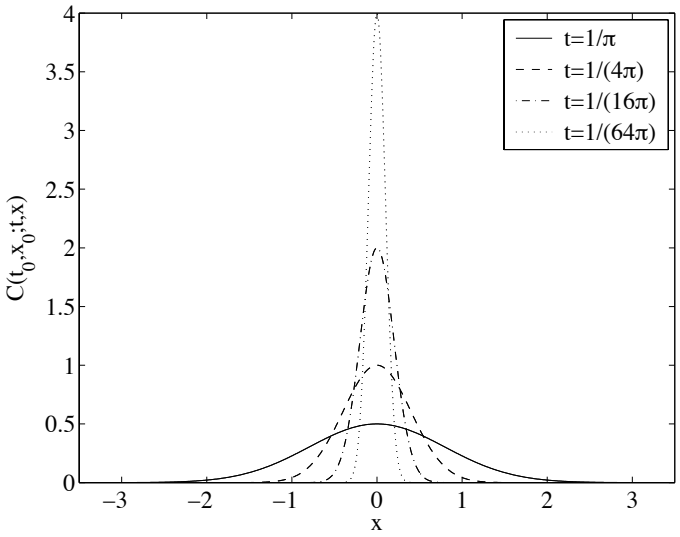


Figure 3.1: The initial delta function can be thought as 'spike' distribution. This illustration uses $M = 1$ kg and $D = 1$ kg/m.

Let us notice that the equations (3.3) and (3.4) describe the spreading of mass in a fluid with no mean velocity.

Example

Suppose that the solute is spreading in one direction. Suppose also that the initial slug of mass M was introduced at time t_0 at the x_0 origin and there are no boundaries to prevent the mass from diffusing to infinity in both directions. Mathematically this initial condition can be written as

$$C(x_0, t_0; x, t_0) = M\delta(x - x_0) \tag{3.5}$$

Physically the delta function represents a unit mass concentrated into an infinitely small space with an infinitely large concentration and $M\delta(x - x_0)$ represents a mass M concentrated into a very small space. For example, if some accident took place and some pollutant spilled we can represent the initial concentration distribution by a delta function. The direct solution of the diffusion equation (3.4) with the initial condition (3.5) is the function

$$C(x_0, t_0; x, t_0) = \frac{M}{\sqrt{4\pi D_m(t - t_0)}} \exp\left(-\frac{(x - x_0)^2}{4D_m(t - t_0)}\right)$$

With $M = 1$kg, $C(t_0, x_0; t, x)$ is the density function of a normal distribution with parameters $(x_0, 2D_m(t - t_0))$. Examples of this concentration for different moment of times t are shown on Figure 3.1.

Now suppose that the fluid is moving with velocity $\boldsymbol{u}$, whose components in the x_1, x_2 and x_3 directions are u_1, u_2 and u_3. We call the transport by the mean motion of the

fluid "advection" and assume that transport by advection and by diffusion are separate, additive processes. We also assume that we are dealing with the molecular diffusion in laminar flow, so that the diffusion coefficient has a constant value D_m in all directions.

First consider again a one-dimensional case. Suppose that the solute is spreading in the x_1-direction. Then the rate of mass transport through a unit area in the x_2x_3-plane is the quantity (u_1C), because the unit area multiplied by the concentration of mass transport is the advective plus the diffusive flux

$$q = \underbrace{u_1C}_{\text{advective flux}} + \underbrace{\left(-D_m\frac{\partial C}{\partial x}\right)}_{\text{diffusion flux}} \tag{3.6}$$

By applying the law of mass conservation we can obtain the diffusion equation as before plus the additional advective term (see, for instance, [43]):

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x_1}(u_1C) + D_m\frac{\partial^2 C}{\partial x^2} \tag{3.7}$$

In the three-dimensional case this equation takes the form

$$\frac{\partial C}{\partial t} = -\frac{\partial(u_1C)}{\partial x_1} - \frac{\partial(u_2C)}{\partial x_2} - \frac{\partial(u_3C)}{\partial x_3} + D_m\left(\frac{\partial^2 C}{\partial(x_1)^2} + \frac{\partial^2 C}{\partial(x_2)^2} + \frac{\partial^2 C}{\partial(x_3)^2}\right) \tag{3.8}$$

3.1.2 Turbulent diffusion

In the previous section we dealt with molecular diffusion in laminar flow. However, the majority of flows in nature is turbulent. It is difficult to define a turbulent fluid motion and determine exact criteria to differ the turbulent flow from laminar.

In practice one can use the Reynolds number to estimate if a fluid flow is laminar or turbulent. It is named after Osborne Reynolds (1842-1912), who proposed it in 1883. Typically it is given as follows

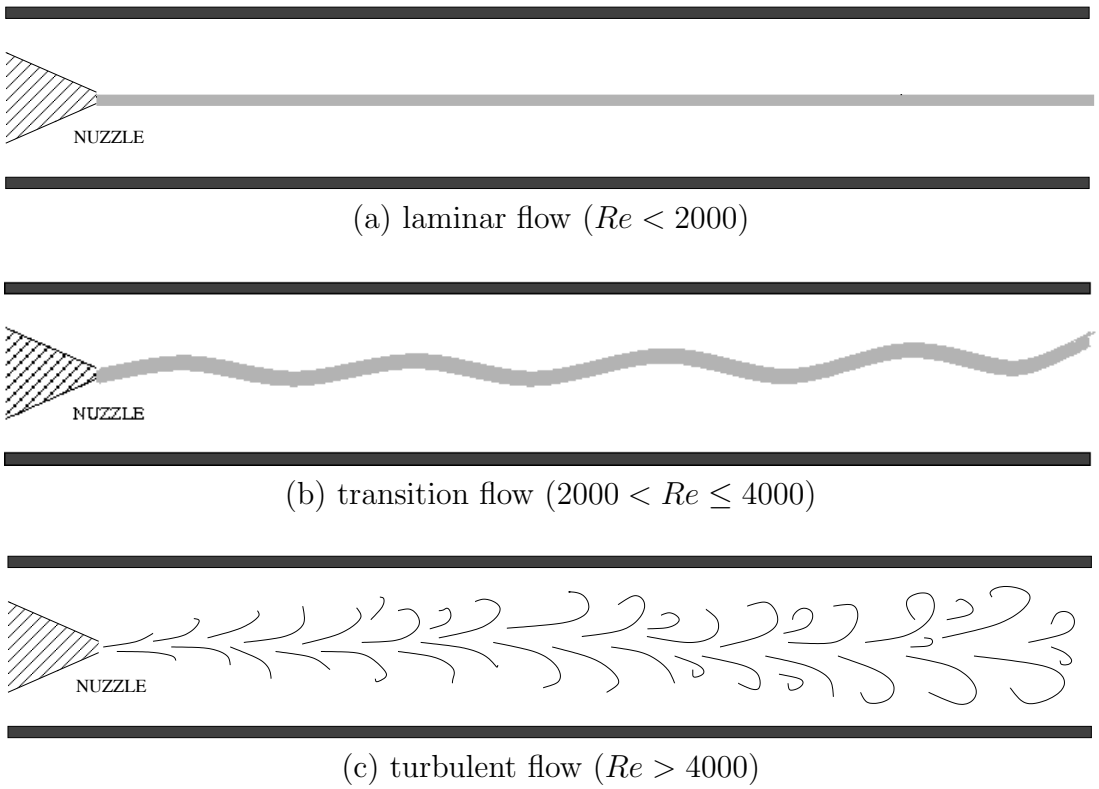
$$Re = \frac{\rho u\mathcal{L}}{\mu}$$

where u means fluid velocity, ρ is fluid density, $\mathcal{L}$ is characteristic length and μ is (absolute) dynamic viscosity.

Osborne Reynold investigated the phenomena of the turbulence in an experiment which has become a classic in fluid mechanics. A filament of dye was introduced on the centerline at the upstream end of a pipe. In a laminar flow the filament spreads along the centerline, but in turbulent flow the dye spreads across the pipe (see Figure 3.2). This experiment shows that the mass introduced at a point will spread in turbulent flow much faster than in laminar flow.

In turbulent flow the velocities measured in the fluid are unsteady and possess a appreciate random components (see Figure 3.3), while in laminar flow the velocity is constant. Often the flow which may fluctuate between laminar and turbulent is marked out and called transition flow.

The idea of the mechanism of the turbulence was first developed by English physicist Geoffrey Taylor [123, 124] and Lewis Richardson [96]. The turbulence consists of



(a) laminar flow ($Re < 2000$)

(b) transition flow ($2000 < Re \leq 4000$)

(c) turbulent flow ($Re > 4000$)

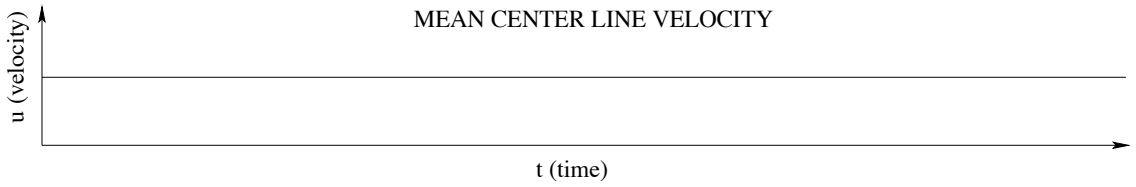
Figure 3.2: Dye introduced at the upstream end of a pipe in (a) laminar, (b) transition and (c) turbulent flow

a hierarchy of "eddies" (i.e. disturbances or nonhomogeneties) of various orders. The eddies of a given size arise as a result of the loss stability of larger eddies of the preceding size, borrowing their energy, and, in their own turn, losing their stability and generating smaller eddies to which they transmit their energy. Thus there arises a so-called "cascade process" of breaking-down of eddies in which the energy of the overall flow is transmitted to motions of smaller and smaller scale, down to motion of the smallest possible scale, which is stable. This stable extremely small-scale motion is characterized by a sufficiently small Reynold number. Thus it follows that the viscosity will play an important role and there will be a considerable dissipation of kinetic energy into heat.

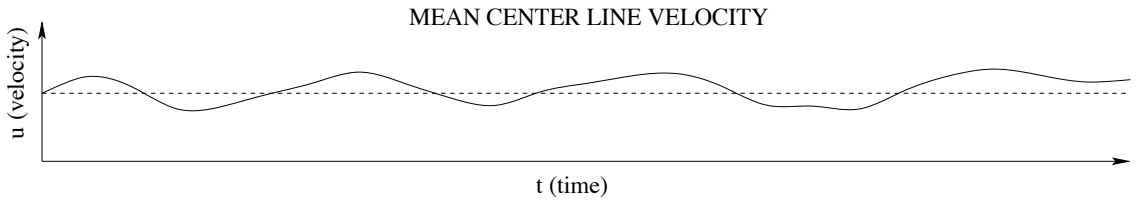
The further development of the theory of the turbulence was given by Kolmogorov and Obukhov in 1941 (see [70, 71]). More about the physics of the turbulence can be found in [6, 87].

Further we assume that the character of the turbulence remains steady, or, in other words, the variance of the velocity is steady and does not change with time or position. Fluid dynamicists call this stationary homogeneous turbulence.

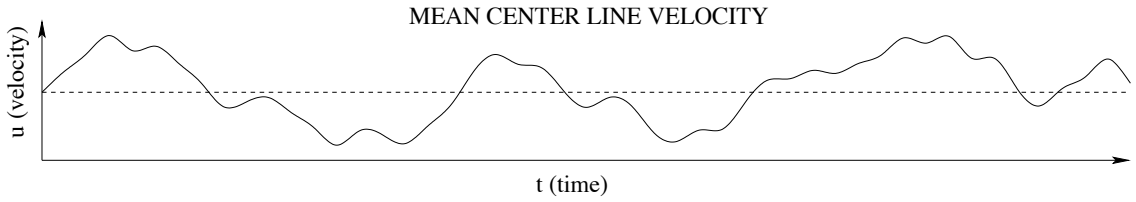
The instantaneous flow $\boldsymbol{\mathcal{U}}$ at any point can be written as the sum of the time-averaged



(a) laminar flow ($Re < 2000$)



(b) transition flow ($2000 < Re \leq 4000$)



(c) turbulent flow ($Re > 4000$)

Figure 3.3: The instantaneous velocity measured at the center of a pipe in (a) laminar, (b) transition and (c) turbulent flow

value $\boldsymbol{u}$ at that point plus a variation $\boldsymbol{\mathcal{U}}'$ from the average

$$\boldsymbol{\mathcal{U}} = \boldsymbol{u} + \boldsymbol{\mathcal{U}}'$$

as well as instantaneous concentration

$$\mathcal{C} = C + \mathcal{C}'$$

Substituting $\boldsymbol{\mathcal{U}}$ and $\mathcal{C}$ into equation (3.8) leads to the following equation

$$\begin{aligned}\frac{\partial \mathcal{C}}{\partial t} = & -\frac{\partial (u_1 + \mathcal{U}_1')(C + \mathcal{C}')}{\partial x_1} - \frac{\partial (u_2 + \mathcal{U}_2')(C + \mathcal{C}')}{\partial x_2} - \frac{\partial (u_3 + \mathcal{U}_3')(C + \mathcal{C}')}{\partial x_3} + \\ & D_m \left(\frac{\partial^2 (C + \mathcal{C}')}{\partial (x_1)^2)} + \frac{\partial^2 (C + \mathcal{C}')}{\partial (x_2)^2} + \frac{\partial^2 (C + \mathcal{C}')}{\partial (x_3)^2} \right)\end{aligned} \tag{3.9}$$

Integrating this equation to obtain the time averaged leads to the the equation (a more detailed discussion can be found in [54])

$$\begin{aligned}\frac{\partial C}{\partial t} = & -\left(\frac{\partial (u_1 C)}{\partial x_1} + \frac{\partial (u_2 C)}{\partial x_2} + \frac{\partial (u_3 C)}{\partial x_3} \right) - \frac{\partial (\overline{\mathcal{U}_1 \mathcal{C}_1})}{\partial x_1} - \frac{\partial (\overline{\mathcal{U}_2 \mathcal{C}_2})}{\partial x_2} - \frac{\partial (\overline{\mathcal{U}_3 \mathcal{C}_3})}{\partial x_3} \\ & + D_m \left(\frac{\partial^2 C}{\partial (x_1)^2} + \frac{\partial^2 C}{\partial (x_2)^2} + \frac{\partial^2 C}{\partial (x_3)^2} \right)\end{aligned} \tag{3.10}$$

Here $\overline{\mathcal{U}_i\mathcal{C}}$ $(i = 1, 2, 3)$ represent the time-averaged turbulent advective transport, i..e, transport associated with the turbulent fluctuations of velocity and concentration. For engineering purposes, these terms are often treated as turbulent diffusion terms and, since, turbulent diffusion is normally much larger than the molecular diffusion, the molecular terms are frequently incorporated into the turbulent diffusion terms

$$D_m\frac{\partial C}{\partial x_i} - \overline{\mathcal{U}_i\mathcal{C}} = k_i\frac{\partial C}{\partial x_i}, \quad i = 1, 2, 3 \tag{3.11}$$

Substitution Equation (3.11) into (3.10) leads to the following transport equation

$$\begin{aligned}\frac{\partial C}{\partial t} = & -\left(\frac{\partial(u_1C)}{\partial x_1} + \frac{\partial(u_2C)}{\partial x_2} + \frac{\partial(u_3C)}{\partial x_3}\right) + \\ & \frac{\partial}{\partial x_1}\left(k_1\frac{\partial C}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k_2\frac{\partial C}{\partial x_2}\right) + \frac{\partial}{\partial x_3}\left(k_3\frac{\partial C}{\partial x_3}\right)\end{aligned} \tag{3.12}$$

3.2 Vertically-integrated transport model

For some transport models the concentration in one of the three directions may be very small as compared with two other directions. The transport models in rivers and in unstratified shallow flow can be taken as example. For these cases it is more convenient to define the depth-averaged concentrations and velocities in two directions (x_1 and x_2). The time-averaged flows u_1 and u_2 at point can be written as follows

$$\begin{aligned}u_1 &= \tilde{u}_1 + u_1' \\ u_2 &= \tilde{u}_2 + u_2'\end{aligned} \tag{3.13}$$

where $\tilde{u}_1$ and $\tilde{u}_2$ are the depth-averaged velocities at point and u_1', u_2' are the variation from the depth average, caused by vertical distribution of velocities u_1 and u_2. Because of transport processes, the concentration is not completely uniform over the depth and, as a result, the concentration C can be written as the sum of depth-averaged concentration $\tilde{C}$ and the vertical variation concentration C'

$$C = \tilde{C} + C' \tag{3.14}$$

By substituting the equation (3.13) into (3.12) the following equation may be obtained

$$\begin{aligned}\frac{\partial\tilde{C} + C'}{\partial t} = & -\left(\frac{\partial((\tilde{u}_1 + u_1')(\tilde{C} + C'))}{\partial x_1} + \frac{\partial((\tilde{u}_2 + u_2')(\tilde{C} + C'))}{\partial x_2} + \frac{\partial(u_3(\tilde{C} + C')}{\partial x_3}\right) + \\ & \frac{\partial}{\partial x_1}\left(k_1\frac{\partial(\tilde{C} + C')}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k_2\frac{\partial(\tilde{C} + C')}{\partial x_2}\right) + \frac{\partial}{\partial x_3}\left(k_3\frac{\partial(\tilde{C} + C')}{\partial x_3}\right)\end{aligned} \tag{3.15}$$

By integrating the equation (3.15) over the depth we will obtain the following equation

$$\begin{aligned}\frac{\partial(\tilde{C}H)}{\partial t} = & -\frac{\partial(\tilde{u}_1\tilde{C}H)}{\partial x_1} - \frac{\partial(\tilde{u}_2\tilde{C}H)}{\partial x_2} - \frac{\partial(\overline{u_1'C'}H)}{\partial x_1} - \frac{\partial(\overline{u_2'C'}H)}{\partial x_2} \\ & + \frac{\partial}{\partial x_1}\left(k_1\frac{\partial\tilde{C}}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(k_2\frac{\partial\tilde{C}}{\partial x_2}\right)\end{aligned} \tag{3.16}$$

Here the expressions $\overline{u_1'C'}$ and $\overline{u_2'C'}$ indicate the depth-averaged values, i.e. the advective transport associated with the variation of velocities in vertical direction. These terms may be represented as Fickian (for more details see [54])

$$\begin{aligned} -k_1 H\frac{\partial \tilde{C}}{\partial x_1} + \overline{u_1'C'}H &= -D_1 H\frac{\partial \tilde{C}}{\partial x_1} \\ -k_2 H\frac{\partial \tilde{C}}{\partial x_2} + \overline{u_2'C'}H &= -D_2 H\frac{\partial \tilde{C}}{\partial x_2} \end{aligned} \tag{3.17}$$

The resulting depth-averaged advection-diffusion equation has the form (for sake of simplicity the tildes have been dropped out)

$$\frac{\partial (CH)}{\partial t} = -\frac{\partial (u_1 CH)}{\partial x_1} - \frac{\partial (u_2 CH)}{\partial x_2} + \frac{\partial}{\partial x_1}\left(D_1 H\frac{\partial C}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(D_2 H\frac{\partial C}{\partial x_2}\right) \tag{3.18}$$

3.3 Random walk models

There are two basic approaches to describe the transport processes. One can adopt the Eulerian view point of view and solve the advection-diffusion equations (3.12) and (3.18) numerically, for instance, using finite difference or finite difference element methods. Another approach is to use random walk models. By interpreting the advection-diffusion equation as a Fokker-Planck equation the corresponding stochastic model is used for modeling the movement of a single particle of pollutant. By simulating the positions of many particles we can describe the pollutant transport process. The stochastic approach is connected with the Lagrangian point of view.

3.3.1 Three-dimensional random walk model

Let consider a three dimensional space with points $\boldsymbol{x} = (x_1, x_2, x_3)$ and define the vector function $\boldsymbol{a}$ and $\boldsymbol{\sigma}$ according to:

$$a(s, \boldsymbol{x}) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \begin{pmatrix} \dfrac{\partial k_1}{\partial x_1} \\ \dfrac{\partial k_2}{\partial x_2} \\ \dfrac{\partial k_3}{\partial x_3} \end{pmatrix} \tag{3.19}$$

$$\sigma(s, \boldsymbol{x}) = \begin{pmatrix} \sqrt{2k_1} & 0 & 0 \\ 0 & \sqrt{2k_2} & 0 \\ 0 & 0 & \sqrt{2k_3} \end{pmatrix} \tag{3.20}$$

By substituting (3.19) and (3.20) into the Fokker-Planck equation (2.39) we will obtain the advection-diffusion equation (3.12) with initial condition

$$C(t, \boldsymbol{x}; s, \boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y})$$

This initial condition means that the pollutant was released at time t at location $\boldsymbol{x}$. The corresponding system of stochastic differential equations has the form

$$\begin{cases} dX_1 = \left(u_1 + \dfrac{\partial k_1}{\partial x_1}\right) ds + \sqrt{2k_1} dW_1 \\ dX_2 = \left(u_2 + \dfrac{\partial k_2}{\partial x_2}\right) ds + \sqrt{2k_2} dW_2 \\ dX_3 = \left(u_3 + \dfrac{\partial k_3}{\partial x_3}\right) ds + \sqrt{2k_3} dW_3 \\ \boldsymbol{X}(t, \boldsymbol{y}) = \boldsymbol{x} \end{cases} \tag{3.21}$$

3.3.2 Two-dimensional random walk model

Now consider two-dimensional space with point $\boldsymbol{x} = (x_1, x_2)$ and define the vector functions $\boldsymbol{a}$ and $\boldsymbol{\sigma}$ as

$$a(s, \boldsymbol{x}) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} \dfrac{D_1}{H}\dfrac{\partial H}{\partial x_1} \\ \dfrac{D_2}{H}\dfrac{\partial H}{\partial x_2} \end{pmatrix} + \begin{pmatrix} \dfrac{\partial D_1}{\partial x_1} \\ \dfrac{\partial D_2}{\partial x_2} \end{pmatrix} \tag{3.22}$$

and

$$\sigma(s, \boldsymbol{x}) = \begin{pmatrix} \sqrt{2D_1} & 0 \\ 0 & \sqrt{2D_2} \end{pmatrix} \tag{3.23}$$

By substituting these functions into the Fokker-Plack equation (2.39) we will obtain the following equation

$$\frac{\partial p}{\partial s} = -\frac{\partial \left((u_1 + \dfrac{\partial D_1}{\partial y_1} + \dfrac{D_1 \partial H}{H \partial y_1})p\right)}{\partial y_1} - \frac{\partial \left((u_2 + \dfrac{\partial D_2}{\partial y_2} + \dfrac{D_2 \partial H}{H \partial y_2})p\right)}{\partial y_2} + \frac{\partial^2 (D_1 p)}{\partial (y_1)^2} + \frac{\partial^2 (D_2 p)}{\partial (y_2)^2} \tag{3.24}$$

Substituting $p = HC$ in this equation and rearranging terms yields

$$\frac{\partial (HC)}{\partial s} = -\frac{\partial (u_1 HC)}{\partial y_1} - \frac{\partial (u_2 HC)}{\partial y_2} + \frac{\partial}{\partial y_1}\left(H D_1 \frac{\partial C}{\partial y_1}\right) + \frac{\partial}{\partial y_2}\left(H D_2 \frac{\partial C}{\partial y_2}\right) \tag{3.25}$$

where C is the particle concentration. This equation is the vertically-integrated advection-diffusion equation (3.18). The corresponding random walk model can be written in the form

$$\begin{cases} dX_1 = \left(u_1 + \dfrac{D_1}{H}\dfrac{\partial H}{\partial x_1} + \dfrac{\partial D_1}{\partial x_1}\right) ds + \sqrt{2D_1} dW_1 \\ dX_2 = \left(u_2 + \dfrac{D_2}{H}\dfrac{\partial H}{\partial x_2} + \dfrac{\partial D_2}{\partial x_2}\right) ds + \sqrt{2D_2} dW_2 \\ \boldsymbol{X}(t, \boldsymbol{y}) = \boldsymbol{x} \end{cases} \tag{3.26}$$

3.4 Numerical methods for solving advection-diffusion equation

The numerical solution of the advection-diffusion equation has been an area of active research for many years. In order to solve the corresponding partial differential equation,

one has to rely on numerical techniques in nearly all cases. Much effort has already been put in developing satisfactory numerical methods for advection-diffusion equations. The main obstacle consists in that we need to treat simultaneously a hyperbolic term (associated with advection) and a parabolic terms (associated with dispersion), a problem that no numerical method has fully overcome [5]. In this section we consider the main approaches for dealing with this problem and discuss their advantages and disadvantages.

Most numerical methods for simulating advection-diffusion processes can be split into three categories: Eulerian, Lagrangian and mixed Eulerian-Lagrangian methods.

3.4.1 Eulerian methods

In Eulerian methods, the transport equation is solved on a fixed spatial grid. The finite element method and finite difference method are primary examples of this class of solution methods. Eulerian methods that have been used for flow simulation (see, for instance, [115]) were the earliest methods applied to transport modelling and they are still used commonly today.

Finite difference method

The classical finite difference method is probably the oldest, most popular, and conceptually simplest of the numerical procedures. The theories and solution techniques of this method have been covered in a number of standard textbooks (e.g. [59, 145, 142, 66, 95, 7]). The finite difference techniques are based upon the approximations that permit replacing differential equations by a set of approximating algebraic equations. A finite difference solution basically involves three steps: dividing the area of problem into rectangular cells which are identified with discrete points or nodes. Then, the partial differential equation is approximated using a truncated Taylor series expansion and, finally, the linear system of equations is solved with prescribed initial conditions.

Integrated finite difference method

The integrated finite difference method is an interesting extension of the classical finite difference method [90]. The domain is discretized into irregular polygons rather than rectangles. This added flexibility is balanced by a lower order of accuracy and a set of algebraic equation that generates a coefficient matrix structure that is less regular and consequently less amenable to efficient solution strategies than the classical finite difference approach [92].

Finite element method

Like the finite-difference method, the finite-element method has been widely used for numerical solution of advection-diffusion equation. The description of this method may be found in a number of comprehensive texts, including [132, 59, 145, 60, 93]. In general, the finite-element method is formulated using the method of weighted residuals which is more difficult to conceptualize than the finite-difference method. The Galerkin finite-element method, which is one of the most popular technique used in subsurface simulation, is an example of the weighted residuals technique. The domain is divided into elements and the numerical solution is the discretizations with basis functions of local support. The main

difference between finite-difference and finite-element methods is that the first method is an approximation to the differential equation, while the second one is an approximation to its solution. The finite-element method is more accurate and offers more flexibility in spatial discretiztion than the finite-difference method, but at the expense of higher numerical complexity and more computational effort.

The Eulerian methods offer the advantages and convenience of a fixed grid and they are also easy to implement. However, these methods applied to the problems with a steep plume front lead to serious problems such as numerical dispersion, artificial oscillation and global mass balance error. For instance, for advection-dominated problems, that exists in many field applications, an Eulerian method is susceptible to excessive numerical dispersion and artificial oscillation [148]. Another situation, that often occur in case of calamities at sea is the advection-diffusion problem with very high initial concentration. In this case the delta like function is assumed to be the initial concentration and the most classical method applied to this problem will produce inaccurate and unrealistic solutions.

The effect of numerical dispersion is similar to physical dispersion, but is caused by truncation error. When physical dispersion is small or negligible, numerical dispersion becomes a serious problem, leading to the smearing of concentration fronts which should have a sharp appearance [149]. The artificial oscillation is typical of some higher order schemes designed to eliminate numerical dispersion and tends to become more severe as the concentration front becomes sharper. This oscillation is exhibited as sinusoidal concentration distributions, often leading to negative concentrations [136]. Many methods for contaminant transport simulations are also non-conservative. Such non-conservative schemes are subject to non-physical loss of mass [116].

The result is that sufficiently complex problems solved by the finite difference or finite element methods may require both prohibitively small time step and fine spatial resolution of complex domain. The global Eulerian methods, limited by computer memory become unattractive once the accuracy and stability constrains require an excessive number of cells or nodes or an excessive small time step. The finite difference and finite elements methods (as well as finite volume method [89]) stay the most popular techniques, a lot of improvements and research into Eulerian methods have been done recently. The examples are Total Variation Dimishing (TVD) schemes [21, 76, 128] and the finite analytic method [79, 80].

3.4.2 Lagrangian methods

The Lagrangian method treats the transport of contaminant by modeling of large number of moving particles, and avoid solving the advection-diffusion equation directly. As a result, these Lagrangian methods are free of numerical dispersion and particularly effective in handling advection-dominated problems and problems with initial high concentration gradient. The typical example of the Lagrangian method is random walk model.

The movement of a particle is modeled with a stochastic differential equation, which is consistent with the advection-diffusion equation. By simulating the positions of many particles the advection-diffusion processes can be described. Because the particles can not leave the domain, the solution is always mass conservative. Moreover, because the concentration of the contaminant is connected with the density probability function of

the underlying stochastic process, the solution is always non-negative. The random walk methods has been applied in a number of scientific disciplines to solve transport problems: in modeling of pollutant transport in surface waters [19, 31, 52, 32, 139, 104, 117, 17], in groundwater modelling [135, 131, 77, 67], in turbulent atmosphere modelling [146, 4, 14, 105, 41, 39], in turbulent flow [125, 126, 75]

Lagrangian methods provide an accurate and efficient solution to advective dominated problems by essentially eliminating the effects of numerical dispersion and artificial oscillation. The random walk methods requires relatively little computer storage as compared with finite difference or finite element methods. However, the lack of a fixed grid or fixed coordinate in Lagrangian method may lead to numerical instability and computational difficulties ([137]). Another source of numerical error is a result of interpolation of numerical flow in arbitrary particle location that can lead to local mass balance error and solution anomalies ([77]). The random walk solution may also not be smooth if the number of particles is not large enough. As a result, it is difficult to characterize the tailing concentration [120]. The last problem can be reduced by increasing the number of particles, however it will require too much CPU time. This disadvantage can be easily compensated by using parallel processing. Because of the independence of each realization of particle movement, the random walk models are extremely suitable for the parallelization (e.g. [23]. Using the parallel processing the computational time can be significantly reduced.

Lagrangian methods are very suitable for the active transport simulation. For this kind of modelling the Eulerian formulation may be difficult, while the Lagrangian formulation requires only some simple modification. For instance, the Lagrangian methods has been applied to model the transport of flatfish larvae ([48]). The larvae are believed to actively modulate their transport process by vertical migration through a sheared flow field, influenced by factors such as light, temperature, pressure, salinity, speed and direction of the flow, availability of food and food odour, turbidity and turbulence. This aspect of the larvae behavior can be included in a Lagrangian model in which the flatfish larvae are represented by particles, which responded differently to tidal cues - in this case, hydrostatic pressure..

3.4.3 Mixed Eulerian-Lagrangian methods

The Lagrangian approach can be an alternative for the Eulerian methods, for instance in case of high initial concentration gradient or for advective dominated problems. On the other hand, the Eulerian approach is more suitable for dispersion-dominated problems, for which it provides accurate solution in reasonable time. The choice of methods depends on the given problem. Sometimes, it is not easy to classify the problem and decide which method should be applied. The mixed Eulerian-Lagrangian methods attempt to combine the advantages of Lagrangian and Eulerian methods. For instance, in case of a high initial concentration gradient, the Lagrangian approach can be applied in order to avoid the negativety of the solution and loss of mass and, then, the Eulerian method may be used [119]. To eliminate numerical dispersion or artificial oscillation of the solution the Lagrangian method is used to solve the advective term and other terms are solved by Eulerian methods. The mixed Eulerian-Lagrangian methods are conceptually attractive

and have been widely applied in field applications.

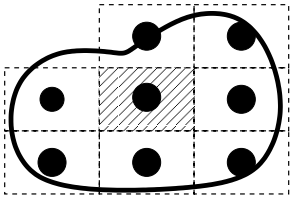
Method of characteristics (MOC)

The method of characteristics was originally applied to transport in porous media for calculation of miscible displacement in reservoir simulation [45]. This method was later implemented in a two-dimensional solute transport model [73] and in a three-dimensional transport problem [147].

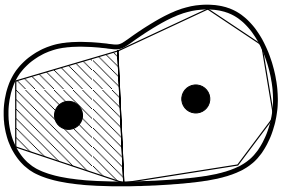
At the beginning of the simulation, a set of moving particles is distributed in the flow field either randomly or with a fixed pattern. Each particle is assigned the concentration associated with that cell (rather than the mass used in random walk methods). These particles are then tracked forward in time using particle-tracking algorithms and new cell concentrations are calculated based on the average concentration associated with the particles within a cell at the end of the time step.

The method of characteristics has been widely used in field studies. Several variations on this scheme, including modified method of characteristics (MMOC) [18, 34] and the hybrid method of characteristics (HMOC) [147, 148] were developed to improve the computational efficiency of traditional MOC and more accurately simulate dispersion-dominated systems. The MMOC is similar to the MOC except in the treatment of the advection term. Unlike the MOC technique, which tracks a large number of moving particles forward in time and keeps track of the concentration and position of each particle, the MMOC technique relies upon backtracking of particle paths, but with multiple characteristics per cell. The HMOC combines the computational advantages of MMOC and the dispersive robustness of traditional MOC by dynamically choosing which method to use depending upon the local behavior of the system. At sharp advective fronts, traditional MOC is used; elsewhere, MMOC is used.

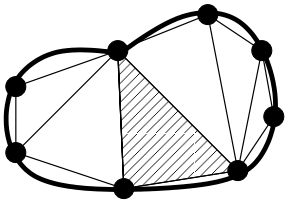
In this section only main and widely used methods for solving advection-diffusion problems were mentioned. The full overview of methods is beyond the scope of this thesis. The comprehensive summary of the numerical procedures for different areas of research are given, for instance, in [148, 138, 40, 92].



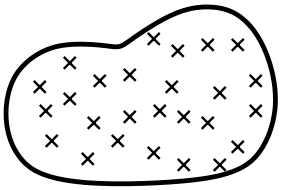
Finite difference (classical)



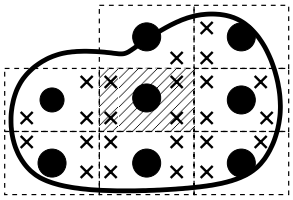
Finite difference (integrated)



Finite element (Galerkin)



Random walk



MOC

Diagrammatic representation of numerical approximation methods. Source [92]

Chapter 4

Random walk model for space-varying diffusivities

4.1 Introduction

It is common knowledge that large-scale diffusion processes in the ocean occur mostly along the isopycnal surfaces, i.e. surfaces of equal density. There is also some diapycnal diffusion. The latter is associated with a diffusion flux orthogonal to isopycnal surfaces. The diapycnal and isopycnal diffusion fluxes are commonly parameterized *ă la Fourier-Fick*, a formulation involving a diffusion tensor that is not diagonal [94]. As was seen [8, 9] many Eulerian discretizations of the isopycnal diffusion term yield discrete operators that are not monotonic - a problem which is particularly annoying. The discrete version of the isopycnal mixing parameterization can produce spurious oscillations in the tracer fields, which disagrees with the well-known properties of diffusion operators [83]. The Lagrangian method applied to simulate the transport processes along isopycnal surfaces should help avoid spurious oscillations and negative values.

In this section, random walk schemes associated with non-diagonal diffusivity tensors whose components vary in space are established for multi-dimensional cases (Section 4.2). These methods are applied for simulating the transport of a passive tracer along the isopycnal surface for two-dimensional (Section 4.3) and three-dimensional linear problems (Section 4.4). In Section 4.5 the random walk algorithm is tested for a one-dimensional settling and a diffusion model for which key properties of the solution can be derived. Finally, conclusions are drawn in Section 4.6.

This chapter is based on the article: Spivakovskaya D., Deleersnijder E. and Heemink A.W. 'Lagrangian modelling of multi-dimensional advection-diffusion with space-varying diffusivities: theory and idealized test cases', Ocean Dynamics. 57(3), pp. 189-203

4.2 Random walk model of multi-dimensional advection-diffusion

Let $\Omega \subset \mathbb{R}^d$ denote the isolated domain of interest. The concentration $C(t, \boldsymbol{x})$ can be found from the advection-diffusion equation

$$\frac{\partial C}{\partial t} = -\sum_{i=1}^{d} \frac{\partial}{\partial x_i}(u_i C) + \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(k_{ij}\frac{\partial C}{\partial x_j}\right) \tag{4.1}$$

We suppose that the velocity vector $\boldsymbol{u}$ is divergence free

$$\sum_{i=1}^{d} \frac{\partial u}{\partial x_i} = 0 \tag{4.2}$$

Also we suppose that on boundary Γ of the domain Ω the following conditions are satisfied

$$\begin{aligned} &\boldsymbol{u} \bullet \boldsymbol{n} = 0 \\ &(-\boldsymbol{K} \bullet \nabla C) \bullet \boldsymbol{n} = 0 \end{aligned} \tag{4.3}$$

Here $\boldsymbol{n}$ denotes the outward unit vector to Γ, $\boldsymbol{K}$ is the diffusivity tensor that is symmetric and positive definite (e.g. [25]). At the initial time the concentration is represented by Dirac delta function

$$C(t_0, \boldsymbol{x}_0; t_0, \boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_0) \tag{4.4}$$

This boundary value problem can be solved with the help of the following system of stochastic differential equations

$$\begin{aligned} dX_1(t) &= \left(u_1 + \sum_{j=1}^{d} \frac{\partial k_{1j}}{\partial x_j}\right) dt + \sqrt{2} \sum_{j=1}^{d} V_{1j} dW_j(t) \\ dX_2(t) &= \left(u_2 + \sum_{j=1}^{d} \frac{\partial k_{2j}}{\partial x_j}\right) dt + \sqrt{2} \sum_{j=1}^{d} V_{2j} dW_j(t) \\ &\dots \\ dX_d(t) &= \left(u_3 + \sum_{j=1}^{d} \frac{\partial k_{dj}}{\partial x_j}\right) dt + \sqrt{2} \sum_{j=1}^{d} V_{dj} dW_j(t) \\ \boldsymbol{X}(t_0) &= \boldsymbol{x}_0 \end{aligned} \tag{4.5}$$

Here the matrix $\boldsymbol{V}$ is the matrix such that $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{V}^T$. Because the matrix $\boldsymbol{K}$ is symmetric and positive definite, it can be always represented via a matrix $\boldsymbol{V}$. Usually, the matrix $\boldsymbol{V}$ is chosen the upper-triangular matrix. The transition density function $p(t_0, \boldsymbol{x}_0; t, \boldsymbol{x})$ of the stochastic process $\boldsymbol{X}(t)$ is the solution of the Fokker-Planck equation, that, in this case, is the advection-diffusion equation. Indeed, the Fokker-Planck equation, which is

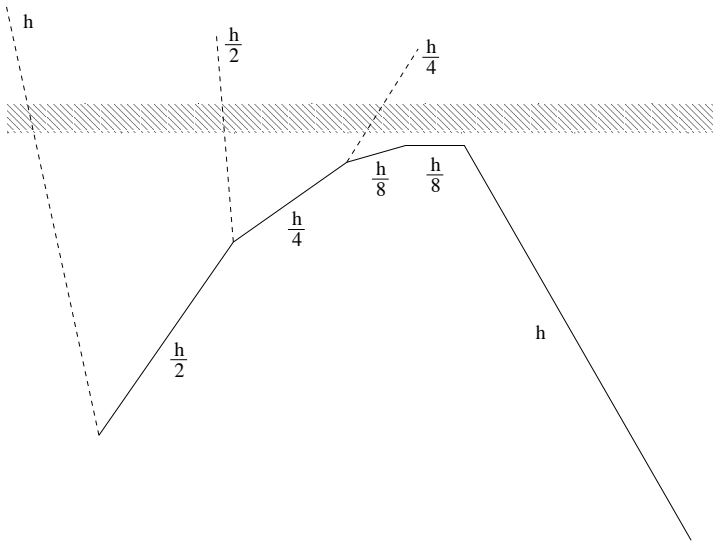


Figure 4.1: The simulation of particle's track near boundary. Source: [117]

consistent with the stochastic system (4.5) has the form

$$\begin{aligned}
&\frac{\partial p}{\partial t} = -\sum_{i=1}^{d}\frac{\partial}{\partial x_i}(u_i p) - \sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left((\sum_{j=1}^{d}\frac{\partial k_{ij}}{\partial x_j})p\right) + \frac{1}{2}\cdot 2\sum_{i,j=1}^{d}\frac{\partial^2(k_{ij}p)}{\partial x_i \partial x_j} \Longrightarrow \\
&\frac{\partial p}{\partial t} = -\sum_{i=1}^{d}\frac{\partial}{\partial x_i}(u_i p) - \sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(\frac{\partial k_{ij}}{\partial x_j}p\right) + \sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(k_{ij}\frac{\partial p}{\partial x_j} + \frac{\partial k_{ij}}{\partial x_j}p\right) \Longrightarrow \\
&\frac{\partial p}{\partial t} = -\sum_{i=1}^{d}\frac{\partial}{\partial x_i}(u_i p) + \sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(k_{ij}\frac{\partial p}{\partial x_j}\right)
\end{aligned}$$

The last equation coincides with the advection-diffusion equation (4.1). In general, a system of stochastic differential equations like (4.5) can not be solved analytically, but only numerically (see Section 2.3 for the overview of the numerical schemes).

One should be careful when particles are close to the boundaries. Theoretically, the particle cannot cross the boundary, however, because of the discretizations of the stochastic system (4.5) particles can occur outside the domain. To prevent this we need to take a sufficiently small time step. This procedure is described in [117, 118]. The original time step is halved, letting the particle travel two short time steps instead of a single big one. This process is repeated until the particle does not cross the boundary anymore. The result is that the particle trajectory bends along a certain boundary, but never crosses it.

It is a common practice to calculate the averages over the grids in the space, i.e. by calculating the number of particles in a box (see, for instance, [97]). The estimation of the concentration is then obtained by multiplying the number of particles in each box with their mass and dividing the total mass by the volume of the box. This estimator depends on the choice of the boxes: their sizes and the centers of the averaging intervals. This method is ideal if the volume average over such box is exactly what the modeler wants.

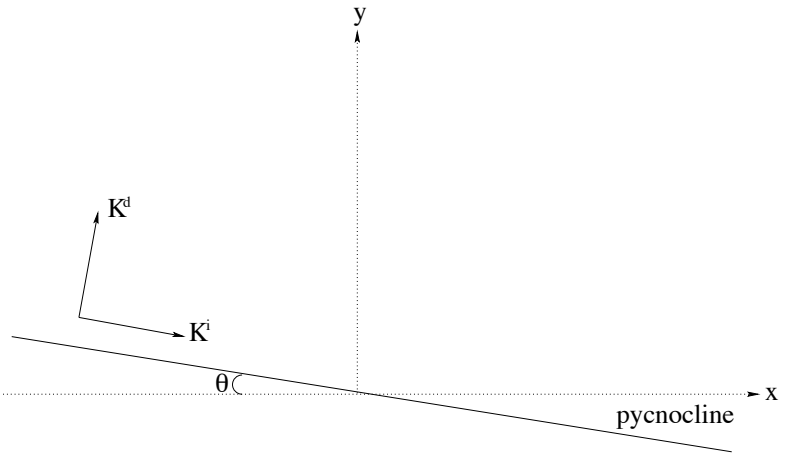


Figure 4.2: The two-dimensional linear model of iso- and diapycnal diffusion

However, for many cases this method is not sufficiently accurate and requires much CPU time. On the one hand, we cannot chose the number of boxes to be too large, otherwise the concentration function becomes very irregular or noisy (having large variance). On the other hand we cannot describe the concentration function in one box more precisely than a constant.

Additional information about the advantages and disadvantages of the kernel estimator as compared with the box estimator for stochastic particle models can be found in [50]. Further, for one of the test problems, the two methods are compared and it can be seen that the kernel estimator gives the better approximation.

4.3 Linear two-dimension iso- and dia-pycnal diffusion problem

First we discuss the Lagrangian method described above for iso- and dia-pycnal diffusion processes, that can be solved analytically. If homogeneity can be assumed along one horizontal coordinate, a two-dimensional problem is to be dealt with. For a large-scale ocean model the formulation of the diffusion model resorts to two diffusivity coefficients, k^i and k^d, which are isopycnal and diapycnal diffusivities, respectively. In the principal axes, the diffusivity tensor reads

$$\boldsymbol{k} = \begin{pmatrix} k^i & 0 \\ 0 & k^d \end{pmatrix} \tag{4.6}$$

The first principal axis is parallel to the isopycnal direction, while the other is orthogonal to it. Let θ denote the angle between the horizontal and the isopycnal direction, in

the horizontal-vertical coordinates (x, z) the diffusivity tensor is ([8, 94])

$$\boldsymbol{k} = \begin{pmatrix} k_{xx} & k_{xz} \\ k_{zx} & k_{zz} \end{pmatrix} = \begin{pmatrix} \cos^2\theta k^i + \sin^2\theta k^d & \sin\theta\cos\theta(k^i - k^d) \\ \sin\theta\cos\theta(k^i - k^d) & \sin^2\theta k^i + \cos^2\theta k^d \end{pmatrix} \tag{4.7}$$

We consider a two-dimensional model and suppose that the advective processes can be neglected and the concentration $C(t, \boldsymbol{x}) := C(0, \mathbf{0}; t, \boldsymbol{x})$ of a passive tracer obeys the following equation:

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}\left(k_{xx}\frac{\partial C}{\partial x} + k_{xz}\frac{\partial C}{\partial z}\right) + \frac{\partial}{\partial z}\left(k_{zx}\frac{\partial C}{\partial x} + k_{zz}\frac{\partial C}{\partial z}\right) \tag{4.8}$$

The domain of interest is assumed to be infinite,

$$-\infty < x < \infty, \quad -\infty < z < \infty$$

and the initial concentration is a Dirac impulse, i.e.

$$C(0, \boldsymbol{x}) = \delta(\boldsymbol{x} - \mathbf{0}) = \delta(x - 0)\delta(z - 0)$$

It is useful to introduce dimensionless variables. The time and space coordinate are transformed as follows:

$$t' = \frac{t}{T}, \quad x' = \frac{x}{\mathcal{L}_h}, \quad z' = \frac{z}{\mathcal{L}_\nu} \tag{4.9}$$

where T, $\mathcal{L}_h$ and $\mathcal{L}_\nu$ denote the appropriate timescale, horizontal length scale and vertical length scale. It is convenient to define the latter in such a way that

$$T = \frac{\mathcal{L}_h^2}{k^i} = \frac{\mathcal{L}_\nu^2}{k^d} \tag{4.10}$$

The ratio of the vertical length scale to the horizontal one, i.e. the aspect ratio, is

$$\alpha = \frac{\mathcal{L}_\nu}{\mathcal{L}_h} \tag{4.11}$$

The concentration is scaled as follows:

$$C' = \frac{C}{1/(\mathcal{L}_h\mathcal{L}_\nu)} \tag{4.12}$$

The equation to be solved now reads

$$\frac{\partial C'}{\partial t'} = \frac{\partial}{\partial x'}\left(k'_{xx}\frac{\partial C'}{\partial x'} + k'_{xz}\frac{\partial C'}{\partial z'}\right) + \frac{\partial}{\partial z'}\left(k'_{zx}\frac{\partial C'}{\partial x'} + k'_{zz}\frac{\partial C'}{\partial z'}\right) \tag{4.13}$$

with

$$\begin{aligned} k'_{xx} &= \frac{k_{xx}}{\mathcal{L}_h^2/T} = \cos^2\theta + \alpha^2\sin^2\theta \\ k'_{xz} &= \frac{k_{xz}}{\mathcal{L}_h\mathcal{L}_\nu/T} = (\alpha^{-1} - \alpha)\sin\theta\cos\theta \\ k'_{zx} &= \frac{k_{zx}}{\mathcal{L}_h\mathcal{L}_\nu/T} = (\alpha^{-1} - \alpha)\sin\theta\cos\theta \\ k'_{zz} &= \frac{k_{zz}}{\mathcal{L}_\nu^2/T} = \cos^2\theta + \alpha^{-2}\sin^2\theta \end{aligned} \tag{4.14}$$

The initial condition is

$$C'(0, \boldsymbol{x}') = \delta(\boldsymbol{x}' - \mathbf{0}) = \delta(x' - 0)\delta(z' - 0) \tag{4.15}$$

From here on, only dimensionless variables quantities will be dealt with. So, for the sake of simplicity, the primes will be dropped.

The general form of the solution to the differential problem (4.13,4.15) is

$$C(t, \boldsymbol{x}) = \frac{1}{4\pi t\sqrt{\det \boldsymbol{k}}} \exp\left[-\frac{\boldsymbol{x}^T \bullet \boldsymbol{k}^{-1} \bullet \boldsymbol{x}}{4t}\right] \tag{4.16}$$

It is readily seen that the determinant of $\boldsymbol{k}$ is equal to unity, i.e.

$$\det \boldsymbol{k} = 1$$

while the inverse simply is

$$\boldsymbol{k}^{-1} = \begin{pmatrix} \cos^2\theta + \alpha^{-2}\sin^2\theta & -(\alpha^{-1} - \alpha)\sin\theta\cos\theta \\ -(\alpha^{-1} - \alpha)\sin\theta\cos\theta & \cos^2\theta + \alpha^2\sin^2\theta \end{pmatrix}$$

Then, the solution may be rewritten as follows:

$$C(t, x, z) = \frac{1}{4\pi t} \exp\left[\frac{-(\cos\theta x + \alpha\sin\theta z)^2 - (\cos\theta z - \alpha^{-1}\sin\theta x)^2}{4t}\right] \tag{4.17}$$

In the ocean, the slope of the isopycnal surfaces and the aspect ratio is usually small - which is why [20] suggested a simplified version of the isopycnal diffusivity tensor. For numerical experiments, the following values can be used [83]

$$\theta \approx 10^{-3} \approx \alpha \tag{4.18}$$

The random walk model (4.5) can now be cast into the following form

$$\begin{aligned} dX(t) &= \sqrt{2}V_{xx}dW_1(t) \\ dZ(t) &= \sqrt{2}V_{xz}dW_1(t) + \sqrt{2}V_{zz}dW_2(t) \\ X(0) &= Z(0) = 0 \end{aligned} \tag{4.19}$$

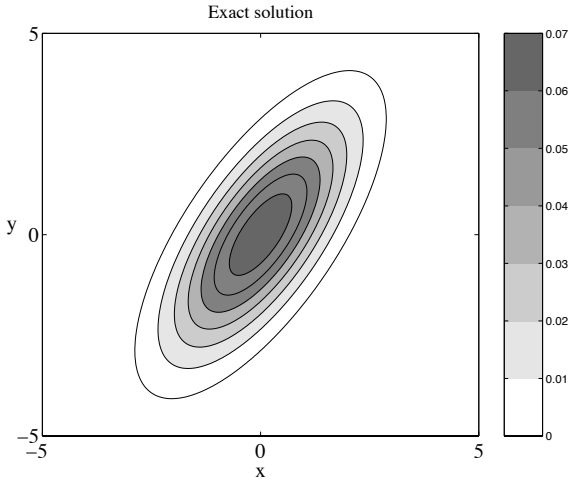
where $\boldsymbol{k} = \boldsymbol{V}\boldsymbol{V}^T$ and the matrix $\boldsymbol{V}$ is

$$\boldsymbol{V} = \begin{pmatrix} V_{xx} & 0 \\ V_{xz} & V_{zz} \end{pmatrix} = \begin{pmatrix} \sqrt{k_{xx}} & 0 \\ \dfrac{k_{xz}}{\sqrt{k_{xx}}} & \sqrt{k_{zz} - \dfrac{k_{xz}^2}{k_{xx}}} \end{pmatrix}$$

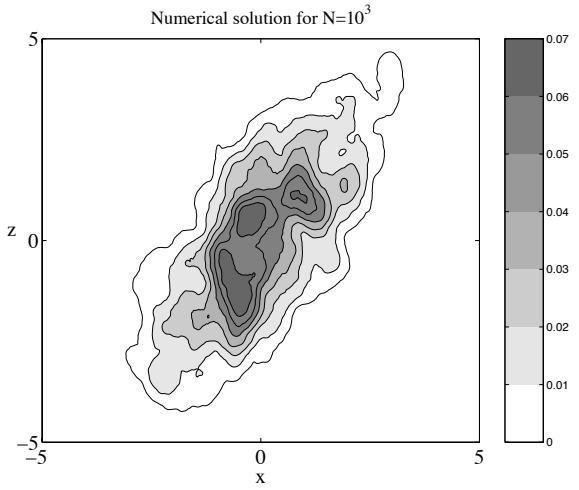
To simulate the track of each particle we used the Euler scheme (2.23)

The following parameters have been used for the present model:

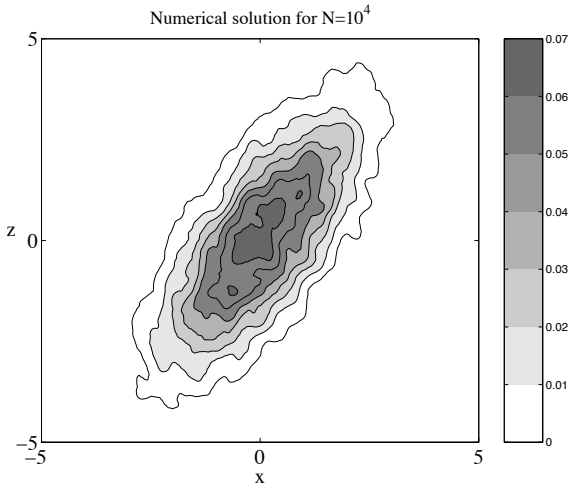
aspect ratio	α	10^{-3}
slope of the isopycnal surface	θ	10^{-3}
timestep	Δt	10^{-2}
time of simulation	T	1
kernel function	$K(\boldsymbol{u})$	Epanechnikov



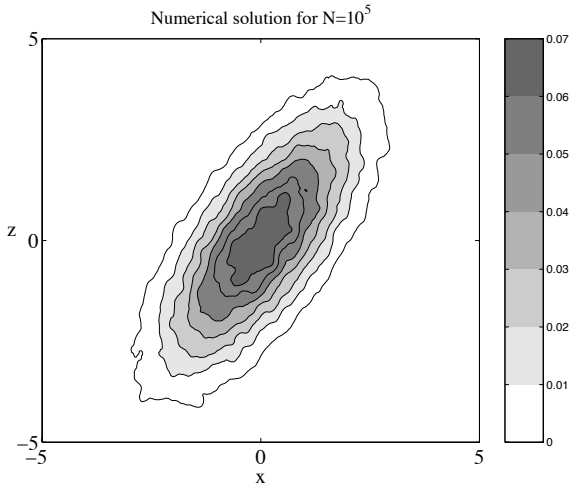
(a) The exact solutions



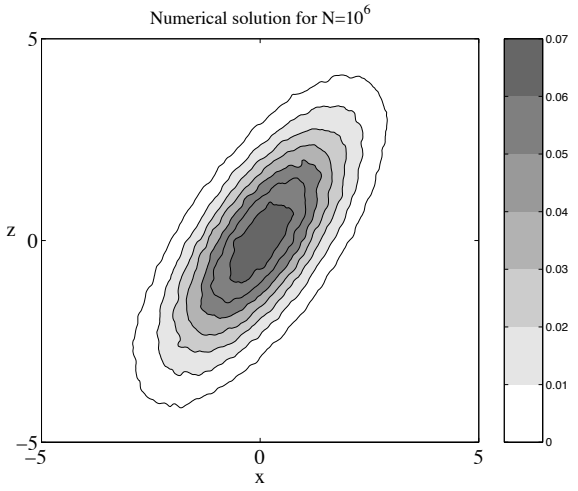
(b) $N = 10^3$



(c) $N = 10^4$



(d) $N = 10^5$



(e) $N = 10^6$

Figure 4.3: The analytical (a) and numerical solutions (for (b) 10^3, (c) 10^4, (d) 10^5 and (e) 10^6 particles) for two-dimensional linear iso- and diapycnal advection-diffusion problem

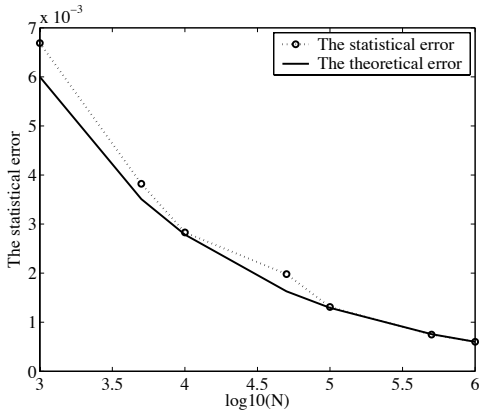


Figure 4.4: The comparison of the statistical error with the theoretical in case of a two-dimensional advection-diffusion process

The exact solution is shown on Figure 4.3(a) and the numerical solutions for different numbers of particles are shown on Figure 4.3(b)-(e). In Table 4.3 the results of the estimation of the concentration at the point are shown. For a fixed number of particles N we repeated the experiment several times in order to calculate the statistical error as follows

$$\epsilon_{\text{statistical}} = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}\left(C^{(m)}(T,\boldsymbol{x}) - \overline{C}(T,\boldsymbol{x})\right)^2} \tag{4.20}$$

where M is a number of experiments (in our case the value $M = 30$ was taken), $C^{(m)}$, $m = 1, \ldots, M$ is the concentration obtained in the mth simulation and

$$\overline{C}(T,\boldsymbol{x}) = \frac{1}{M}\sum_{m=1}^{M} C^{(m)}(T,\boldsymbol{x}) \tag{4.21}$$

is the average value of concentration. The formula (2.47) suggests that the rate of the convergence of $\epsilon_{\text{statistical}}$ is supposed to be $\mathcal{O}(N^{-1/3})$. On the Figure 4.4 it can be seen the statistical and the theoretical errors. It is clear that the statistical error has the same order of convergence as the theoretical error.

N	Average value	Statistical error	exact
10^3	0.0295	0.0067	0.0313
$5 * 10^3$	0.0313	0.0038	0.0313
10^4	0.0308	0.0028	0.0313
$5 * 10^4$	0.0303	0.0020	0.0313
10^5	0.0312	0.0013	0.0313
$5 * 10^5$	0.0314	0.0007	0.0313
10^6	0.0312	0.0006	0.0313

Table 4.1: The concentration at the location $(1.8, 2.5)$

4.4 Linear three-dimension iso- and dia-pycnal diffusion problem

Now consider the three-dimensional problem, for which we again suppose that the advection process can be neglected and $\boldsymbol{u} = \mathbf{0}$). As for the previous two-dimensional problem, the formulation of diffusion model resorts to two diffusivity coefficients, k^i and k^d , which are the isopycnal and diapycnal diffusivities, respectively. In the principal axes, the diffusivity tensor reads

$$\boldsymbol{k} = \begin{pmatrix} k^i & 0 & 0 \\ 0 & k^i & 0 \\ 0 & 0 & k^d \end{pmatrix} \tag{4.22}$$

The z-principal axe is perpendicular to the isopycnal plane. To rotate the coordinate system associated with the isopycnal surface into the geodesic (x, y, z)-system we need two angles θ and γ ([94]) and the diffusivity tensor takes the form

$$\boldsymbol{k} = \begin{pmatrix} k_{xx} & k_{xy} & k_{xz} \\ k_{yx} & k_{yy} & k_{yz} \\ k_{zx} & k_{zy} & k_{zz} \end{pmatrix} \tag{4.23}$$

where

$$\begin{aligned}
k_{xx} &= k^i \cos^2\theta + \sin^2\theta(k^i \sin^2\gamma + k^d \cos^2\gamma) \\
k_{xy} &= k_{yx} = -\cos\gamma \sin\gamma \sin^2\theta(k^i - k^d) \\
k_{xz} &= k_{zx} = \cos\gamma \sin\theta \cos\theta(k^i - k^d) \\
k_{yy} &= k^i \cos^2\theta + \sin^2\theta(k^i \cos^2\gamma + k^d \sin^2\gamma) \\
k_{yz} &= k_{zy} = \sin\gamma \sin\theta \cos\theta(k^i - k^d) \\
k_{zz} &= k^i \sin^2\theta + k^d \cos^2\theta
\end{aligned}$$

As a consequence, the advection-diffusion equation may be written in the form

The section is based on the article : Spivakovskaya D., Deleersnijder E. and Heemink A.W. 2006 'Random walk model in case of iso- and diapycnal diffusion' Proceedings of European Conference on Computational Fluid Dynamics ECCOMAS CFD 2006, P. Wesseling, E. Oñate and J. Périaux (Eds)

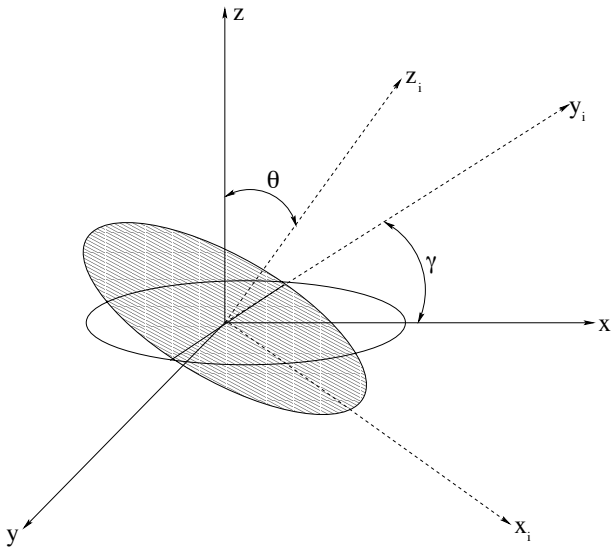


Figure 4.5: The three-dimensional linear model of iso- and diapycnal diffusion. Source: [94]

$$\begin{aligned}\frac{\partial C}{\partial t} = \quad & \frac{\partial}{\partial x}\left(k_{xx}\frac{\partial C}{\partial x} + k_{xy}\frac{\partial C}{\partial y} + k_{xz}\frac{\partial C}{\partial z}\right) + \\ & \frac{\partial}{\partial y}\left(k_{yx}\frac{\partial C}{\partial x} + k_{yy}\frac{\partial C}{\partial y} + k_{yz}\frac{\partial C}{\partial z}\right) + \\ & \frac{\partial}{\partial z}\left(k_{zx}\frac{\partial C}{\partial x} + k_{zy}\frac{\partial C}{\partial y} + k_{zz}\frac{\partial C}{\partial z}\right)\end{aligned} \tag{4.24}$$

The domain of interest is assumed to be infinite,

$$-\infty < x < \infty, \quad -\infty < y < \infty, \quad -\infty < z < \infty$$

and the initial concentration is a Dirac impulse, i.e.

$$C(0.\boldsymbol{x}) = \delta(\boldsymbol{x} - \mathbf{0}) = \delta(x-0)\delta(y-0)\delta(z-0)$$

It is useful to introduce dimensionless variables. The time and space coordinate are transformed as follows:

$$t' = \frac{t}{T}, \quad x' = \frac{x}{\mathcal{L}_h}, \quad y' = \frac{y}{\mathcal{L}_h}, \quad z' = \frac{z}{\mathcal{L}_\nu} \tag{4.25}$$

where T, $\mathcal{L}_h$ and $\mathcal{L}_\nu$ denote the appropriate timescale, horizontal length scale and vertical length scale. It is convenient to define the latter in such a way that

$$T = \frac{\mathcal{L}_h^2}{k^i} = \frac{\mathcal{L}_\nu^2}{k^d} \tag{4.26}$$

The ratio of the vertical length scale to the horizontal one, i.e. the aspect ratio, is

$$\alpha = \frac{\mathcal{L}_\nu}{\mathcal{L}_h} \tag{4.27}$$

The concentration is scaled as follows:

$$C' = \frac{C}{1/(\mathcal{L}_h^2 \mathcal{L}_\nu)} \tag{4.28}$$

The equation to be solved now reads

$$\begin{aligned} \frac{\partial C'}{\partial t'} = & \frac{\partial}{\partial x'}\left(k'_{xx}\frac{\partial C'}{\partial x'} + k'_{xy}\frac{\partial C'}{\partial y'} + k'_{xz}\frac{\partial C'}{\partial z'}\right) + \\ & \frac{\partial}{\partial y'}\left(k'_{yx}\frac{\partial C'}{\partial x'} + k'_{yy}\frac{\partial C'}{\partial y'} + k'_{yz}\frac{\partial C'}{\partial z'}\right) + \\ & \frac{\partial}{\partial z'}\left(k'_{zx}\frac{\partial C'}{\partial x'} + k'_{zy}\frac{\partial C'}{\partial y'} + k'_{zz}\frac{\partial C'}{\partial z'}\right) \end{aligned} \tag{4.29}$$

with

$$\begin{aligned} k'_{xx} &= \frac{k_{xx}}{\mathcal{L}_h^2/T} = \cos^2\theta + \sin^2\theta(\sin^2\gamma + \alpha^2\cos^2\gamma) \\ k'_{xy} = k'_{yx} &= \frac{k_{xy}}{\mathcal{L}_h^2/T} = -\sin^2\theta\cos\gamma\sin\gamma(1-\alpha^2) \\ k'_{xz} = K'_{zx} &= \frac{k_{xz}}{\mathcal{L}_h\mathcal{L}_\nu/T} = \cos\theta\sin\theta\cos\gamma(\alpha^{-1}-\alpha) \\ k'_{yy} &= \frac{k_{yy}}{\mathcal{L}_h^2/T} = \cos^2\theta + \sin^2\theta(\cos^2\gamma + \alpha^2\sin^2\gamma) \\ k'_{yz} = k'_{zy} &= \frac{k_{yz}}{\mathcal{L}_h\mathcal{L}_\nu/T} = \cos\theta\sin\theta\sin\gamma(\alpha^{-1}-\alpha) \\ k'_{zz} &= \frac{k_{zz}}{\mathcal{L}_\nu^2/T} = \cos^2\theta + \alpha^{-2}\sin^2\theta \end{aligned} \tag{4.30}$$

The initial condition is

$$C'(0, \boldsymbol{x}') = \delta(\boldsymbol{x}' - \mathbf{0}) = \delta(x'-0)\delta(y'-0)\delta(z'-0) \tag{4.31}$$

From here on, only dimensionless variables quantities will be dealt with. So, for the sake of simplicity, the primes will be dropped. It is easily can be shown that

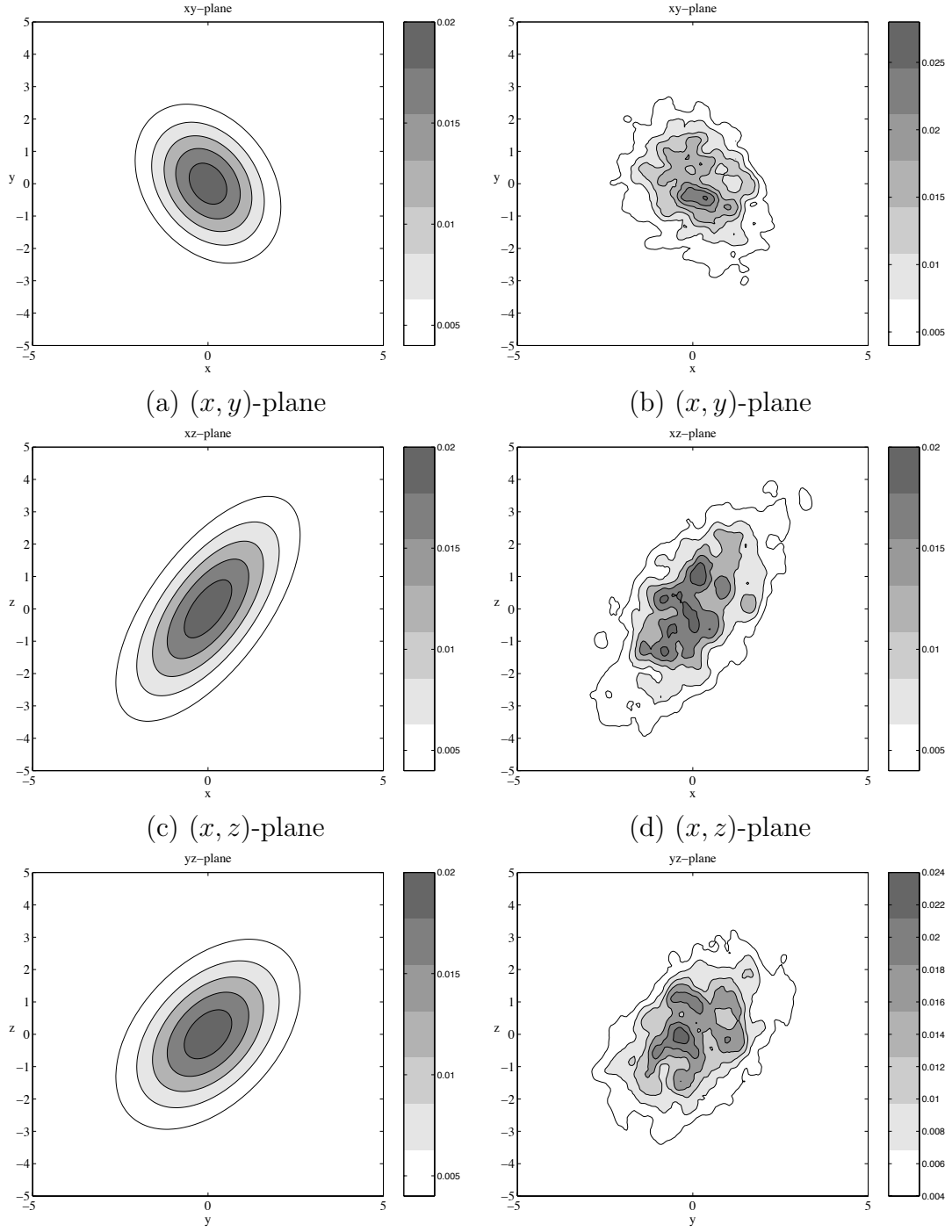
$$det(\boldsymbol{k}) = 1$$

and the inverse matrix

$$\boldsymbol{k}^{-1} = \begin{pmatrix} \tilde{k}_{xx} & \tilde{k}_{xy} & \tilde{k}_{xz} \\ \tilde{k}_{yx} & \tilde{k}_{yy} & \tilde{k}_{yz} \\ \tilde{k}_{zx} & \tilde{k}_{zy} & \tilde{k}_{zz} \end{pmatrix}$$

where

$$\begin{aligned} \tilde{k}_{xx} &= \cos^2\theta + \sin^2\theta(\sin^2\gamma + \alpha^{-2}\cos^2\gamma) \\ \tilde{k}_{xy} &= -\sin^2\theta\cos\gamma\sin\gamma(1-\alpha^{-2}) \\ \tilde{k}_{xz} = \tilde{k}_{zx} &= -\cos\theta\sin\theta\cos\gamma(\alpha^{-1}-\alpha) \\ \tilde{k}_{yy} &= \cos^2\theta + \sin^2\theta(\cos^2\gamma + \alpha^{-2}\sin^2\gamma) \\ \tilde{k}_{yz} = \tilde{k}_{zy} &= -\cos\theta\sin\theta\sin\gamma(\alpha^{-1}-\alpha) \\ \tilde{k}_{zz} &= \cos^2\theta + \alpha^2\sin^2\theta \end{aligned} \tag{4.32}$$

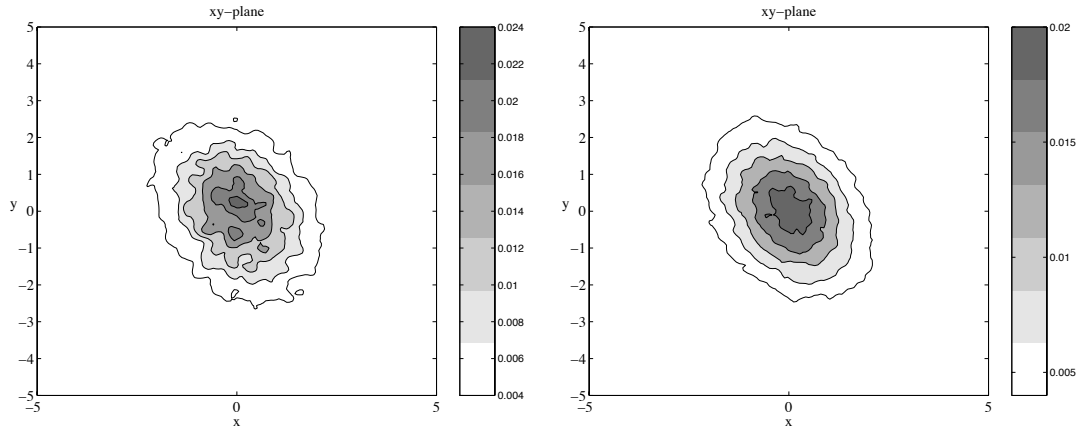


(a) (x, y)-plane (b) (x, y)-plane

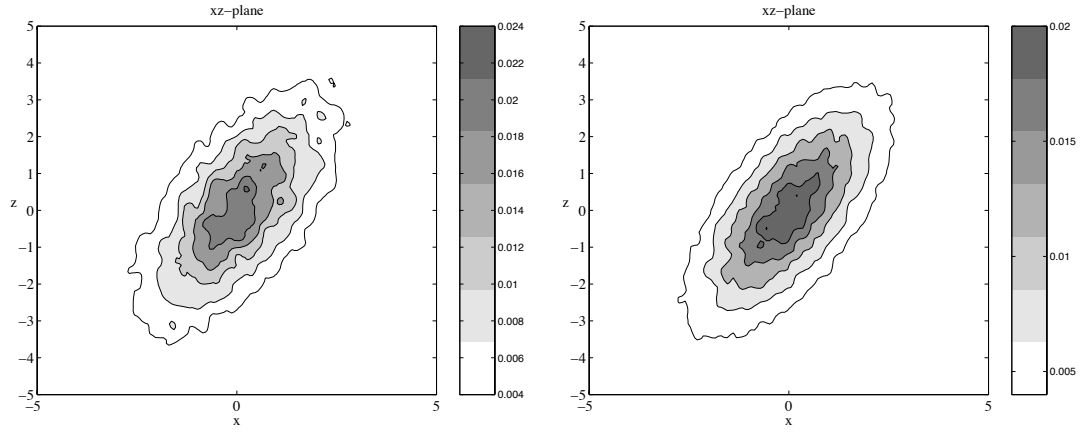
(c) (x, z)-plane (d) (x, z)-plane

(e) (y, z)-plane (f) (y, z)-plane

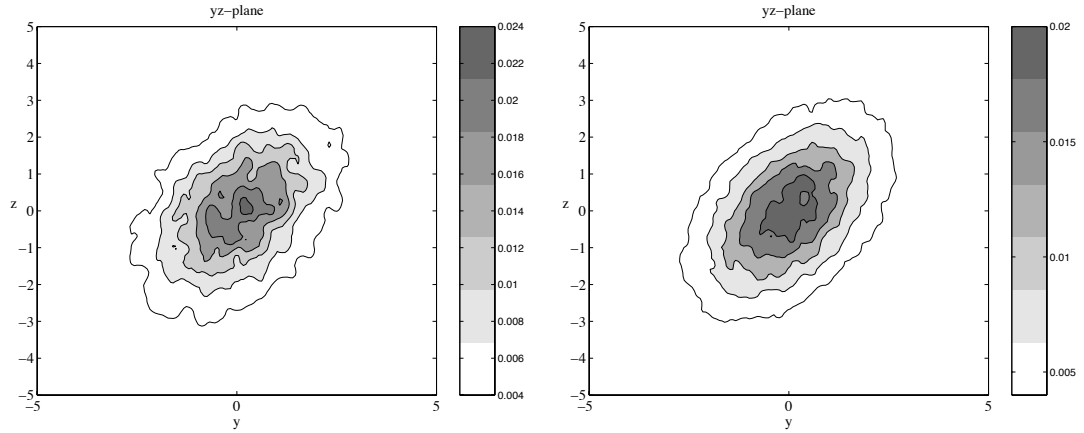
Figure 4.6: The analytical (a),(c),(e) and numerical solutions (for (b),(d),(f) 10^4 particles) for three-dimensional linear iso- and diapycanl advection-diffusion problem. The values of parameters $\alpha = 10^{-3}$, $\gamma = \pi/6$, $\theta = 10^{-3}$, $T = 1$.



(a) (x, y)-plane (b) (x, y)-plane

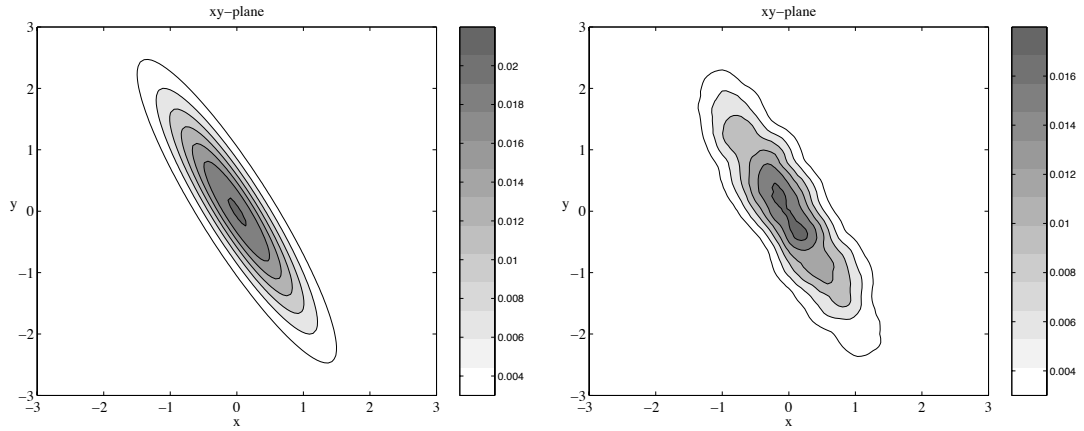


(c) (x, z)-plane (d) (x, z)-plane



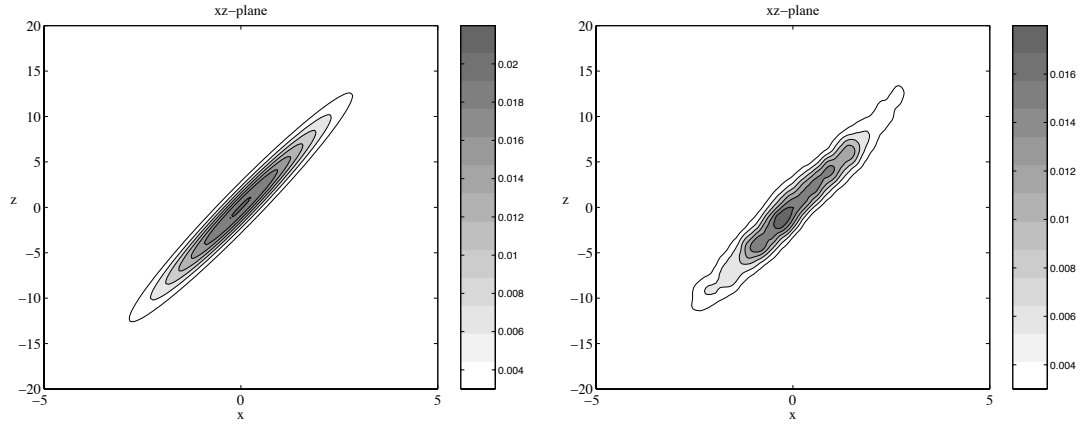
(e) (y, z)-plane (f) (y, z)-plane

Figure 4.7: The numerical solutions for $N = 10^5$ particles (a), (c), (e) and for $N = 10^6$ particles (b), (c), (e) of three-dimensional linear iso- and diapycanl advection-diffusion problem. The values of parameters $\alpha = 10^{-3}$, $\gamma = \pi/6$, $\theta = 10^{-3}$, $T = 1$.



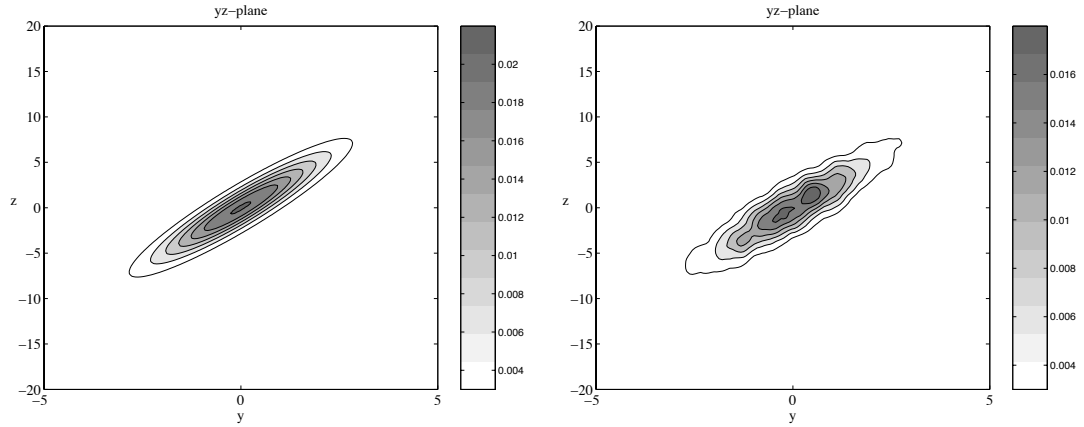
(a) (x, y)-plane

(b) (x, y)-plane



(c) (x, z)-plane

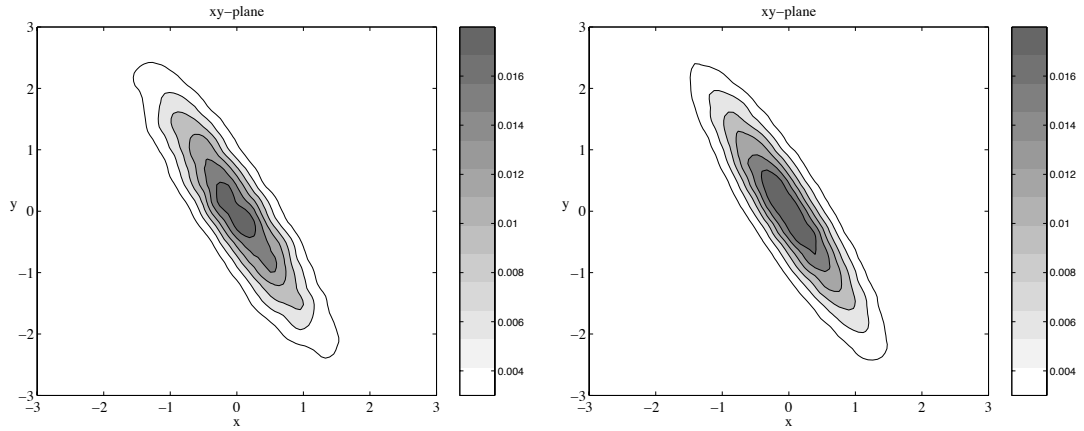
(d) (x, z)-plane



(e) (y, z)-plane

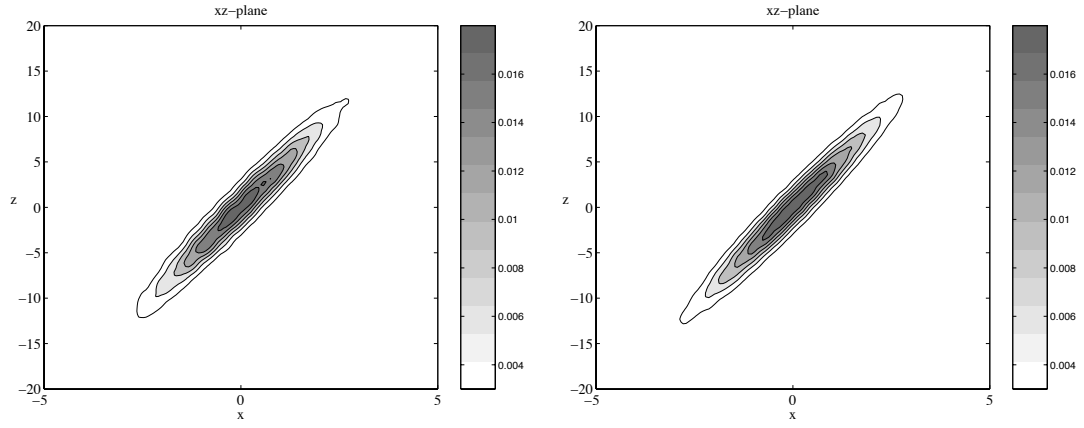
(f) (y, z)-plane

Figure 4.8: The analytical (a),(c),(e) and numerical solutions (for (b),(d),(f) 10^4 particles) for three-dimensional linear iso- and diapycanl advection-diffusion problem. The values of parameters $\alpha = 10^{-3}$, $\gamma = \pi/6$, $\theta = 5 \cdot 10^{-3}$, $T = 1$.



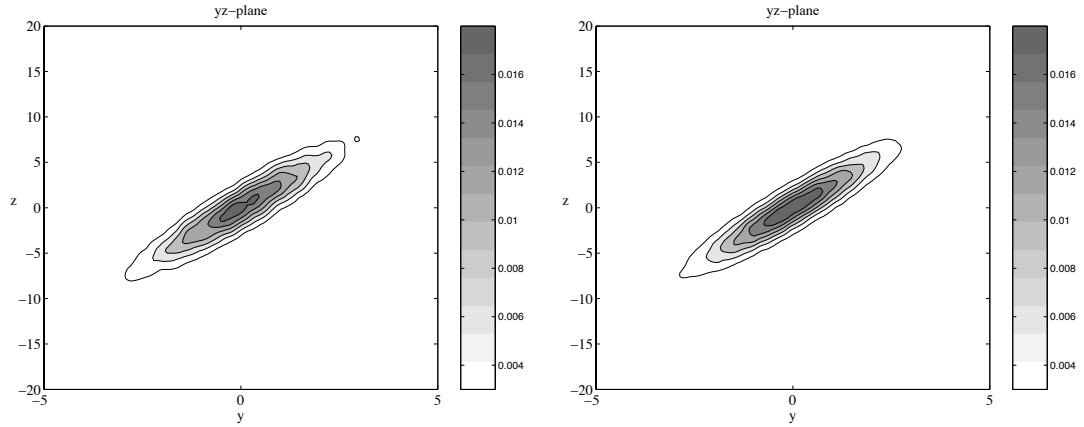
(a) (x, y)-plane

(b) (x, y)-plane



(c) (x, z)-plane

(d) (x, z)-plane



(e) (y, z)-plane

(f) (y, z)-plane

Figure 4.9: The numerical solutions for $N = 10^5$ particles (a), (c), (e) and for $N = 10^6$ particles (b), (c), (e) of three-dimensional linear iso- and diapycanl advection-diffusion problem. The values of parameters $\alpha = 10^{-3}$, $\gamma = \pi/6$, $\theta = 5 \cdot 10^{-3}$, $T = 1$.

Then, the solution may be rewritten as follows:

$$C(t,x,y,z) = \frac{1}{(4\pi t)^{3/2}} \exp\{-\frac{1}{4t} \times [(z\cos\theta - \alpha^{-1}(y\sin\theta\sin\gamma + x\sin\theta\cos\gamma))^2 + (z\alpha\sin\theta + x\cos\theta\cos\gamma + y\cos\theta\sin\gamma)^2 + (x\sin\gamma - y\cos\gamma)^2]\} \tag{4.33}$$

As it was mentioned in the previous section we can use the following values of parameters

$$\theta \approx 10^{-3} \approx \alpha \tag{4.34}$$

The random walk model (4.5) has the form

$$\begin{aligned} dX(t) &= \sqrt{2}V_{xx}dW_1(t) \\ dY(t) &= \sqrt{2}V_{yx}dW_1(t) + \sqrt{2}V_{yy}dW_2(t) \\ dZ(t) &= \sqrt{2}V_{zx}dW_1(t) + \sqrt{2}V_{zy}dW_2(t) + \sqrt{2}V_{zz}dW_3(t) \\ X(0) &= Y(0) = Z(0) = 0 \end{aligned} \tag{4.35}$$

where $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{V}^T$ and the matrix $\boldsymbol{V}$ has the following form

$$\boldsymbol{V} = \begin{pmatrix} V_{xx} & 0 & 0 \\ V_{yx} & V_{yy} & 0 \\ V_{zx} & V_{zy} & V_{zz} \end{pmatrix}$$

Here

$$\begin{aligned} V_{xx} &= \sqrt{k_{xx}} \\ V_{yx} &= k_{xy}/\sqrt{k_{xx}} \\ V_{zx} &= k_{xz}/\sqrt{k_{xx}} \\ V_{yy} &= \sqrt{\frac{k_{yy}k_{xx} - k_{xy}^2}{k_{xx}}} \\ V_{zy} &= \frac{k_{yz}k_{xx} - k_{xy}k_{xz}}{\sqrt{k_{xx}(k_{xx}k_{yy} - k_{xy}^2)}} \\ V_{zz} &= \frac{k_{zz}k_{yy}k_{xx} + 2k_{xy}k_{xz}k_{yz} - k_{xz}^2k_{yy} - k_{yz}^2k_{xx} - k_{xy}^2k_{zz}}{\sqrt{k_{yy}k_{xx} - k_{xy}^2}} \end{aligned}$$

To calculate the concentration the extrapolation method (2.29) has been used with the following parameters

$$\alpha = 10^{-3}, \quad \gamma = \pi/6, \quad T = 1$$

Two cases were considered with two different values of angle $\theta = 10^{-3}$ (the figures 4.6,4.7) and $\theta = 5 \cdot 10^{-3}$ (the figures 4.8, 4.9). From the figures it can be easily seen that the random walk model gives a good approximation of the exact solution in both cases.

In Table 4.2 the results of estimation of the concentration $C(t, \boldsymbol{x})$ ($\theta = 10^{-3}$) at given point $\boldsymbol{x} = (-1, 1.5, 0)$ are shown. For a fixed value N of particles we calculate the statistical error using the formula (4.20).

The theoretical rate of the convergence of statistical error $\epsilon_{\text{statistical}}$ is $\mathcal{O}(N^{-2/7})$ (for the three-dimensional model). From Figure 4.10 it can be seen that the statistical error agrees with the theoretical one as it was for the two-dimensional model.

N	Average value	Statistical error	Exact value
10^3	0.0103	0.0044	0.0099
$5 \cdot 10^3$	0.0091	0.0028	0.0099
10^4	0.0102	0.0023	0.0099
$5 \cdot 10^4$	0.0100	0.0011	0.0099
10^5	0.0098	0.0011	0.0099
$5 \cdot 10^5$	0.0099	0.0006	0.0099
10^6	0.0097	0.0005	0.0099

Table 4.2: The concentration at the location $\boldsymbol{x} = (-1, 1.5, 0)$

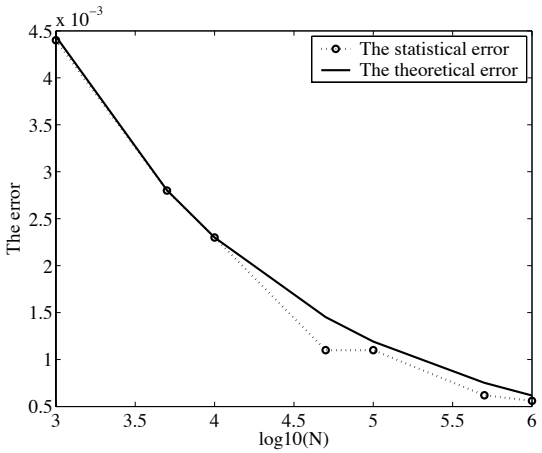


Figure 4.10: The comparison of the statistical error with the theoretical in case of three-dimensional advection-diffusion problem

4.5 Settling and diffusion model

In the test case considered in the previous section the area of interest is assumed infinite. However, in the most applications we have to deal with the bounded area. That is the reason why it is important to test the random walk scheme for the model with space-varying diffusivity in the presence of boundaries (see Figure 4.11).This model has been introduced and investigated in [26].

Suppose that the settling velocity w is zero. In this case the concentration of constituent under consideration can be found from the following differential equation ($0 \leq t$,

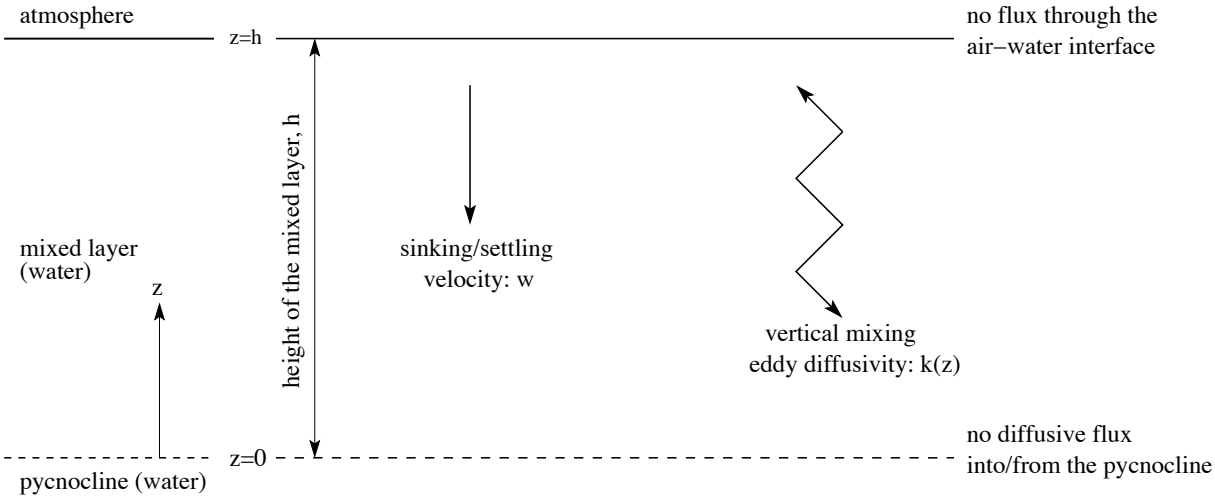


Figure 4.11: Sinking-diffusion model: illustration of its geometry, parameters and boundary conditions. Source: [26]

$0 \leq z \leq h$)

$$\begin{aligned} \frac{\partial C}{\partial t} &= \frac{\partial}{\partial z}\left(k\frac{\partial C}{\partial z}\right) \\ \left[k\frac{\partial C}{\partial z}\right]_{z=0,h} &= 0 \\ C(0,z) &= \delta(z-z_0) \end{aligned} \tag{4.36}$$

where $k(z)$ is the eddy diffusivity. The latter is positive and non-zero in the interval $0 < z < h$.

For the sake of generality, it is convenient to reformulate the problem above using dimensionless variables. The latter are defined to be

$$t' = \frac{t}{h^2/k}, \quad z' = \frac{z}{h}, \quad k' = \frac{k}{\overline{k}} \tag{4.37}$$

where $\overline{k}$ denotes the average over the domain of interest of the eddy diffusivity i.e.

$$\overline{k} = h^{-1}\int_0^h k(z)dz \tag{4.38}$$

From here on, only dimensionless quantities will be used. This is why we will drop the primes. So, using dimensionless variables, the domain of interest and the problem to be solved may be rewritten as follows:

$$\begin{aligned} &0 \le t, \quad 0 \le z \le 1 \\ &\frac{\partial c}{\partial t} = \frac{\partial}{\partial z}\left(k\frac{\partial C}{\partial z}\right) \\ &\left[k\frac{\partial C}{\partial z}\right]_{z=0,1} = 0 \\ &C(0,z) = \delta(z - z_0) \end{aligned} \tag{4.39}$$

where the domain-averaged values of the eddy diffusivity must be equal to unity, i.e.

$$\overline{k} = \frac{1}{h}\int_0^h k(z)dz = 1$$

As an example of the diffusivity profile the following function may be chosen

$$k(z) = 6z(1-z) \tag{4.40}$$

This choice is consistent with the diffusion processes in the upper mixed layer. The diffusivity profile $k(z)$ should lead to zero in the neigboorhood of the bottom of the mixed layer and the maximum of the diffusivity should not occur at surface [26]. It is clear that the diffusivity function (4.40) satisfies these conditions. It can be shown that, in this case, the analytical solution can be obtained in the form

$$C(t,z) = 1 + \sum_{n=1}^{\infty}(2n+1)P_n(2z-1)P_n(2z_0-1)e^{-6n(n+1)t} \tag{4.41}$$

where $P_n(z)$ denotes the nth order Legendre polynomial. Figure 4.12(a) presents the analytical solution for different moments of time obtained for $z_0 = 1/2$.

The random walk model (4.5) in this case has the following form

$$\begin{aligned} &dZ = \frac{dk}{dz}dt + \sqrt{2k(z)}dW(t) \\ &Z(0) = z_0 \end{aligned} \tag{4.42}$$

For the numerical solution (see Figures 4.12 and 4.13) we apply the Euler scheme with timestep $\Delta t = 3\cdot 10^{-5}$ (2.23). For this example the kernel estimator (2.51) with the Gaussian kernel (2.44) was compared with the box counting method. Figure 4.12 illustrates the results of the kernel estimator with the following choice of the bandwidth

$$\begin{aligned} &\lambda = \mathcal{K}N^{-\frac{1}{5}}, \\ &\mathcal{K} = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\left(\overline{Z}^{(j)}(t) - \frac{1}{N}\sum_{l=1}^{N}\overline{Z}^{(l)}(t)\right)^2} \end{aligned} \tag{4.43}$$

where $\mathcal{K}$ is the standard deviation of the sample $\overline{Z}^{(i)}$, $i = 1, \ldots, N$.

The numerical solution shown on Figure 4.13 was obtained as a result of box counting approach. The domain was divided into 50 intervals, so the size of each interval is 0.02. It is clear that the solution obtained by using the kernel estimator is more accurate than by using the box counting approach by at least a order of magnitude.

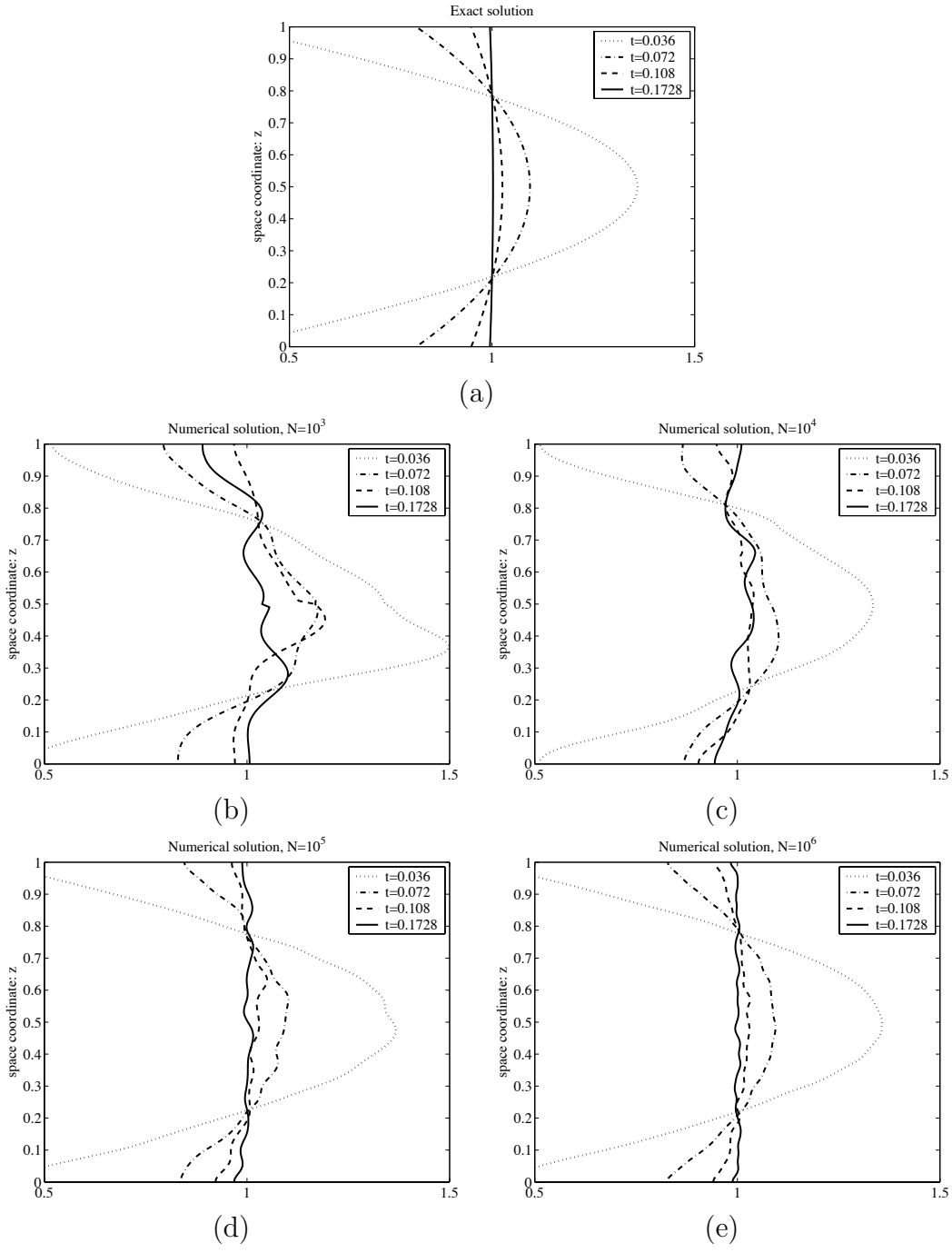


Figure 4.12: (a) The concentration profile for different moments of the simulation, obtained by using the explicit expression (4.41). The numerical approximation of the concentration profile for (b) $N = 10^3$, (c) $N = 10^4$, (d) $N = 10^5$ and (e) $N = 10^6$. The kernel estimator with the Gaussian function was used.

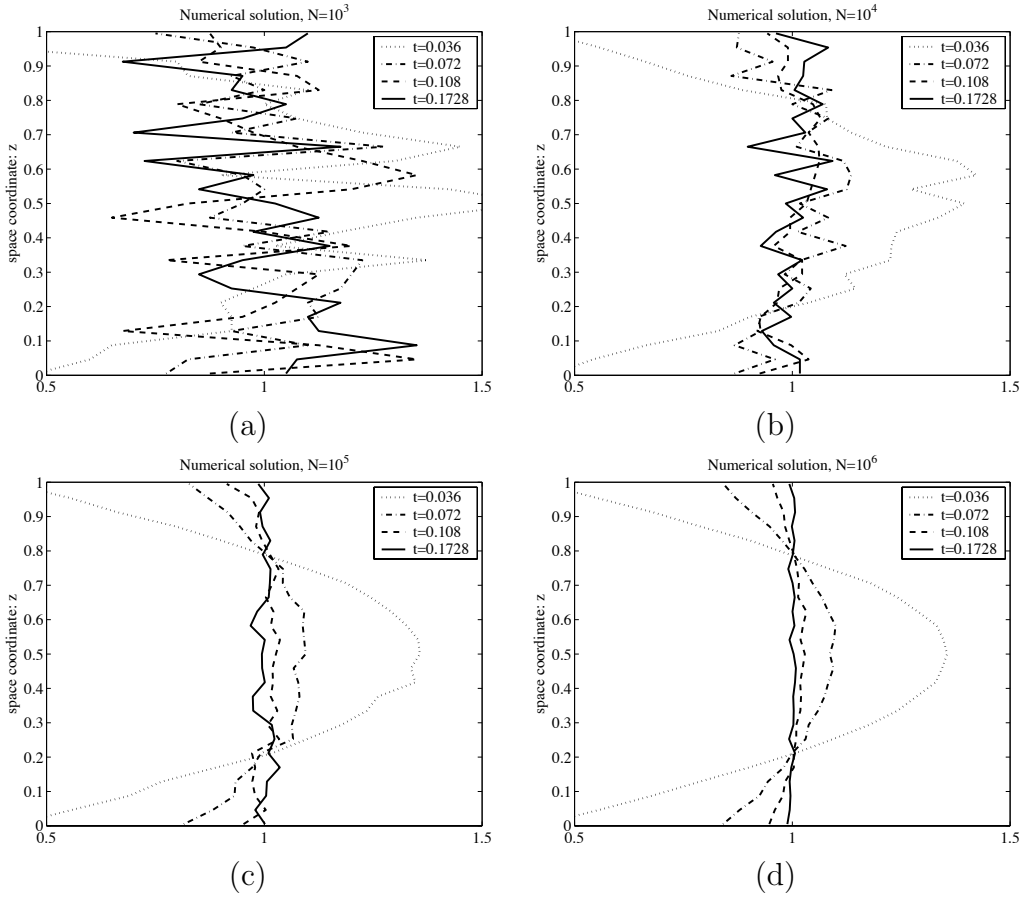


Figure 4.13: The numerical approximation of the concentration profile for the eddy diffusivity function (4.41) for (a) $N = 10^3$, (b) $N = 10^4$, (c) $N = 10^5$ and (d) $N = 10^6$. The box estimator was used

We calculated the concentration at the middle of the boundary layer and at the boundary (using the kernel estimator) and repeated each simulation 30 times in order to find the statistical error. The results for different moments of time and different number of particles are presented in Table 4.5. In Tables 4.4 and 4.5 one can see the concentration calculated at the boundary $z = 0$ using the reflection method (Table 4.4) and the boundary kernel (2.55) (Table 4.5).

In Section 2.4.3 it was mentioned that the reflection of the kernel function near the boundary allows to achieve the consistency with the true value of the concentration, but still has a large bias of order $\mathcal{O}(\lambda)$. In Section 2.4.3 it was shown that the bias is proportional to the concentration gradient. When the concentration gradient at the boundary is high (for instance, for $t = 0.018$ or $t = 0.036$) one can see from Table 4.5 that the estimated concentration is larger than the true value of the concentration. For larger N, the optimal bandwidth $\lambda = \mathcal{O}(N^{-1/5})$ becomes smaller and the numerical concentration converges to the true value.

time t	$N = 10^3$	$N = 10^4$	$N = 10^5$	exact
0.0036	3.762 ± 0.192	3.856 ± 0.053	3.873 ± 0.011	3.881
0.018	1.812 ± 0.087	1.799 ± 0.031	1.811 ± 0.013	1.814
0.036	1.326 ± 0.074	1.356 ± 0.026	1.359 ± 0.010	1.359
0.054	1.163 ± 0.055	1.174 ± 0.029	1.180 ± 0.012	1.081
0.072	1.079 ± 0.058	1.089 ± 0.024	1.092 ± 0.009	1.094
0.108	1.025 ± 0.060	1.020 ± 0.021	1.023 ± 0.011	1.023
0.172	0.990 ± 0.063	1.003 ± 0.025	1.000 ± 0.008	1.003

Table 4.3: The exact and numerical concentration and the statistical error at the location $z = 0.5$ (middle of the boundary layer). All variables are dimensionless

time t	$N = 10^3$	$N = 10^4$	$N = 10^5$	exact
0.018	0.134 ± 0.034	0.097 ± 0.014	0.074 ± 0.004	0.040
0.036	0.502 ± 0.057	0.465 ± 0.017	0.425 ± 0.009	0.360
0.054	0.735 ± 0.068	0.712 ± 0.032	0.689 ± 0.011	0.647
0.072	0.883 ± 0.053	0.844 ± 0.030	0.839 ± 0.014	0.813
0.108	0.970 ± 0.078	0.953 ± 0.029	0.950 ± 0.013	0.949
0.172	1.003 ± 0.098	0.995 ± 0.036	0.992 ± 0.013	0.995

Table 4.4: The exact and numerical concentration and the statistical error at the location $z = 0.0$ (the boundary). All variables are dimensionless. The kernel estimator was transformed near boundary using reflection technique based (see Section 2.4.3).

time t	$N = 10^3$	$N = 10^4$	$N = 10^5$	exact
0.018	0.072 ± 0.095	0.031 ± 0.028	0.040 ± 0.012	0.040
0.036	0.348 ± 0.204	0.372 ± 0.065	0.366 ± 0.027	0.360
0.054	0.695 ± 0.192	0.641 ± 0.102	0.640 ± 0.027	0.647
0.072	0.844 ± 0.247	0.826 ± 0.136	0.802 ± 0.042	0.813
0.108	0.980 ± 0.270	0.993 ± 0.096	0.954 ± 0.055	0.949
0.172	1.065 ± 0.307	0.983 ± 0.100	0.986 ± 0.049	0.995

Table 4.5: The exact and numerical concentration and the statistical error at the location $z = 0.0$ (the boundary). All variables are dimensionless. The boundary kernel (2.55) was used.

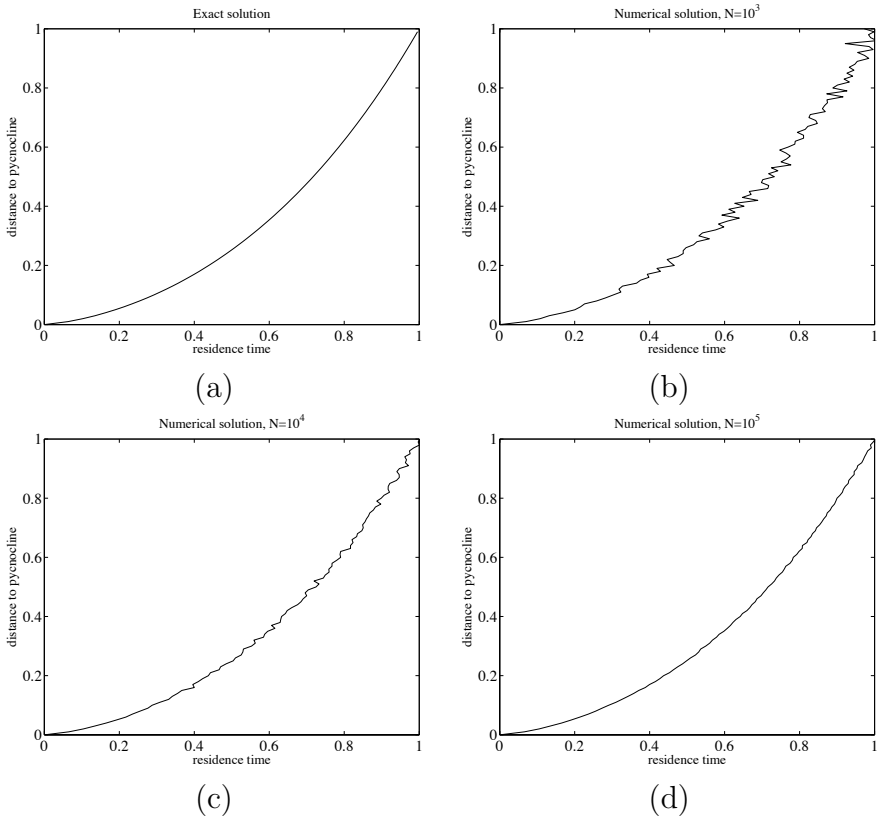


Figure 4.14: The residence time for diffusivity profile (4.40): exact (a) and numerical approximation for (b) $N = 10^3$ (c) $N = 10^4$ (d) $N = 10^5$ particles

4.5.1 The residence time

If the settling velocity w is nonzero, the analytical solution cannot be obtained. However, the exact solution for the adjoint problem of finding the residence time $\theta(z_0)$ is known ([26, 27]). To obtain the residence time $\theta(z_0)$, a unit amount of tracer is released at the initial time at a distance z_0 to the pycnocline. In other words, the initial condition reads

$$C(0, z) = \delta(z - z_0) \tag{4.44}$$

As it has been done in the previous examples, it is convenient to introduce dimensionless variables:

$$t' = \frac{t}{h/w}, \quad (z', z_0') = \frac{(z, z_0)}{h}, \quad k' = \frac{k}{\overline{k}}, \quad \theta' = \frac{\theta}{h/w} \tag{4.45}$$

Finally, the Peclet number is defined to be

$$Pe = \frac{wh}{\overline{k}} \tag{4.46}$$

and because only the dimensionless variables will be used from now on we will drop all primes. Accordingly, the domain of the interest is now defined as $0 \leq t$ and $0 \leq z \leq 1$

In [26] it is shown that the residence time corresponding to the parabolic profile (4.40) (Figure 4.14(a)) is

$$\theta(z) = z + \left(\frac{z}{1-z}\right)^{\mu} B_{1-z}(1+\mu, 1-\mu) \tag{4.47}$$

where $B_{1-z}(1+\mu, 1-\mu)$ is a generalized incomplete beta function, i.e.

$$B_{1-z}(1+\mu, 1-\mu) = \int_0^{1-z} \sigma^{\mu}(1-\sigma)^{-\mu} d\sigma \tag{4.48}$$

and $\mu = Pe/6$.

In this test problem the random walk algorithm is given by

$$\begin{aligned} dZ(t) &= \left(w + \frac{dk}{dz}\right) dt + \sqrt{2k(z)} dW(t) \\ Z(0) &= z_0. \end{aligned} \tag{4.49}$$

We release particles at the position z_0 and model their movement using a Euler scheme similar to (7.36). For each realization of the particle track we calculate the residence time, i.e. how much time the particle needs to leave the domain. By averaging the results, the average residence time can be obtained. The numerical results for $N = 10^3$, $N = 10^4$ and $N = 10^5$ are shown on Figure 4.14 (b)-(d). For this numerical simulation the following parameters were used

Timestep	Δt	10^{-4}
Peclet number	Pe	5

Comparing the exact solution (Figure 4.14 (a)) with the numerical solutions (Figure 4.14 (b)-(d)) it is clear that already the numerical solution for $N = 10^3$ particles provides a good approximation of the exact solution.

4.6 Conclusion

In this chapter the random walk model for the simulation of diffusion processes with space-varying diffusivity is introduced and analyzed. The kernel estimator is applied instead of the traditional box counting method. It is shown that the kernel estimator allows to reduce the number of particles by one order of magnitude as compared with the box counting method. This Lagrangian model is applied to several test problems and results show that this random walk model may be a good alternative to commonly-used Eulerian models, for instance, in case of space-varying non-diagonal diffusivity matrix.

Chapter 5

Simulation of the transport of particles in coastal waters using forward and reverse time diffusion

5.1 Introduction

In Section 2.5.1 it was shown that the efficiency of Monte Carlo methods can be improved by variance reduction. Here an approximation of the Kolmogorov backwards equation is required. In some simple applications analytical approximations of this partial differential equation can be used. In general, however, a numerical solution is required. For high dimensional systems this may become very time consuming [103]. Recently, G. Milstein, J.G.M. Schoenmakers and V. Spokoiny [85] introduced the concept of reverse time diffusion. The classical Monte Carlo estimator is based on forward realizations of the original stochastic model. In [85] a reverse system is derived from the original model and it is shown that the classical Monte Carlo estimator can also be based on realizations of this reverse system. For many problems it is more efficient to use realizations of the reverse system instead of the original forward model. The most efficient implementation is, however, obtained if the forward realizations and the reverse system realizations are combined. This is called the forward reverse estimator. In this chapter, the forward reverse method is applied to estimate the concentration of a pollutant in coastal waters at a number of given critical locations, given the fact that a calamity has taken place at another given location. In Section 5.2.1 the concept of the reverse diffusion is introduced. The forward reverse estimator based on both forward and reverse systems is considered in Section 5.2.2. In Section 5.3, we illustrate the new method with a relatively simple example. In next Sections 5.4-5.5 the new method is applied to estimate the concentration of a pollutant in Dutch coastal waters. The reverse-time diffusion may be effectively used

This chapter is based on the papers:

Spivakovskaya D., Heemink A.W., Milstein G. and Schoenmakers J.G.M. 2005 'Simulation of the transport of particles in coastal waters using forward and reverse time diffusion', Advances in Water Resources, 28, pp.927-938.

Spivakovskaya D., Heemink A.W., Schoenmakers J.G.M and Milstein G. 2004 'Stochastic modeling of transport in coastal waters using forward and reverse time diffusion', in 'Proceedings of Computational Methods in Water Resources XVI', C.T. Miller, M.W. Farthing, W.G. Gray, and G.F. Pinder (Eds.)

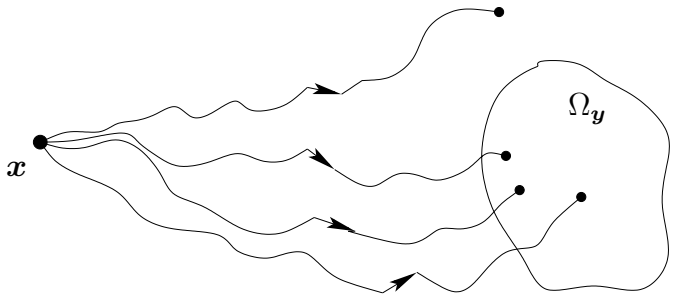


Figure 5.1: The forward approach

to reduce the possible risk of calamities as it is shown in Section 5.6. Some aspects of the implementation of the forward-reverse estimator are discussed in Section 5.7. Finally, a conclusion is drawn in Section 5.8.

5.2 The forward-reverse estimator

5.2.1 Forward and reverse probabilistic representation

In Section 2.4.2 we considered the kernel estimator (2.51) for calculation the density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$. This formula can be applied for a more general case to estimate the functional

$$w(t,T) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(\boldsymbol{x})p(t,\boldsymbol{x};T,\boldsymbol{y})f(\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y}, \quad T \geq t \tag{5.1}$$

Here $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ is the transition density function corresponding to the solution of stochastic differential equation (2.11), g is a fixed initial probability density and f is a continuous function. The problem of evaluating (5.1) is encountered in many applications, for example, in the estimation of the concentration $w(t, T)$ at a future time T at a location around some $\boldsymbol{y}$, described by the kernel $f(\boldsymbol{y})$, given a release of pollution at time t with the probability density $g(\boldsymbol{x})$.

Clearly, (5.1) has a probabilistic representation [2, 91]

$$w(t,T) = Ef(\boldsymbol{X}_{t,\boldsymbol{\zeta}}(T)) \tag{5.2}$$

where $\boldsymbol{\zeta}$ is a random variable in R^d with density $g(\boldsymbol{x})$, which is independent of the Wiener process $\boldsymbol{W}(s)$, $t \leq s \leq T$ that is used in the stochastic differential equation (2.51).

The functional $w(t, T)$ can be approximated straightforwardly by the Monte Carlo estimator,

$$w(t,T) \approx \frac{1}{N}\sum_{n=1}^{N} f(\bar{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{\zeta}^{(n)}}(T)) \tag{5.3}$$

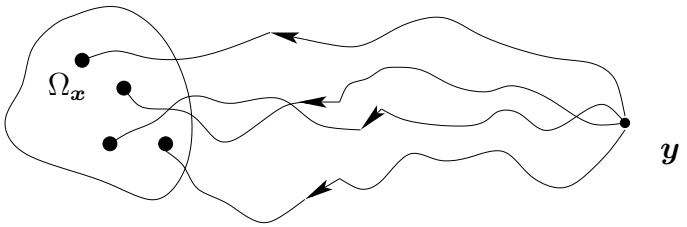


Figure 5.2: The reverse approach

where $\boldsymbol{\zeta}^{(n)}, \quad n = 1, \ldots, N$, are independent identically distributed (i.i.d.) copies of $\boldsymbol{\zeta}$, and $\bar{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{\zeta}^{(n)}}(T)$ are independent approximations of the solution $\boldsymbol{X}(T)$ of the stochastic differential equation (2.11), starting at t in $\boldsymbol{\zeta}^{(n)}$. Because for the estimators (2.51) and (5.3) we used the forward trajectories of the solution of the stochastic differential equation (2.11), these estimators are usually referred as forward estimators.

The forward approach is attractive, for instance, for problems where the calamity is sharply located at $\boldsymbol{x}$, and where the critical area $\Omega_{\boldsymbol{y}}$ (see Figure 5.1) is rather large. In such a case the function $f(\boldsymbol{y})$ can be chosen as

$$f(\boldsymbol{v}) = \begin{cases} 1, & \boldsymbol{v} \in \Omega_{\boldsymbol{y}} \\ 0, & \boldsymbol{v} \notin \Omega_{\boldsymbol{y}} \end{cases}$$

g is a delta function concentrated at the point $\boldsymbol{x}$, and the integral (5.1) may be considered for fixed T as a function of t and $\boldsymbol{x}$,

$$w(t, \boldsymbol{x}; T) = Ef(\boldsymbol{X}_{t,\boldsymbol{x}}(T)) = \int\limits_{R^d} p(t, \boldsymbol{x}; T, \boldsymbol{y}) f(\boldsymbol{y}) d\boldsymbol{y} \tag{5.4}$$

For fixed T, $w(t, \boldsymbol{x}; T)$ satisfies the backward Kolmogorov equation (2.42). Thus, the forward approach is related to the backward Kolmogorov equation.

Now, suppose that it is required to evaluate the effect in a critical location $\boldsymbol{y}$ of a calamity that may occur in a large initial area $\Omega_{\boldsymbol{x}}$ (see Figure 5.2). So, f is a delta function concentrated at $\boldsymbol{y}$, and g may be be chosen as

$$g(u) = \begin{cases} 1, & \boldsymbol{u} \in \Omega_{\boldsymbol{x}} \\ 0, & \boldsymbol{u} \notin \Omega_{\boldsymbol{x}} \end{cases}$$

In this case the forward approach breaks down since the variance of the estimator (5.3) is infinity. As a remedy, we may apply the reverse approach explained below.

For a fixed function g (not necessarily a density), a fixed initial time t, and a delta function f concentrated at $\boldsymbol{y}$, the integral (5.1) may be considered as a function of s and $\boldsymbol{y}$,

$$w(t; s, \boldsymbol{y}) = \int\limits_{\mathbb{R}^d} g(\boldsymbol{x}) p(t, \boldsymbol{x}; s, \boldsymbol{y}) d\boldsymbol{x} \tag{5.5}$$

In particular, the transition density function $p(t, \boldsymbol{x}; s, \boldsymbol{y})$ satisfies the forward Kolmogorov equation (2.39). Then, multiplying this equation by $g(\boldsymbol{x})$ and integrating with respect to $\boldsymbol{x}$ yields the following Cauchy problem for the function $w(t; s, \boldsymbol{y})$

$$\begin{aligned} &\frac{\partial w}{\partial s} + \sum_{i=1}^{d} \frac{\partial}{y_i}(a_i(s, \boldsymbol{y})w) - \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial y_i \partial y_j}(b_{ij}(s, \boldsymbol{y})w) = 0 \\ &w(t; t, \boldsymbol{y}) = g(\boldsymbol{y}) \end{aligned} \tag{5.6}$$

In [85] it is shown that the solution of the problem (5.6) has a probabilistic representation of the form

$$w(t; s, \boldsymbol{y}) = Eg(\boldsymbol{Y}_{t,\boldsymbol{y}}(s))\mathcal{Y}_{t,\boldsymbol{y}}(s), \quad t \le s \le T \tag{5.7}$$

where the process $(\boldsymbol{Y}_{t,\boldsymbol{y}}, \mathcal{Y}_{t,\boldsymbol{y}})$ in $\mathbb{R}^{d+1}$ is the solution of a so called reverse(-time) stochastic differential equation system connected with (2.11) which is described in details in the following section.

In case the function g tends to a delta function, similar as the forward approach the reverse approach breaks down also since the variance of the random variable $g(\boldsymbol{Y})\mathcal{Y}$ in (5.7) tends to infinity as well.

In general, in case the initial location and the critical location are both small areas represented by delta functions, the combined use of forward and reverse probabilistic representations developed in [85] for estimating the transition density $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ provides the most efficient approach.

5.2.2 The forward-reverse estimator

The probabilistic representation (5.7) is based on a so called reverse-time diffusion and has been introduce by Thomson [125] (see also [74, 75]) The reverse-time system, however, can be seen as a generalization of Thomson's approach and is derived in [85] in a more transparent and more rigorous way.

Suppose that t is a fixed initial time parameter and $w(s, \boldsymbol{y}) := w(t; s, \boldsymbol{y})$. We introduce a reverse-time variable $\tilde{s} = T + t - s$ and define the functions

$$\begin{aligned} &\tilde{a}_i(\tilde{s}, \boldsymbol{y}) = a_i(T + t - \tilde{s}, \boldsymbol{y}), \\ &\tilde{b}_{ij}(\tilde{s}, \boldsymbol{y}) = b_{ij}(T + t - \tilde{s}, \boldsymbol{y}) \\ &\tilde{\sigma}_{ij}(\tilde{s}, \boldsymbol{y}) = \sigma_{ij}(T + t - \tilde{s}, \boldsymbol{y}) \\ &\tilde{w}(\tilde{s}, \boldsymbol{y}) = w(T + t - \tilde{s}, \boldsymbol{y}) \end{aligned} \tag{5.8}$$

Clearly, $\tilde{w}(T, \boldsymbol{y}) = w(t, \boldsymbol{y})$. Substituting the variable $\tilde{s}$ into the Fokker-Planck equation (5.6) yields

$$\begin{aligned} &\frac{\partial \tilde{w}}{\partial \tilde{s}} + \sum_{i=1}^{d} \frac{\partial}{y_i}(\tilde{a}_i(\tilde{s}, \boldsymbol{y})\tilde{w}) - \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial y_i \partial y_j}(\tilde{b}_{ij}(\tilde{s}, \boldsymbol{y})\tilde{w}) = 0 \\ &\tilde{w}(T, \boldsymbol{y}) = g(\boldsymbol{y}) \end{aligned} \tag{5.9}$$

By rearranging terms the equation (5.9) we can write

$$\begin{aligned} &\frac{\partial \tilde{w}}{\partial \tilde{s}} = -\frac{1}{2}\sum_{i,j=1}^{d} \tilde{b}_{ij}\frac{\partial^2 \tilde{w}}{\partial y_i \partial y_j} - \sum_{i=1}^{d} \alpha_i \frac{\partial \tilde{w}}{\partial y_i} - c(\tilde{s}, \boldsymbol{y})\tilde{w} \\ &\tilde{w}(T, \boldsymbol{y}) = g(\boldsymbol{y}) \end{aligned} \tag{5.10}$$

where

$$\begin{aligned} \alpha_i(\tilde{s}, \boldsymbol{y}) &= \sum_{j=1}^{d} \frac{\partial \tilde{b}_{ij}}{\partial y_j} - \tilde{a}_i \\ c(\tilde{s}, \boldsymbol{y}) &= \frac{1}{2} \sum_{i,1=1}^{d} \frac{\partial^2 \tilde{b}_{ij}}{\partial y_i \partial y_j} - \sum_{i=1}^{d} \frac{\partial \tilde{a}_i}{\partial y_i} \end{aligned} \tag{5.11}$$

So, we obtain a Cauchy problem in reverse time and may state the following result [85, 74]

$$w(T, \boldsymbol{y}) = \tilde{w}(t, \boldsymbol{y}) = E\left[g\left(\boldsymbol{Y}_{t,\boldsymbol{y}}(T)\right) \mathcal{Y}_{t,\boldsymbol{y}}(T)\right] \tag{5.12}$$

where $(\boldsymbol{Y}, \mathcal{Y})$ is the solution of the following system of the stochastic differential equations called reverse-time system

$$\begin{aligned} &d\boldsymbol{Y} = \boldsymbol{\alpha}(\tilde{s}, \boldsymbol{Y}) d\tilde{s} + \tilde{\boldsymbol{\sigma}}(\tilde{s}, \boldsymbol{Y}) d\tilde{\boldsymbol{W}}(\tilde{s}) \\ &d\mathcal{Y} = c(\tilde{s}, \boldsymbol{Y}) \mathcal{Y} d\tilde{s} \\ &\boldsymbol{Y}(t) = \boldsymbol{y}, \quad \mathcal{Y}(t) = 1 \end{aligned} \tag{5.13}$$

Further, for the sake of simplicity the tildes will be dropped.

Now we are ready to introduce the Forward Reverse Estimator (FRE). We follow the heuristic discussion, that can be found in [85]. Let t^* be an internal point of the interval $[t, T]$. Then the transition density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ satisfies the Kolmogorov-Chapman equation (2.30)

$$p(t, \boldsymbol{x}; T, \boldsymbol{y}) = \int\limits_{\mathbb{R}^d} p(t, \boldsymbol{x}; t^*, \boldsymbol{v}) p(t^*, \boldsymbol{v}; T, \boldsymbol{y}) d\boldsymbol{v} \tag{5.14}$$

By assuming that $g(\boldsymbol{v}) = p(t, \boldsymbol{x}; t^*, \boldsymbol{v})$ and applying the equation (5.12), it follows that

$$p(t, \boldsymbol{x}; T, \boldsymbol{y}) = Ep\left(t, \boldsymbol{x}; t^*, \boldsymbol{Y}_{t^*,\boldsymbol{y}}(T)\right) \tag{5.15}$$

We cannot apply the Monte Carlo estimator directly because, in general, the density $\boldsymbol{v} \to p(t, \boldsymbol{x}; t^*, \boldsymbol{v})$ is unknown. The key idea is now to estimate the function $p(t, \boldsymbol{x}; t^*, \boldsymbol{v})$ from a sample of independent realizations of the stochastic process $\boldsymbol{X}_{t,\boldsymbol{x}}$ on the interval $[t, t^*]$ using the kernel estimator (2.51). Then, replacing the unknown function in (5.14) by its numerical approximation $\widehat{p}$ and applying the Monte Carlo method yield to the following estimator

$$\widehat{p}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \frac{1}{M} \left[\frac{1}{N\lambda^d} \sum_{m=1}^{M} \sum_{n=1}^{N} K\left(\frac{\bar{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(t^*) - \bar{\boldsymbol{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T)}{\lambda} \right) \bar{\mathcal{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T) \right] \tag{5.16}$$

A formal description of the forward reverse estimator can be found in [85]. By taking $t^* = T$ the estimator (5.16) collapses to the pure forward estimator (2.51). By taking $t^* = t$ we will obtain the estimator that is based only on the reverse-time system $(\boldsymbol{Y}_{t,\boldsymbol{y}}, \mathcal{Y}_{t,\boldsymbol{y}})$

$$\widehat{p}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \frac{1}{M\lambda^d} \sum_{m=1}^{M} K\left(\frac{\boldsymbol{x} - \bar{\boldsymbol{Y}}_{t,\boldsymbol{y}}^{(m)}(T)}{\lambda} \right) \bar{\mathcal{Y}}_{t,\boldsymbol{y}}^{(m)}(T) \tag{5.17}$$

So, we can consider the forward reverse estimator as an extension of the classical kernel estimator and, as a result, it has superior properties in comparison with density estimators based solely on the forward estimator or solely on the reverse estimator.

5.2.3 Accuracy and complexity of the forward reverse estimator

In Section 2.4.2 the accuracy of the kernel estimator based on the forward system was introduced

$$\epsilon(p) := \sqrt{E(\hat{p}-p)^2} = \sqrt{E(\hat{p}-E\hat{p})^2 + (E\hat{p}-p)^2} = \sqrt{Var(\hat{p}) + Bias^2(\hat{p})}$$

Loosely speaking, for a second order kernel applied in (5.16) under $M = N$ and any choice of $t < t^* < T$, the FRE has root-N ($O(N^{-1/2})$) accuracy for dimension $d \leq 4$ (see [85]). For $d > 4$ root-N accuracy is lost but then the FRE accuracy order is still the square of the FE/RE accuracy order (see Table 5.1.). Moreover, it can be shown that root-N accuracy of (5.16) can also be achieved for $d > 4$ by using higher order kernels.

By definition of the accuracy of the estimator it is possible to relate the "expected" accuracy of the different density estimators to the number of simulated trajectories involved. However, simulating trajectories is not the only costly issue in the density estimation. For all estimators one has to evaluate a functional of the simulated trajectories which consists of a single sum for the FE/RE estimators but a more complicated double sum for the FRE estimator. Therefore, for a proper comparison it is better to consider the *complexity* of the different estimators which is defined as the required computation cost for reaching a given accuracy ϵ. For instance, naive evaluation of the double sum in (5.16) would require a computational cost of order $O(MN)$ in contrast to $O(N)$ for the FE/RE estimators! Clearly, such an naive approach would have a dramatic impact on the complexity of the FRE. Fortunately, there exists smarter procedures for evaluating this double sum which utilize the small support of the kernel K. Particularly, in [49] it is shown for Gaussian kernels that this sum can be evaluated at a cost of $O(N \log N)$ in the case $M = N$ and it is not difficult to see that for other kernels with compact support a cost not more than $O(N \log N)$ can be achieved as well. In Table 1. we resume the results of [85] and list the accuracy and complexity of the forward estimator (FE), reverse estimator (RE) and the forward-reverse estimator (FRE) where the latter is implemented with an efficient procedure for evaluating the double sum, for instance, according to a method of Greengard and Strain. For the FRE estimator we assumed $N = M$ and a second order kernel. Due to the second order kernel we have to distinguish for the FRE between $d < 4$ and $d \geq 4$.

Estimator	δ_N	*Accuracy*	*Complexity*	$\frac{Compl.\{FE,RE\}}{Compl.\{FRE\}}$
FE/RE	$N^{-\frac{1}{4+d}}$	$O(N^{-\frac{2}{4+d}})$	$O(\epsilon^{-2-\frac{d}{2}})$	
FRE $d \leq 4$	$N^{-\frac{1}{d}} \log^{\frac{1}{d}} N$	$O(N^{-\frac{1}{2}})$	$O(\lvert \log \varepsilon \rvert \epsilon^{-2})$	$\lvert \log \varepsilon \rvert^{-1} \epsilon^{-\frac{d}{2}}$
FRE $d > 4$	$N^{-\frac{2}{4+d}}$	$O(N^{-\frac{4}{4+d}})$	$O(\lvert \log \varepsilon \rvert \epsilon^{-1-\frac{d}{4}})$	$\lvert \log \varepsilon \rvert^{-1} \epsilon^{-1-\frac{d}{4}}$

Table 5.1: Accuracy and complexity of FE, RE and FRE

t!

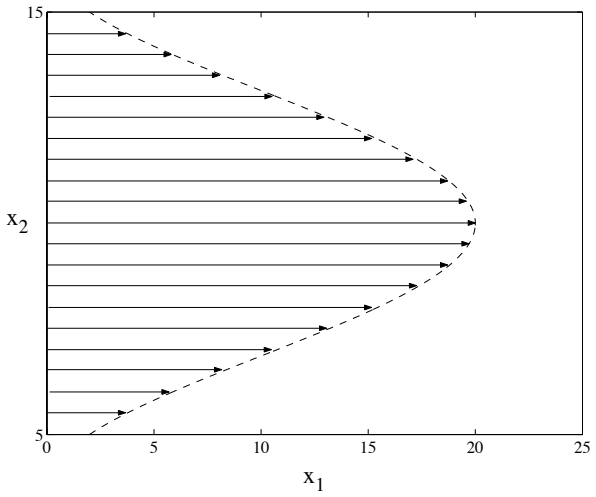


Figure 5.3: The flow field

5.3 Illustrative example

In this section we compare the FE and the FRE in calculating the transition density function for a simple two-dimensional particle model. We assume that the water depth H and diffusion coefficient D are constants. In this case the model can be rewritten as

$$\begin{aligned} dX_1(s) &= u_1 ds + \sqrt{2D} dW_1(s) \\ dX_2(s) &= u_2 ds + \sqrt{2D} dW_2(s) \\ \boldsymbol{X}(t) &= \boldsymbol{x} = (x_1, x_2) \end{aligned} \tag{5.18}$$

The reverse system of (5.18) is:

$$\begin{aligned} dY_1(s) &= -u_1 ds + \sqrt{2D} d\widetilde{W}_1(s) \\ dY_2(s) &= -u_2 ds + \sqrt{2D} d\widetilde{W}_2(s) \\ d\mathcal{Y}(s) &= \left(-\tfrac{\partial u_1}{\partial y_1} - \tfrac{\partial u_2}{\partial y_2} \right) \mathcal{Y} ds \\ \boldsymbol{Y}(t^*) &= \boldsymbol{y} = (y_1, y_2), \quad \mathcal{Y}(t^*) = 1 \end{aligned} \tag{5.19}$$

The velocities u_1 and u_2 have to satisfy the condition of free divergence:

$$\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \equiv 0$$

In this case the last equation in system (5.19) becomes trivial and the weighting coefficient $\mathcal{Y}_{t^*,\boldsymbol{y}} = 1$ for each realization of stochastic process $\boldsymbol{Y}(s)$. As an example, we consider the following divergence free flow (see figure 5.3)

$$\begin{aligned} u_1 &= f(x_2) = 10(cos(0.5(x_2 - 10)) + 1) \\ u_2 &\equiv 0 \end{aligned}$$

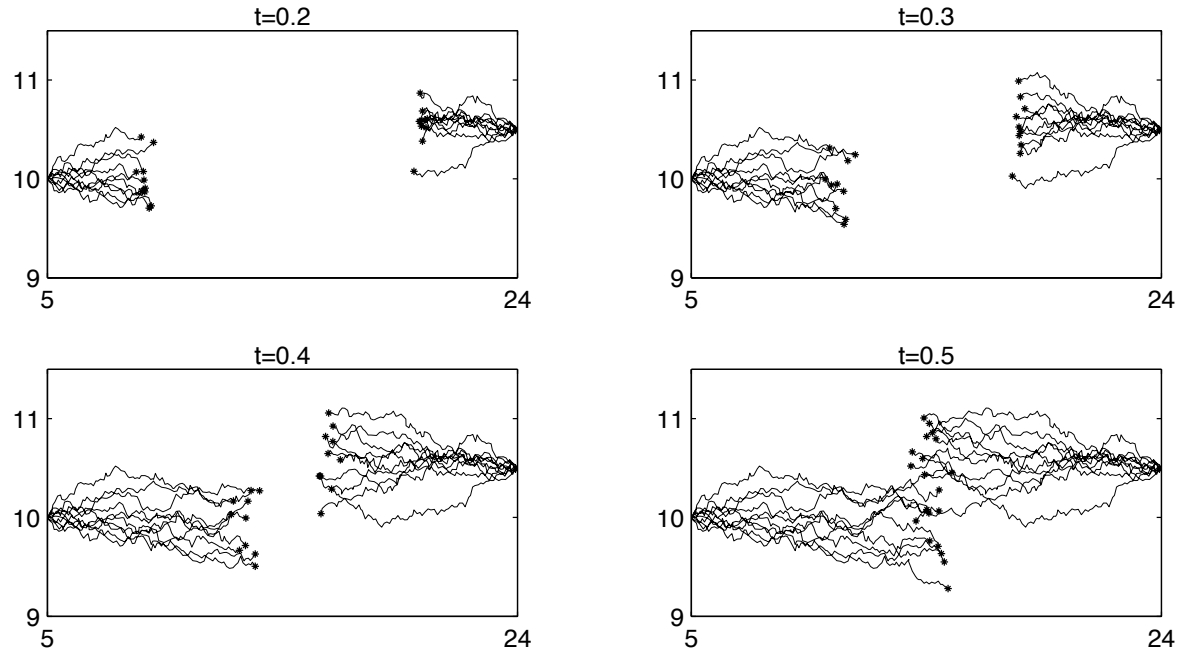


Figure 5.4: The simulation of forward and reverse-time tracks for 10 particles

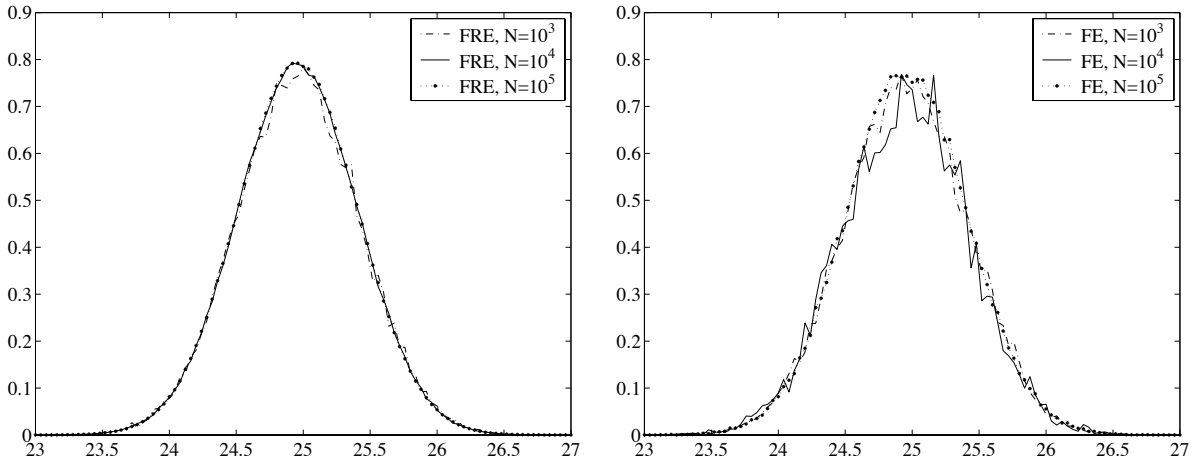


Figure 5.5: The transition density function via the FRE and the FE. The value of parameters are $t = 0$, $T = 1$, $\boldsymbol{x} = (15, 10)$

and we apply both the FE and the FRE to the following model

$$\begin{aligned} dX_1(s) &= f(X_2(s))ds + \sqrt{2D}dW_1(s) \\ dX_2(s) &= \sqrt{2D}dW_2(s) \\ \boldsymbol{X}(t) &= \boldsymbol{x} = (x_1, x_2) \end{aligned}$$

method	N	$p(t, \boldsymbol{x}; T, \boldsymbol{y})$	*method*	N	$p(t, \boldsymbol{x}; T, \boldsymbol{y})$
FE	10^3	0.0819 ± 0.0267	FRE	10^3	0.0790 ± 0.0087
FE	10^4	0.0787 ± 0.0097	FRE	10^4	0.0729 ± 0.0033
FE	10^5	0.0732 ± 0.0028	FRE	10^5	0.0721 ± 0.0011
FE	10^6	0.0728 ± 0.0013	FRE	10^6	0.0721 ± 0.0003
FE	10^7	0.0722 ± 0.0006	FRE	10^7	0.0721

Table 5.2: The value of the transition density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ for the test problem

with reverse system

$$\begin{aligned} dY_1(s) &= -f(Y_2(s))ds + \sqrt{2D}d\widetilde{W}_1(s) \\ dY_2(s) &= \sqrt{2D}d\widetilde{W}_s^2 \\ \boldsymbol{Y}(t^*) &= \boldsymbol{y} = (y_1, y_2) \end{aligned}$$

For this model the following parameters were chosen:

- diffusion coefficient - $D = 0.1$ kg/m;
- initial and end positions of particle - $\boldsymbol{x} = (x_1, x_2) = (5, 10)$, $\boldsymbol{y} = (y_1, y_2) = (24, 10.5)$;
- start and end time - $t = 0$, $T = 1$;
- internal point - $t^* = \frac{1}{2}(T - t) = 0.5$;
- Gaussian kernel function - $K(v_1, v_2) = \frac{1}{2\pi}\exp(-(v_1^2 + v_2^2)/2)$
- bandwidth - for FE $\lambda = \mathcal{K}N^{-\frac{1}{6}}$, where $\mathcal{K} = \left(\frac{1}{N-1}\sum_{n=1}^{N}\left|X^{(n)} - \frac{1}{N}\sum_{n=1}^{N}X^{(n)}\right|^2\right)^{\frac{1}{2}}$, for FRE $\lambda = \sqrt{\frac{\log(N)}{N}}$.

We use the Euler scheme with timestep $\Delta t = 0.005$ to obtain a numerical solution of the systems (5.18) and (5.19) and to simulate the forward and reverse-time particle's tracks. On Figure 5.3 the simulation results of 10 particles are shown. Using FE and FRE estimating procedures we determine the value of transition density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$. The relation between the probability density and the particle concentration described in Section 3.3.2, $p(t, \boldsymbol{x}; T, \boldsymbol{y})/H$ is the particle concentration in point $\boldsymbol{y}$ at time T, given a release of the particles in $\boldsymbol{x}$ at time t.

The results of the FE and the FRE are shown in Table 5.2. For different values of sample size N we have done the simulation 10 times to gain insight into the error of both the FE and the FRE methods.

In order to demonstrate the efficiency of the FRE in comparison with the FE, another example was examined too. We fix $t = 0$, $x_1 = 5$, $x_2 = 10$, $T = 1$, $y_2 = 10.0$ and calculate the transition density function $p(t, x_1, x_2, T, y_1, y_2) = p(y_1)$ using both the FE and the FRE with the same parameters as in the previous case.

Figure 5.5 shows the results of the two methods for $N = 10^3$, $N = 10^4$ and $N = 10^5$ particles. Because $p(t, \boldsymbol{x}; T, \boldsymbol{y})/H$ is the particle concentration Figure (5.5) shows the particle concentration at time T. By comparing the figures of the density functions it is clear that the FRE is much more accurate than the classical FE.

5.4 Tidally-averaged model

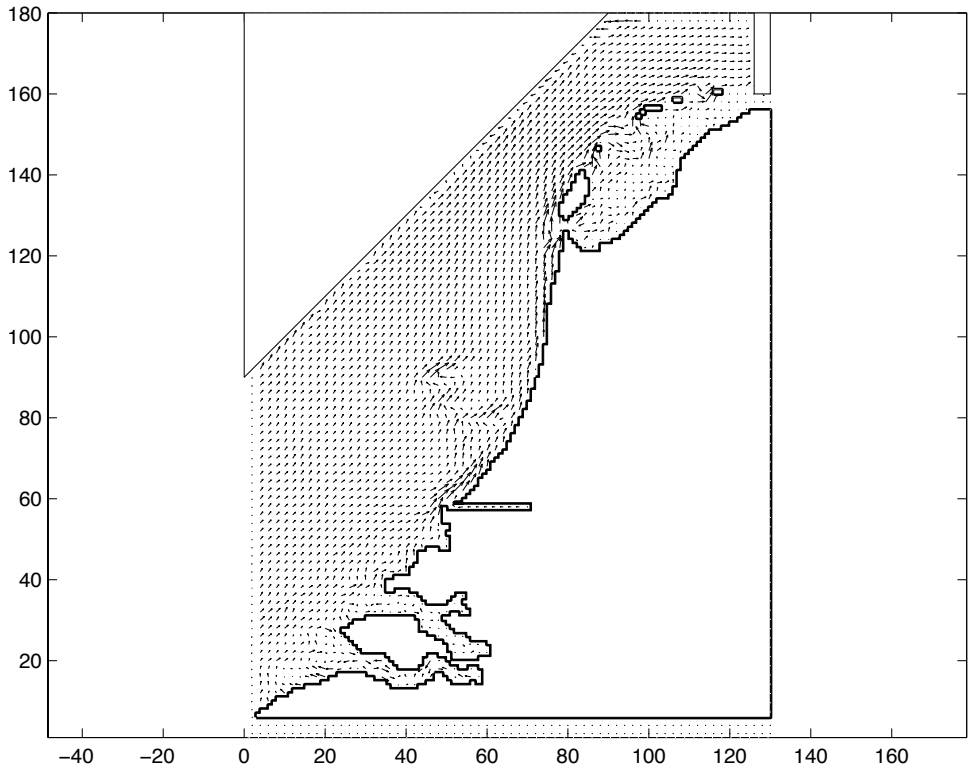
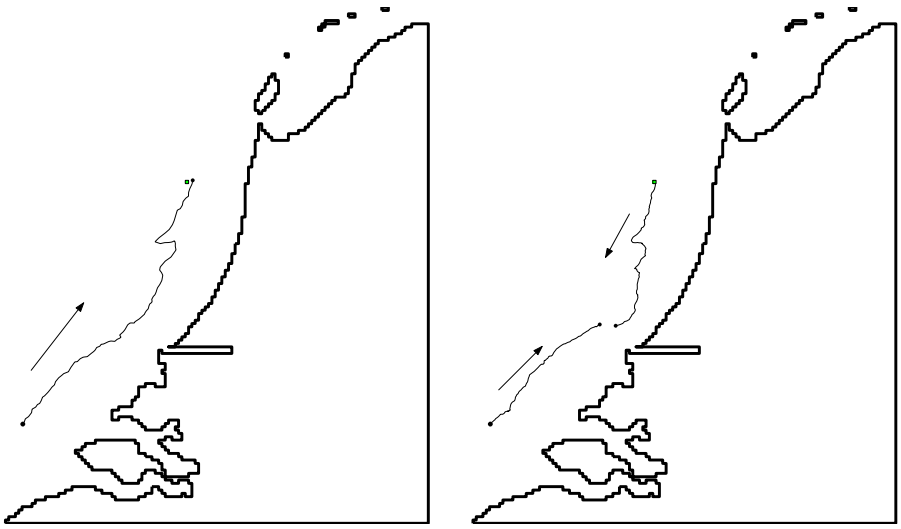


Figure 5.6: The tidally-averaged flow

The random walk model described in the previous section is applied to the tidally-averaged numerical model. This numerical flow was done from the tidal velocity field from the two-layer model of the Dutch coastal zone, including stratified areas in the Rhine outflow region, with vertical diffusion coefficients below 10^{-4} m^2/s. In this region, the velocity differences between the layers are most of the time larger than 0.5 m/s and the vertical differences of the residuals are about 0.1 m/s. Depth range is of 10-30 m. An instantaneous horizontal diffusion of 5 m^2/s was used in all transport simulations. $\Delta X = \Delta Y = 1600\,m \times 1600\,m$

The used 'tidal' velocities were not determined by tide only, but also by varying wind and density structure of the water. They were taken from the last part of a simulated event with large Rhine outflow (4000 m^3/s). After a period of moderate northeasterly winds the wind veered to SSW. A large baroclinic eddy was present in the computed residual velocity field (see Figure 5.6). Because the velocities and water depth are known only in the knots of the grid, we use bilinear interpolation to obtain the velocities and water depth in the arbitrary location. This tidally-averaged model of the Dutch coastal zone was constructed by de Kok in [33]. Additional information can be also found in



(a) (b)

Figure 5.7: The simulation of some particle's tracks (a) forward track (b) forward and reverse-time tracks

[32, 53].

5.5 Application (the FRE)

Because the diffusion coefficient D is a constant, the random walk model can be written as follows:

$$\begin{aligned} dX_1(s) &= \left(u_1 + \frac{D\partial H}{H\partial x_1}\right) ds + \sqrt{2D} dW_1(s) \\ dX_2(s) &= \left(u_2 + \frac{D\partial H}{H\partial x_2}\right) ds + \sqrt{2D} dW_2(s) \\ \boldsymbol{X}(t) &= \boldsymbol{x} = (x_1, x_2) \end{aligned} \tag{5.20}$$

and the corresponding reverse-time system system becomes:

$$\begin{aligned} dY_1(s) &= -\left(u_1 + \frac{D\partial H}{H\partial y_1}\right) ds + \sqrt{2D} d\widetilde{W}_1(s) \\ dY_2(s) &= -\left(u_2 + \frac{D\partial H}{H\partial y_2}\right) ds + \sqrt{2D} d\widetilde{W}_2(s) \\ d\mathcal{Y}(s) &= \left(-\frac{\partial(u_1 + \frac{D\partial H}{H\partial y_1})}{\partial y_1} - \frac{\partial(u_2 + \frac{D\partial H}{H\partial y_2})}{\partial y_2}\right) \mathcal{Y} ds \\ \boldsymbol{Y}(t^*) &= \boldsymbol{y} = (y_1, y_2), \quad \mathcal{Y}(t^*) = 1 \end{aligned} \tag{5.21}$$

We use the Euler scheme (with L time-steps) to obtain the numerical solutions of systems (5.20) and (5.21). An example of a simulation of a particle's forward track is

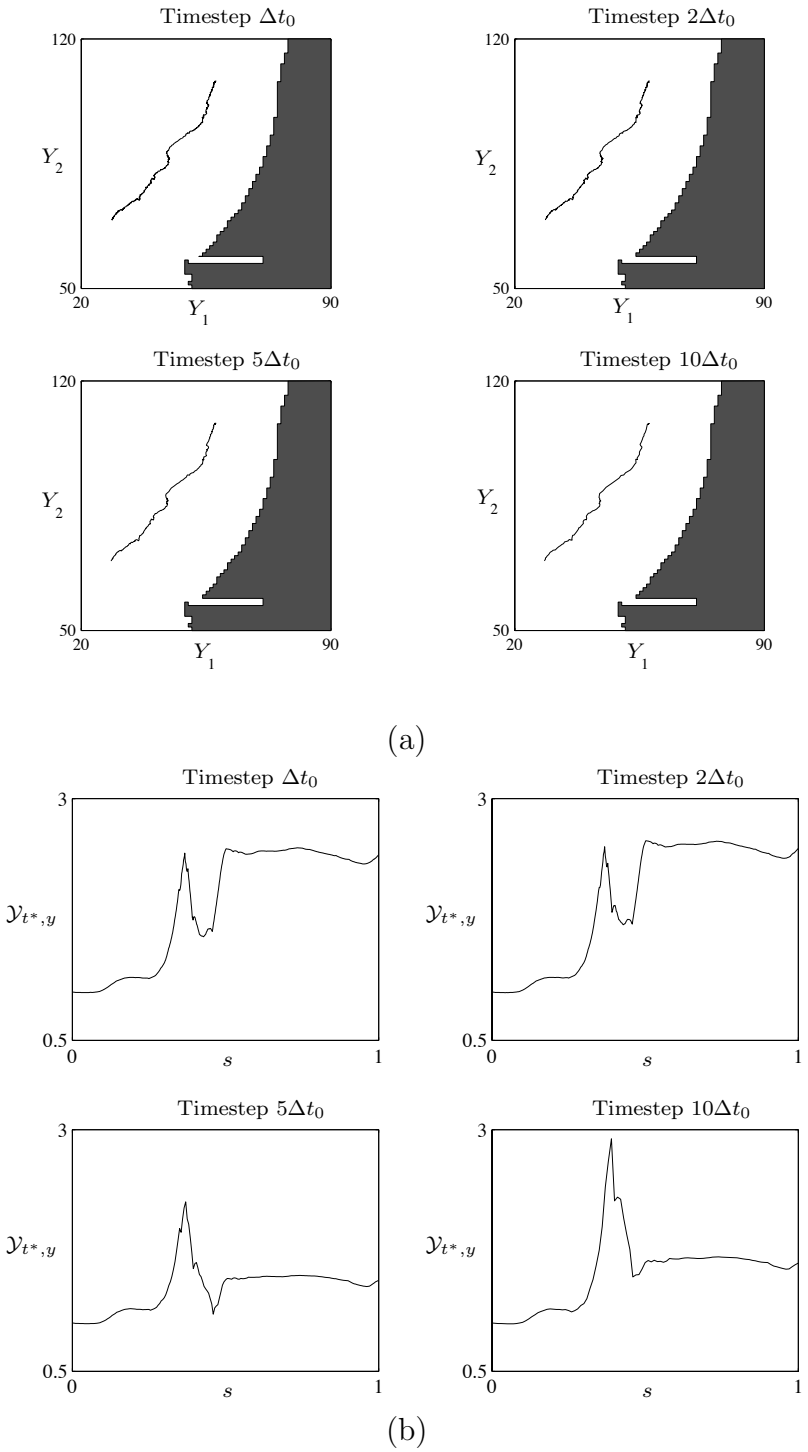


Figure 5.8: (a) The reverse-time tracks of particle for different time-steps and (b) the function $\mathcal{Y}_{t^*,\boldsymbol{y}}(s)$ along the particle tracks $\boldsymbol{Y}_{t^*,\boldsymbol{y}}(s)$ for different time steps

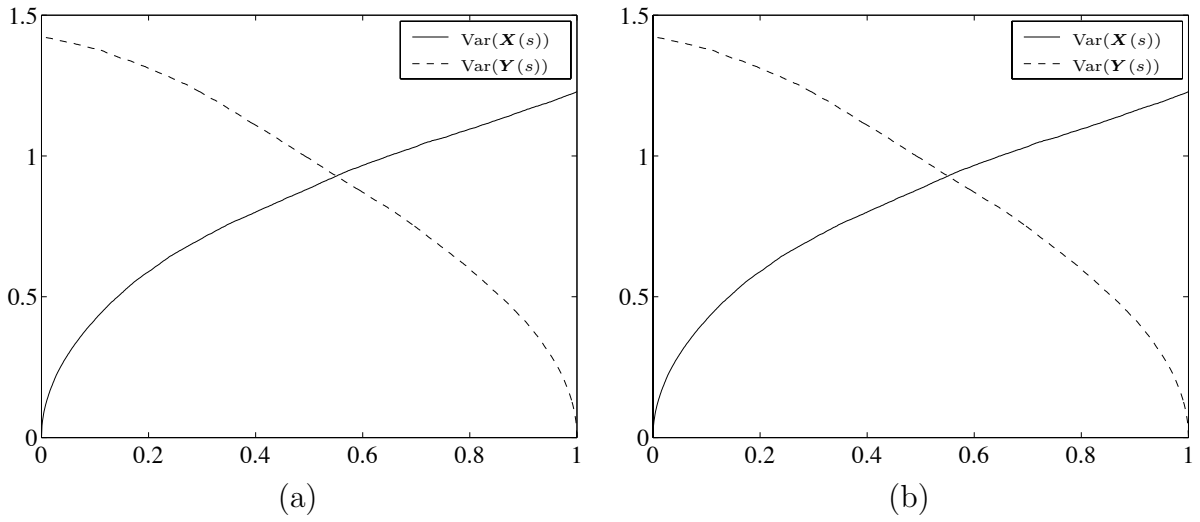


Figure 5.9: The variance of forward and reverse-time processes (a) $\boldsymbol{y} = (14.0, 43.0)$ $t^* = 0.55(T-t)+t$ and (b) $\boldsymbol{y} = (57.8, 108.1)$ $t^* = 0.75(T-t)+t$

shown in Figure 5.7 (a) and a simulation of a forward and a reverse-time particle's track is shown in Figure 5.7 (b).

In order to get an accurate approximation of the stochastic process (5.21) we have to choose the time-step $\Delta t = (T-t)/L$ sufficiently small, because the function $\mathcal{Y}(s)$ is very sensitive to the choice of Δt in areas where the flow is very turbulent. In Figure 5.8 we show an example of realizations of respectively the stochastic process $\boldsymbol{Y}_{t^*,\boldsymbol{y}}(s)$ and the function $\mathcal{Y}_{t^*,\boldsymbol{y}}(s)$ with different time-steps but with the same driving Wiener process. Figure 5.8(a) shows that changing the timestep from $\Delta t = \Delta t_0 = 2250\,s$ to $\Delta t = 10\Delta t_0 = 22500\,s$ does hardly influence the particle's reverse-time track, while Figure 5.8(b) shows that the value of $\mathcal{Y}_{t^*,\boldsymbol{y}}(T)$ changes substantially.

Unlike the FE, the FRE has the internal point t^*, which can be chosen in the interval $[t, T]$. As was mentioned before, if $t^* = t$, then the FRE will be equivalent to a pure RE and if $t^* = T$ we will get a pure FE. Van den Berg [10] showed that an efficient choice of the internal point t^* is where $Var(\boldsymbol{X}_{t,\boldsymbol{x}})(t^*) = Var(\boldsymbol{Y}_{t^*,\boldsymbol{y}})(T)$ and we used this rule as starting point for our application (see Figure 5.9).

From Figure 5.9 we see that the weighting coefficient is very irregular. Therefore, in order to get an accurate solution for the weighting coefficient $\mathcal{Y}$ we need to simulate the time-reverse particle's tracks with a very small timestep Δt. As a result, the reverse simulation takes a lot of CPU time. However, we can reduce this effect if we choose t^* closer to the end point T. The FRE is efficient if all weighting coefficients are close to 1 because in this case all reverse-time particles contribute equally to the final results. If there is a lot of variation in the weighting coefficients then a few reverse time tracks dominate the problem. Variations in the weighting coefficient can be reduced by, again, choosing internal time point t^* closer to the end time point T.

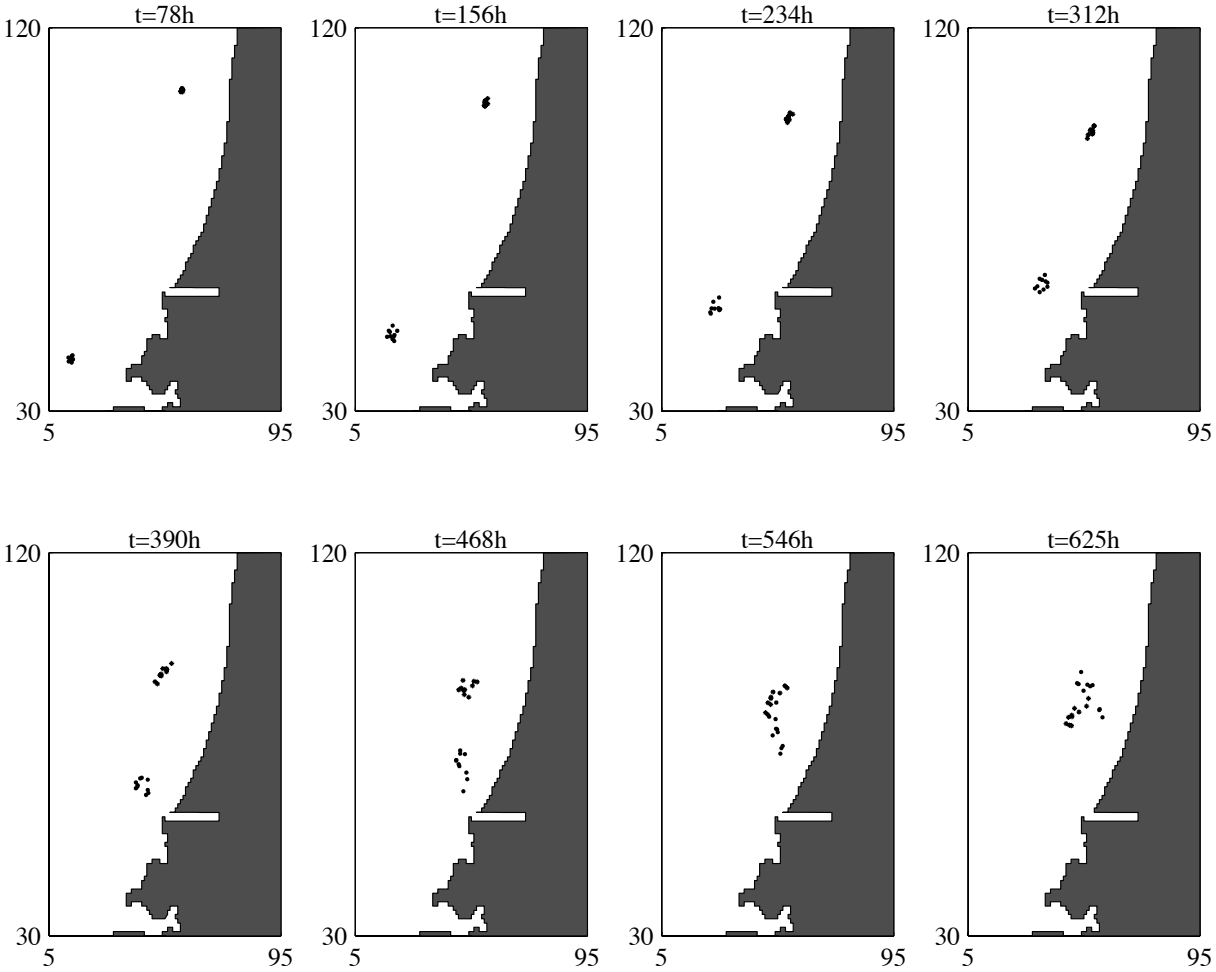


Figure 5.10: The realizations of forward and reverse-time systems for 10 *particles*

We compared the both methods for two different locations:

- $\boldsymbol{x} = (8.2, 35.8), \quad \boldsymbol{y} = (14.0, 43.0),$
  $t = 0, \quad T = 22.5 \cdot 10^4\,\text{sec}, (\approx 62.5\,\text{hours}), \quad t^* = 0.55 * (T - t) + t, \quad L = 100$
- $\boldsymbol{x} = (8.2, 35.8), \quad \boldsymbol{y} = (57.8, 108.1),$
  $t = 0, \quad T = 22.5 \cdot 10^5\,\text{sec}\,(\approx 625\,\text{hours}), \quad t^* = 0.75 * (T - t) + t, \quad L = 1000$

In Figure 5.19 10 particle tracks are shown for the case $\boldsymbol{y} = (57.8, 108.1)$. For this application we used the Epanechnikov kernel (2.45) because it is more efficient (see [106]).

In Table 5.3 the results of applying the FE and FRE are shown for different locations. We did not apply the FRE for a number of particles more than 10^6. Assuming that the results for $N = 10^6$ are close to the exact solution, we can conclude that the FRE is again much more accurate than the FE.

$\boldsymbol{y}$	method	$N = 10^4$	$N = 10^5$	$N = 10^6$	$N = 10^7$	$N = 10^8$
$(14.0, 43.0)$	FE	0.2024	0.1808	0.1831	0.1864	0.1885
	FRE	0.1886	0.1879	0.1886		
$(57.8, 108.1)$	FE	0.0056	0.0065	0.0076	0.0067	0.0069
	FRE	0.0073	0.0070	0.0069		

Table 5.3: The numerical results of the forward and the forward-reverse estimators

5.6 Risk analysis in a coastal zone

Suppose that for some critical locations a high concentration of the pollutant can be dangerous or even fatal for life of some species. To prevent ecological disasters we need to determine which areas in the sea are potentially dangerous for the critical location and which areas are safe, in other words we need to construct a risk map. By the risk map for a given location of interest along the coast we mean the concentration of the pollutant in this area for any locations $\boldsymbol{x} = (x_1, x_2)$ where the pollutant may be released. From a risk map we can immediately determine the most dangerous as well as the safe location of the pollutant release. One of the most effective approaches to solve this problem is to use the reverse time particle models based on the solution of the backward Kolmogorov equation [1, 38, 101].

The reverse-time system (5.21) may be written in the matrix form as follows:

$$\begin{aligned} d\boldsymbol{Y}(s) &= \left(-\boldsymbol{u} - \frac{D}{H}\frac{\partial H}{\partial \boldsymbol{z}}\right) ds + (\sqrt{2D}\boldsymbol{I}_2) d\boldsymbol{W}(s) \\ d\mathcal{Y}(s) &= -\frac{\partial\left(\boldsymbol{u} + \frac{D}{H}\frac{\partial H}{\partial \boldsymbol{z}}\right)}{\partial \boldsymbol{z}^T}\mathcal{Y}(s) ds \\ \boldsymbol{Y}(t) &= \boldsymbol{y},\ \mathcal{Y}(t) = 1 \end{aligned} \tag{5.22}$$

where $\dfrac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}^T} = \dfrac{\partial \boldsymbol{F}}{\partial x_1} + \dfrac{\partial \boldsymbol{F}}{\partial x_2}$. This reverse-time system was used to construct the risk maps for different moments of time at several critical locations along Dutch seaside.

To solve this equation we use the Euler scheme with time step $\Delta t = 3$ min. The reverse estimator (5.17) with the parabolic kernel and the bandwidth λ

$$\lambda = \mathcal{K} N^{-\frac{1}{6}}, \quad \mathcal{K} = \left(\frac{1}{N-1}\sum_{n=1}^{N} \|\overline{\boldsymbol{Y}}_{t,\boldsymbol{y}}^{(n)} - \frac{1}{N}\sum_{n=1}^{N} \overline{\boldsymbol{Y}}_{t,\boldsymbol{y}}^{(n)}\|^2\right)^{\frac{1}{2}}$$

This section is based on the article:

Spivakovskaya D., Heemink A.W. and Schoenmakers J.G.M. 2006 'The risk analysis in coastal zone using reverse-time diffusion'. In 'Proceedings of the XVI International Conference on Computational Methods in Water Resources', P.J. Binning, P.K. Engesgaard, H.K. Dahle, G.F. Pinder and W.G. Gray (Eds.). Copenhagen, Denmark. http://proceedings.cmwr-xvi.org/getFile.py/access?contribId=100&resId=0&materialId=paper&confId=a051

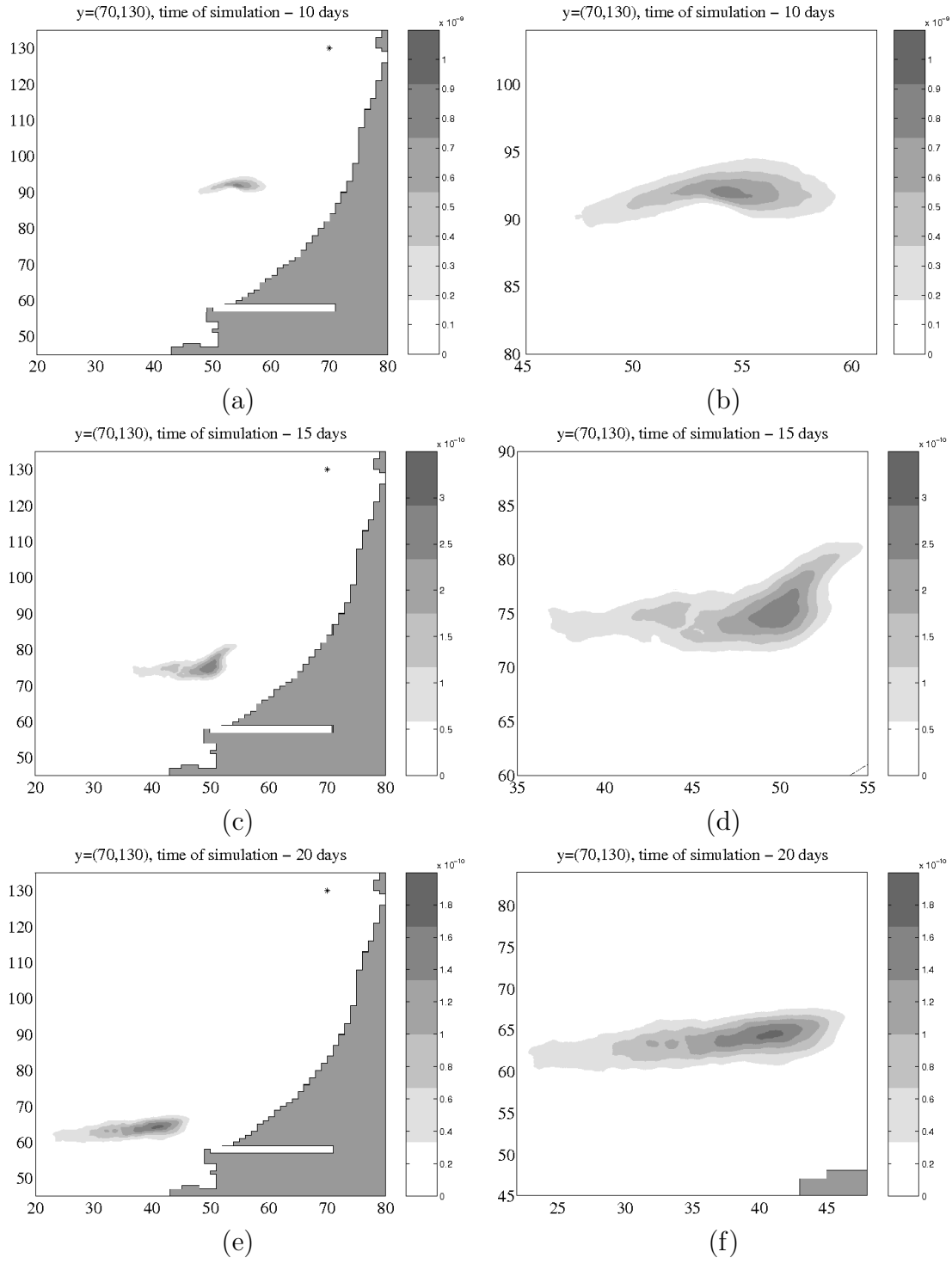


Figure 5.11: The risk map for $\boldsymbol{y} = (70, 130)$ for (a),(b) 10 days, (c),(d) 15 days, (e), (f) 20 days

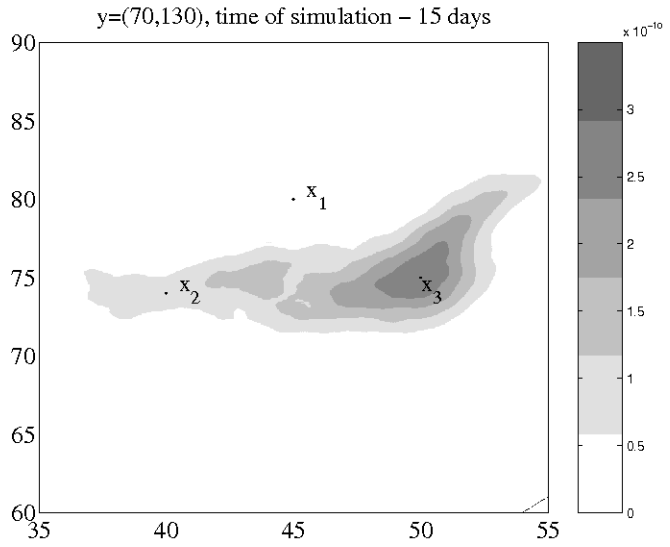


Figure 5.12: The risk map for critical location $\boldsymbol{y} = (70, 130)$ *for* $T = 15$ *days*

was used.

Figure 5.11 represents the risk maps for the location $\boldsymbol{y} = (70, 130)$ for 10 days (Figures 5.11 (a),(b)), 15 days (Figures 5.11 (c),(d)) and 20 days (Figures 5.11 (e),(f)). They have been constructed using $N = 10^5$ particles.

These maps give us an idea which locations of release may be dangerous for the fixed critical location $\boldsymbol{y}$. For instance, let us consider the risk map for 15 days (see Figure 5.12). If the pollutant was released at point $\boldsymbol{x}_1 = (45, 80)$, then the concentration at critical location $\boldsymbol{y}$ after 15 days would be almost 0. It means that the release of the pollutant at $\boldsymbol{x_1}$ will not affect the point $\boldsymbol{y}$. The location $\boldsymbol{x}_2 = (40, 74)$ is not safe, although the release of the pollutant at this location may lead to the concentration less than $M10^{-10}\text{kg/m}^3$ (M is mass of initially released pollutant). The release of the pollutant at point $\boldsymbol{x}_3 = (50, 75)$ however would affect the location $\boldsymbol{y}$ very strongly.

In Table 5.4 one can see the concentration of the pollutant estimated for different locations of release $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$. To calculate the concentration the forward (2.51) and the reverse (5.17) estimators were applied with different numbers of particles. Each realization was repeated 30 times to calculate the statistical error of the estimation. From the results in Table 5.4 it is clear that the statistical error has the same order for the both forward and reverse estimators and converges to zero when the number of particles increases.

$\boldsymbol{x}$	method	$N = 10^3$	$N = 10^4$	$N = 10^5$
$\boldsymbol{x}_1$	FE	0.003 ± 0.011	0.004 ± 0.005	0.003 ± 0.003
	RE	0.019 ± 0.068	0.005 ± 0.002	0.003 ± 0.001
$\boldsymbol{x}_2$	FE	1.025 ± 0.172	0.958 ± 0.095	0.926 ± 0.046
	RE	0.893 ± 0.247	0.923 ± 0.153	0.904 ± 0.060
$\boldsymbol{x}_3$	FE	2.878 ± 0.295	2.870 ± 0.109	2.897 ± 0.072
	RE	2.552 ± 0.320	2.772 ± 0.120	2.791 ± 0.062

Table 5.4: The concentration $C(t, \boldsymbol{x}, T, \boldsymbol{y})$ (10^{-10}), kg/m^3

5.7 The parallel implementation of the FE and the FRE

Because all realizations of particle's paths are independent the random walk model can be easily parallelized. It makes the particle modelling very attractive for a lot of applications. In this section we consider the algorithms of the parallelization for the forward and for the forward-reverse estimators.

5.7.1 The parallelization of the FE/RE estimators

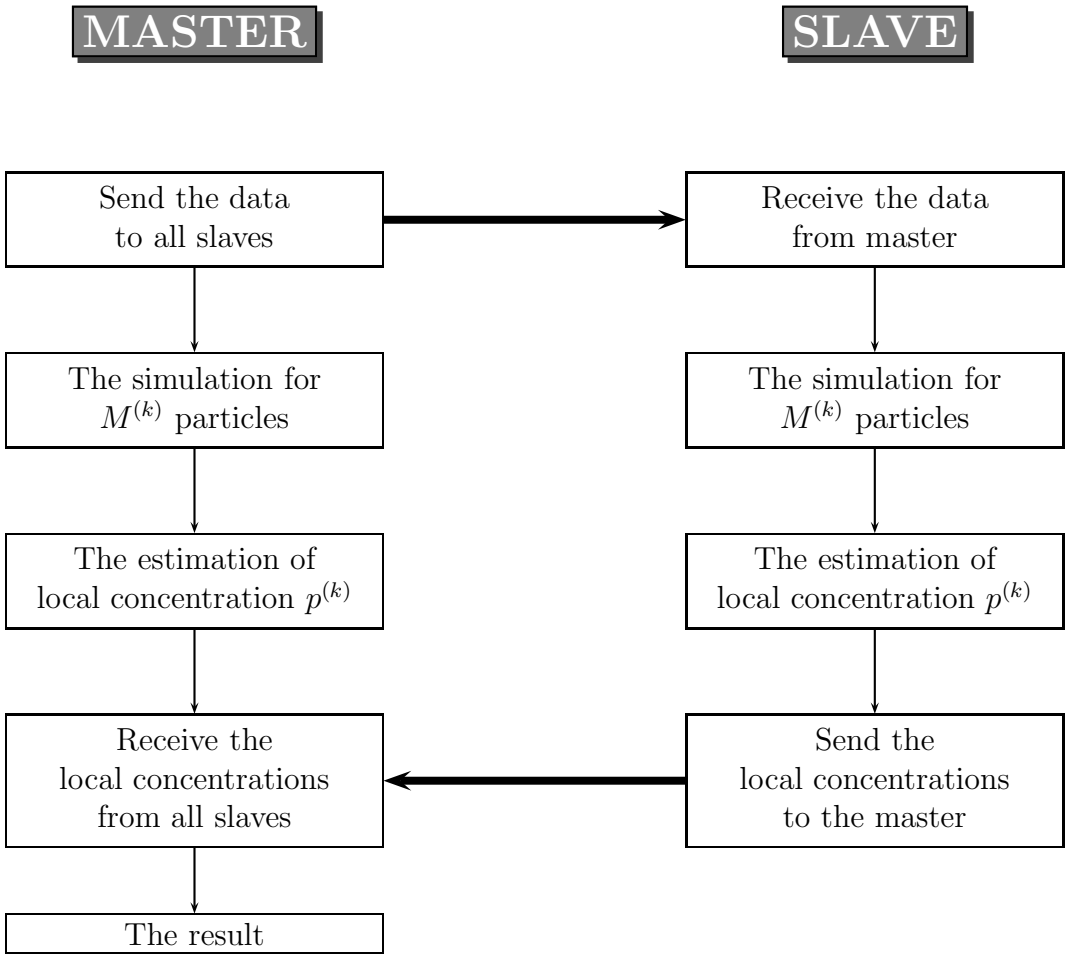
It is very simple to parallelize the FE/RE estimators (2.51) and (5.17). As all particles are independent and do not interact during the simulation, each processor can simulate the movements of subgroup of particles without communications. However, communications are needed to locate particles on the different processors in the beginning and to gather the results from each process. Equation (2.51) can be rewritten as follows

$$\begin{aligned} \widehat{p}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \; & \frac{1}{N\lambda^d} \sum_{n=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(T) - \boldsymbol{y}}{\lambda}\right) = \\ & \frac{1}{N\lambda^d} \sum_{k=1}^{Nproc} \left(\sum_{i=1}^{M^{(k)}} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n,k)}(T) - \boldsymbol{y}}{\lambda}\right)\right) = \sum_{n=1}^{Nproc} p^{(k)} \end{aligned} \tag{5.23}$$

where $Nproc$ is the number of processors, $M^{(k)}$ is a number of particles handled by kth processor, $\sum_{k=1}^{Nproc} M^{(k)} = N$. Each processor gets approximately the same amount of particles, so $M^{(k)} \sim \dfrac{N}{Nproc}$. Finally, $p^{(k)} = \dfrac{1}{\lambda^d} \sum_{i=1}^{M^{(k)}} K\left(\dfrac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n,k)}(T) - \boldsymbol{y}}{\lambda}\right)$ is the contribution of the kth processor to the sum $\widehat{p}$.

This section is based on the article: Huber E., Spivakovskaya D., Lin H.X. and Heemink A.W. 2006 'The parallel implementation of forward-reverse estimator'. In: Proceedings of European Conference on Computational Fluid Dynamics ECCOMAS CFD 2006, P. Wesseling, E. Oñate and J. Périaux (Eds).

The following algorithm can be used for the parallelization of the FE/RE estimators. We need only two communications and the load is equally distributed. That's why the efficiency of this algorithm is close to 1.



5.7.2 The parallelization of the FRE

The FRE cannot be parallelized so easy as it has been done for the FE/RE. It is clear that the parallelization method should be based on the particle decomposition.

For the forward-reverse estimator (5.16) we assume that $M = N$, so we need to simulate the same number of the forward and reverse particles. Therefore, if we evaluate the double sum (5.16) directly, we need $\mathcal{O}(N^2)$ elementary computations. It makes the forward-reverse estimator too computationally expensive. However, it is possible to speedup the calculation, if the kernel function $K(\boldsymbol{x})$ has a compact support, as, for instance, Epanechnikov function (2.45). If we want to use the Gaussian function (2.44), we need first to modify it

$$K_\lambda(\boldsymbol{x}) = \begin{cases} K(\boldsymbol{x}) \text{ as in (2.44)} & \sqrt{\boldsymbol{x}^T\boldsymbol{x}} \le 6\lambda \\ 0 & \text{otherwise} \end{cases} \tag{5.24}$$

This modification does not affect the final result, because for the values $\boldsymbol{x}$, such that $\sqrt{\boldsymbol{x}^T\boldsymbol{x}} > 6\lambda$ the value of Gaussian kernel is very small and can be neglected. As a

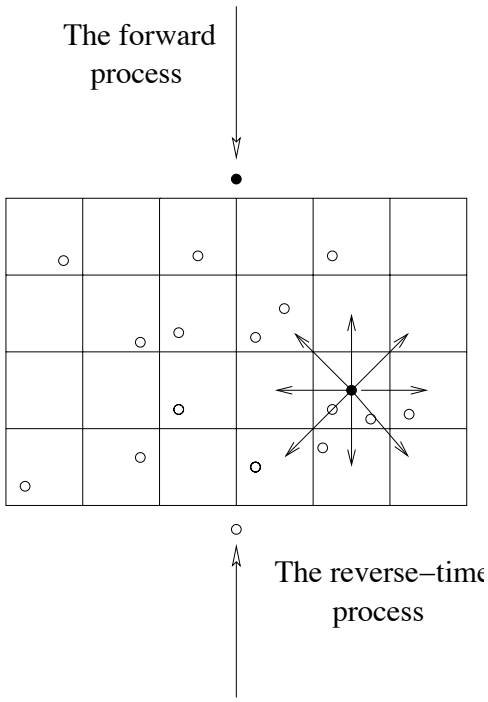


Figure 5.13: The two-dimensional grid

result, in formula (5.16) we need to consider only pairs $\left(\overline{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{x}}(t^*), \overline{\boldsymbol{Y}}^{(m)}_{t^*,\boldsymbol{y}}(T)\right)$ for which $\left|\overline{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{x}}(t^*) - \overline{\boldsymbol{Y}}^{(m)}_{t^*,\boldsymbol{y}}(T))\right| \leq \alpha\lambda$, for some α.

The algorithm proceeds as follows. For a given forward sample $\left\{\boldsymbol{X}^{(n)}\right\}$, $n = 1, \ldots, N$ we define the bounding hyper-rectangle, i.e. the smallest hyper-rectangle in which all realizations lie. This hyper-rectangle is divided in d-dimensional hypercubes with the length of a side equal to $\alpha\lambda$. The size of each cell in the obtained grid can be enlarged to make sure that our grid consists of an integer number of cells. Each realization of the forward process $\boldsymbol{X}^{(n)}$ is stored in cell is which enclosed it. For each realization of the reverse process $\boldsymbol{Y}^{(m)}$ we consider only sample pairs $(\boldsymbol{X}^{(n)}, \boldsymbol{Y}^{(k)})$ where $\boldsymbol{X}^{(n)}$ and $\boldsymbol{Y}^{(k)}$ are in the same grid cell or in neighboring cells. Figure 5.13 illustrates the example of a two-dimensional grid.

Of course, some cells in the grid contain no particles and it makes no sense to store them. In case of a high-dimensional system the number of empty cells is quite large and saving of such a grid requires a lot of memory. To reduce the memory needed for the grid hash-tables can be used. Further discussion about using hash-tables for the forward-reverse estimator can be found in [10].

Next we consider the parallelization of the calculations of the forward-reverse estimator. Three different approaches to particle decomposition were considered and implemented.

1. Each processor gets the same amount of particles assigned for the forward as well for the reverse simulation and simulates the random walk of the forward and the reverse particles. Minimal and maximal positions of the forward simulation particles are evaluated for all processors and the corresponding grid is constructed and broadcasted to all processors. The positions of the reverse particles are then broadcasted to all processors and the kernel is estimated locally. In the end, the resulting local kernel estimations are brought together to make the overall result.

2. Each processor is either computing forward or reverse particles. Minimal and maximal positions of forward simulation particles are exchanged among forward processors. The corresponding grid is calculated and sent to all forward processors. Reverse particle positions are broadcasted and the kernel is estimated locally. Finally, the resulting local kernels are brought together.

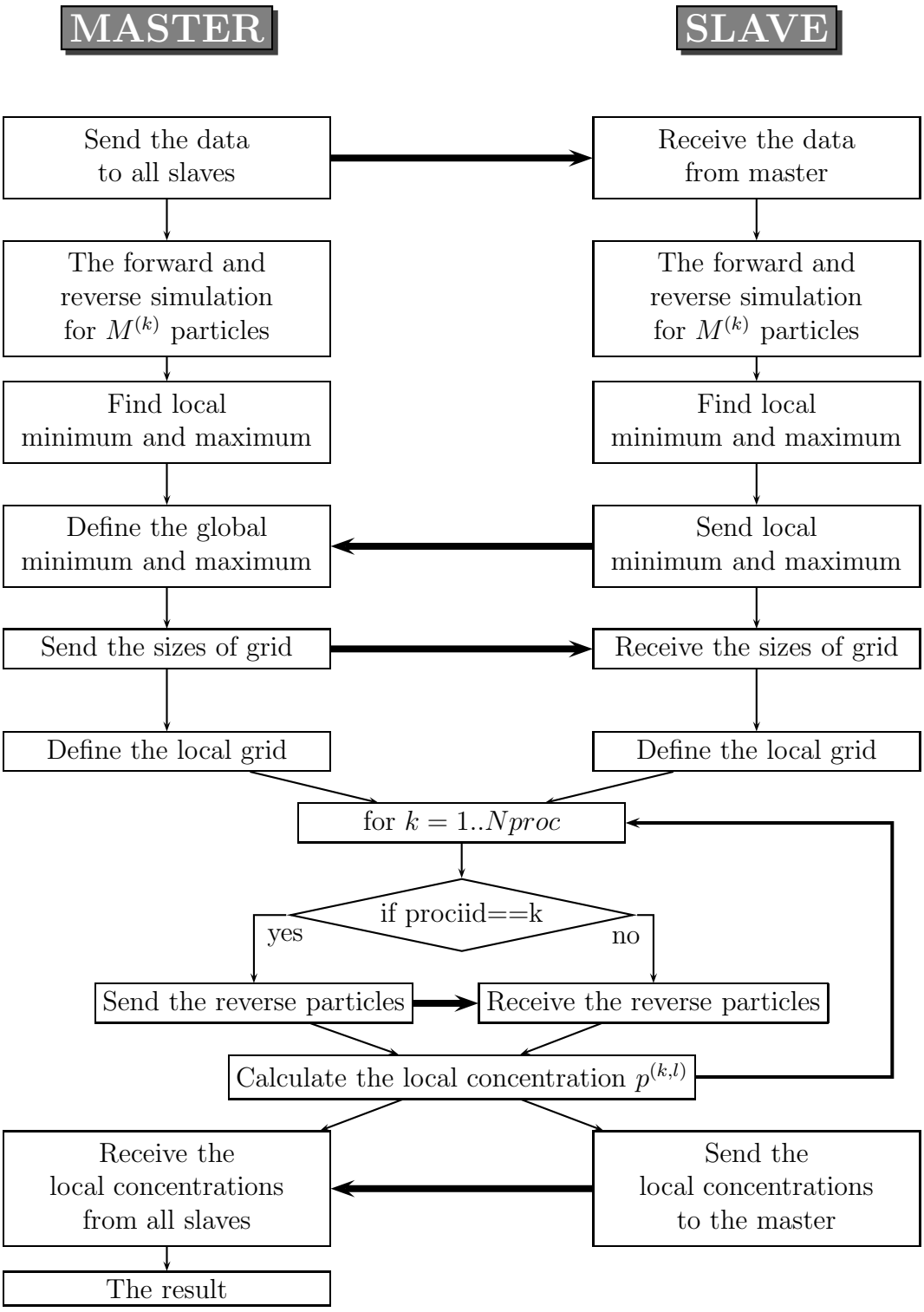
3. One of the processors is assigned to simulate the reverse particles while all the other processors run the forward simulation. The reverse processor keeps simulating while the forward processors exchanges grid information.Then the positions of the reverse particles are broadcasted and followed by the estimation of the kernel. The more particles are used for the simulation the more runs can be made to balance the work load.

Comparing the three different approaches, the first one seems to have the fewest drawbacks. In the second approach the problem that slows down the calculation is that the reverse processors are lazy while the forward processors estimate the kernel, which means that half of the processors are idle during the kernel estimation. In the third approach it is hard to determine how many runs must be made to balance the work load equally. This changes with the number of particles and must be adaptive. The first version seems to be the simplest and since all processors are running the same program it is therefore expected to perform the best.

As it has been done for the FE/RE estimator the forward reverse estimator (5.16) can be rewritten as

$$\widehat{p}(t,\boldsymbol{x},T,\boldsymbol{y}) = \frac{1}{N^2\lambda^d}\sum_{k=1}^{Nproc}\sum_{l=1}^{Nproc}\left(\sum_{n=1}^{M^{(l)}}\sum_{m=1}^{M^{(k)}} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(t^*)-\overline{\boldsymbol{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T)}{\lambda}\right)\overline{\mathcal{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T)\right) = \sum_{k=1}^{Nproc}\sum_{l=1}^{Nproc} p^{(k,l)}, \tag{5.25}$$

where $p^{(k,l)} = \dfrac{1}{N^2\lambda^d}\displaystyle\sum_{n=1}^{M^{(l)}}\sum_{m=1}^{M^{(k)}} K\left(\frac{\overline{\boldsymbol{X}}_{t,\boldsymbol{x}}^{(n)}(t^*)-\overline{\boldsymbol{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T)}{\lambda}\right)\overline{\mathcal{Y}}_{t^*,\boldsymbol{y}}^{(m)}(T)$



In the above described algorithm all the processors are running the same program which makes it particularly easy to implement and the load is equally distributed. The particles are distributed equally over all the processors. Forward as well as reverse simulation is run on each processor. The number of processors and particles can therefore be changed independently without any modifications in the code. It's the simplest approach, but through its simplicity it performs the best. Also the computation on each processor

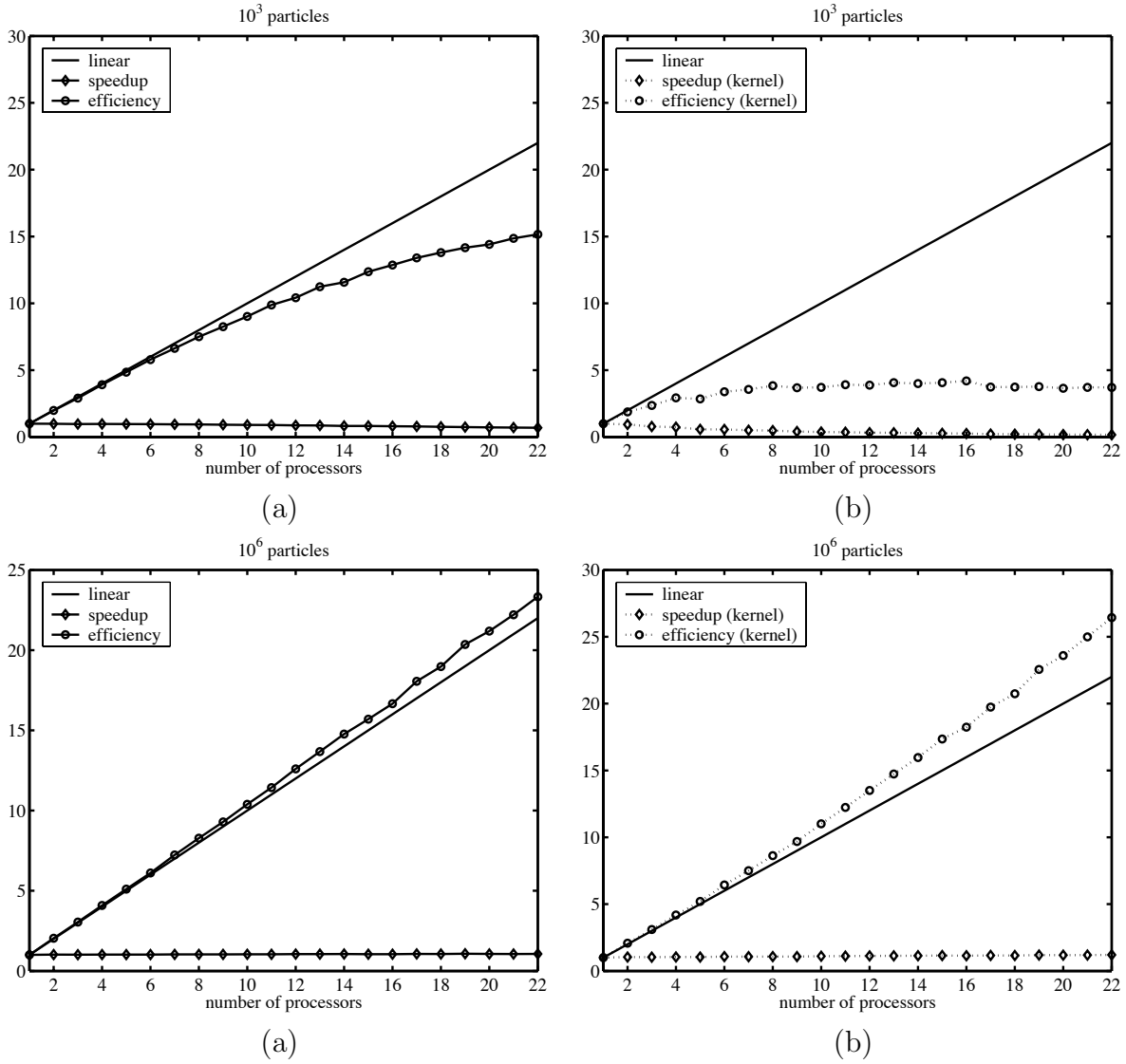


Figure 5.14: The speedup and efficiency of the FRE run on DAS2 for (a) 10^3 particles ((b) only for calculating the double sum) and for (c) 10^6 particles ((d) only for calculating the double sum)

takes about the same amount of time as they are all running the same piece of code. Only when the positions of the forward and reverse simulation particles are known, communication between the processors starts. In case of a long simulation this approach performs especially well. A grid is computed relying on the forward simulation positions. This grid is sent to all processors. Now the position of the reverse simulation is introduced into the grid. This happens by broadcasting the reverse particles. The estimation can now be computed on each processor locally. Then the results are summed together into a global kernel estimation.

To investigate the behavior of the parallelization with more processors the parallelization was run on the DAS2 cluster. As it can be seen from Figure 5.14 the results run on DAS2 are consistent with the previous ones. With 1000 particles the parallelization of the simulation is not very effective and efficiency decreases quickly as the number of processors increases.

Because there are no communications between processes during the forward and reverse simulations, we are interested also in the speedup of a pure kernel estimator. Figure 5.14 also illustrates the speedup for the kernel simulation. When increasing the number of particles to 10^6 a linear speedup can be observed, so the parallelization shows to be very effective. Better results are achieved when the number of particles is increased to a million. There is even superlinear speedup which happens due to some cache effects.

5.8 Conclusion

In this chapter the concept of reverse time diffusion was introduced. The reverse-time diffusion can be used to reduce computational time in a number of applications, for instance, to construct the risk maps for certain locations along the coastline. In this chapter the concept of reverse time diffusion was used to derive an efficient algorithm for estimating particle concentrations in coastal waters. This FRE approach is especially attractive for simulating concentrations caused by a calamity at sea and if one, in addition, is only interested in the results only at a few critical locations. For these problems the use of reverse time particle tracks reduces the computational effort to determine the concentration compared with the classical particle approach at least an order of magnitude. The parameter $t^* \in [t, T]$, the time where the forward and reverse-time tracks meet each other, has to be chosen with care. From a more theoretical point of view t* is optimal if the variance of the forward system is equal to the variance of the reverse system. From the practical point of view it is also important that the weighting coefficients $\mathcal{Y}$ are all close to 1. As a result t^* may have to be chosen closer to T.

Chapter 6

Two-particle models for the estimation of the mean and standard deviation of concentrations in coastal waters

6.1 Introduction

One of the advantages of random walk particle model is that it is a natural way to study not only the mean ensemble concentration, but also higher order moments of the concentration. For instance, the standard deviation of the concentration is connected with the statistics of the trajectories of pairs of particles. The idea of using two-particle simulation to obtain the standard deviation of the concentration was first formulated by Durbin [35]. More recently a number of papers in which Lagrangian models have been applied to investigate the standard deviation of the concentration have been published [12, 64, 75, 126].

In this chapter the multiple particle model is formulated and the forward reverse estimator is applied for the estimation of the mean ensemble concentration and the standard deviation of the concentration of the pollutant at a number of given critical locations. The multiple particle models is introduced in Section 6.2. In Section 6.3 some properties of the distribution of the concentration are discussed for a test problem. Finally, in Section 6.4 we describe an application of the two-particle model and the forward reverse estimator to calculate the mean and standard deviation of the concentration of a pollutant in the Dutch coastal waters.

This chapter is based on articles:

Spivakovskaya D., Heemink A.E. and Schoenmakers J.G.M. 2007 'Two-particle models for the estimation of the mean and standard deviation of concentrations in coastal waters', Stochastic Environmental Resources and Risk Assessment, 21(3), pp.235-253.

Spivakovskaya D., Heemink A.E. and Schoenmakers J.G.M. 2005 'A two particle model for the estimation of the mean and standard deviation of concentrations in coastal waters'. In: Proceedings of International Symposium of Stochastic Hydraulics 2005.

6.2 Two-particle model

6.2.1 Multiple particle model

In the previous section we considered one-particle model of pollutant transport in shallow water, which in general form can be written as follows:

$$\begin{aligned} d\boldsymbol{X}(s) &= \left(\boldsymbol{u} + \frac{\partial D}{\partial \boldsymbol{y}} + \frac{D}{H}\frac{\partial H}{\partial \boldsymbol{y}}\right) ds + \sqrt{2D}\boldsymbol{I}_2 d\boldsymbol{B}(s) \\ \boldsymbol{X}(t) &= \boldsymbol{x} \end{aligned} \tag{6.1}$$

This random walk model does not take into account the spatial correlation of the particle behavior and, therefore, can be applied only for estimating the ensemble mean concentration. However, the turbulent behavior of particles is correlated in space. Thus, the actual concentration at certain locations may be much higher or lower than the ensemble mean concentration. For instance, the ensemble mean concentration may be an average of a large number of zeros (realizations when the cloud of pollutant do not reach the location) and a few very large values. This type of averaging may be meaningless, because the few high concentrations may kill the organisms in a certain area and the large number of zeros can not bring them to life again.

So, if we want to describe the dispersion process in more detail we must use multiple particle models [4]. In a K-particle model, K particles are released at the same time and their behavior is correlated with each other. The correlation between any two particles is assumed to depend only on the distance between them. As it has been mentioned before the turbulence diffusion is correlated in space. It means the closer the particles are located from each other, the less difference in their movement we can observe. Over large distances particles behave almost independently, but when the distance is close to zero the behavior of the particles is highly correlated. As a result, the problem of particles separation when the distance is close to zero arises. According to Thompson [126] we should include the effect of molecular diffusion. In one-particle models the effect of molecular diffusion is negligible as compared with the turbulent diffusion, but in multiple particle models the molecular diffusion plays a significant role. Moreover, if the particles coincide, they can only be separated again by molecular diffusion. As a result, we propose to write the equation (6.1) in the form

$$\begin{aligned} d\boldsymbol{X}(s) &= \left(\boldsymbol{u} + \frac{\partial D}{\partial \boldsymbol{y}} + \frac{D}{H}\frac{\partial H}{\partial \boldsymbol{y}}\right) ds + \sqrt{2D}\boldsymbol{I}_2 \left(\sqrt{1-\beta^2} d\boldsymbol{B}(s) + \beta d\boldsymbol{W}(s)\right) \\ \boldsymbol{X}(t) &= \boldsymbol{x} \end{aligned} \tag{6.2}$$

where $0 \leq \beta \leq 1$ is constant, $\boldsymbol{B}$ is a standard Brownian motion as in the one-particle model, and $\boldsymbol{W}$ is a two-dimensional correlated Brownian motion independent of $\boldsymbol{B}$. In (6.2), $\sqrt{1-\beta^2}\boldsymbol{B}(s)$ represents the diffusion caused by the molecular diffusion and the small scale turbulence and $\beta\boldsymbol{W}(s)$ represents large scale turbulence.

Now we are ready to introduce the K-particle model. The behavior of K particles can be described by the following $2K$-dimensional systems of the stochastic differential

equations

$$\begin{aligned} d\boldsymbol{X}^{[i]}(s) &= \left(\boldsymbol{u}^{[i]} + \frac{\partial D}{\partial \boldsymbol{y}^{[i]}} + \frac{D}{H^{[i]}}\frac{\partial H^{[i]}}{\partial \boldsymbol{y}^{[i]}}\right) ds + \sqrt{2D}\boldsymbol{I}_2 \left(\sqrt{1-\beta^2} d\boldsymbol{B}^{[i]}(s) + \beta d\boldsymbol{W}^{[i]}(s)\right) \\ \boldsymbol{X}^{[i]}(t) &= \boldsymbol{x} \end{aligned} \tag{6.3}$$

With the superscript $[i]$, $i = 1, \ldots, K$ we indicate the functions that depend on the ith particle $\boldsymbol{X}^{[i]}$

$$\begin{aligned} \boldsymbol{u}^{[i]} &= \boldsymbol{u}(s, \boldsymbol{X}^{[i]}) \\ H^{[i]} &= H(\boldsymbol{X}^{[i]}) \end{aligned}$$

The Brownian motion processes $\boldsymbol{B}^{[i]}(s)$ are mutually independent, while any pair processes $\boldsymbol{W}^{[i]}$ and $\boldsymbol{W}^{[j]}$, $(i \neq j)$ is correlated with covariance matrix

$$E\left(\boldsymbol{W}^{[i]}(ds)\boldsymbol{W}^{[j]}(ds)^T\right) = f(r_{ij})\boldsymbol{I}_2 ds \tag{6.4}$$

where $f(r)$ is a covariance function that depends on the distance r between particles $\boldsymbol{X}^{[i]}$ and $\boldsymbol{X}^{[j]}$ and

$$r_{ij} = \|\boldsymbol{X}^{[i]} - \boldsymbol{X}^{[j]}\|$$

The covariance function is related to the spectrum of the turbulence. We assume that the correlation function f satisfies several conditions. First of all, we assume that this function is sufficiently smooth, at least its second derivative is continuous and bounded. We further assume that

1. $f(0) = 1$,
2. $f(r) \to 0$, $r \to \infty$ (monotonically),
3. The matrix $[f(\|\boldsymbol{y}^i - \boldsymbol{y}^j\|)]_{1 \leq i \leq q, 1 \leq j \leq q}$ is positive definite for any choice of points $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^q$ in $\mathbb{R}^2$.

For example, we can use the following function [22]

$$f(r) = \exp(-\alpha r^2), \quad \alpha > 0 \tag{6.5}$$

6.2.2 The standard deviation of the concentration

Multiple particle models may be used to find the distribution of the concentration. The K-particle model allows to find the kth moment of the distribution of the concentration for all $1 \leq k \leq K$. However, it is too time consuming to apply the multiple particles model in numerical applications. If we want to find only the mean concentration and the standard deviation of the concentration distribution we can apply the two-particle model. The behavior of a pair of particles is simulated by Equation (6.3) for $K = 2$. The probability density function $p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}^{[1]}, \boldsymbol{y}^{[2]})$ (or the joint probability function of the stochastic processes $\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$) gives us the information about the standard deviation of the concentration.

Let us consider the neighborhood of the point $\boldsymbol{y}$

$$O_\epsilon(\boldsymbol{y}) = \{(\tilde{y}_1, \tilde{y}_2) : |\tilde{y}_1 - y_1| < \epsilon/2, |\tilde{y}_2 - y_2| < \epsilon/2\}$$

The probability that two particles $\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$ will occur in $O_\epsilon(\boldsymbol{y})$ is given by $P(t, \boldsymbol{x}, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y}))$. The value

$$\frac{1}{\nu^2(O_\epsilon(\boldsymbol{y}))} P(t, \boldsymbol{x}, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y})) = \frac{1}{\epsilon^4} P(t, \boldsymbol{x}, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y})) \tag{6.6}$$

is the concentration of the pairs of particles $(\boldsymbol{X}^{[1]}, \boldsymbol{X}^{[2]})$ in $O_\epsilon(\boldsymbol{y})$. Here $\nu(O_\epsilon(\boldsymbol{y}))$ is the area of the neighborhood $O_\epsilon(\boldsymbol{y})$. When the particles ($\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$) are independent this concentration can be found as

$$\left(\frac{1}{\nu(O_\epsilon(\boldsymbol{y}))} P(t, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y}))\right)^2 = \frac{1}{\epsilon^4} P^2(t, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y})) \tag{6.7}$$

The difference between the values (6.6) and (6.7)

$$\overline{c^2}(T, O_\epsilon(\boldsymbol{y})) = \frac{1}{\epsilon^4} \left| P(t, \boldsymbol{x}, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y}), O_\epsilon(\boldsymbol{y})) - P^2(t, \boldsymbol{x}; T, O_\epsilon(\boldsymbol{y})) \right| \tag{6.8}$$

gives us the information about the fluctuation of the concentration in $O_\epsilon(\boldsymbol{y})$. Proceed to the limit when $\epsilon \to 0$ we will receive the value $\overline{c^2}(T, \boldsymbol{y})$ at the point $\boldsymbol{y}$

$$\overline{c^2}(T, \boldsymbol{y}) = \left| p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}, \boldsymbol{y}) - p^2(t, \boldsymbol{x}; T, \boldsymbol{y}) \right| \tag{6.9}$$

In [35] it is shown that the fluctuation defined by (6.9) leads to a slightly lower concentration fluctuation than the usual definition and can be used as a measure of the fluctuation amplitude. Taking into account that we use the depth averaged model and that the concentration of the pollutant is connected with the density probability function as $C(T, \boldsymbol{y}) = p(t, \boldsymbol{x}; T, \boldsymbol{y})/H(\boldsymbol{y})$ ($H(\boldsymbol{y})$ is a depth at location $\boldsymbol{y}$) we define the standard deviation $Dev(T, \boldsymbol{y})$ at the point $\boldsymbol{y}$ at time T as

$$Dev(T, \boldsymbol{y}) = \frac{1}{H(\boldsymbol{y})} \sqrt{\overline{c^2}(T, \boldsymbol{y})} = \sqrt{\left| \frac{p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}, \boldsymbol{y})}{H^2(\boldsymbol{y})} - C^2(T, \boldsymbol{y}) \right|} \tag{6.10}$$

In Section 6.3 we study some properties of the concentration distribution for a relatively simple example.

6.3 Test problem

In this section first we study the pure dispersion process with zero velocity vector $\boldsymbol{u} = 0$, constant water depth $H = 1\,\mathrm{m}$ and constant dispersion coefficient $D = 5\,\mathrm{kg/m^3}$. The domain of interest (see Figure 6.1) is scaled with the grid size ΔX. The constant $\beta = \sqrt{0.9}$ (see Equation (6.2)).

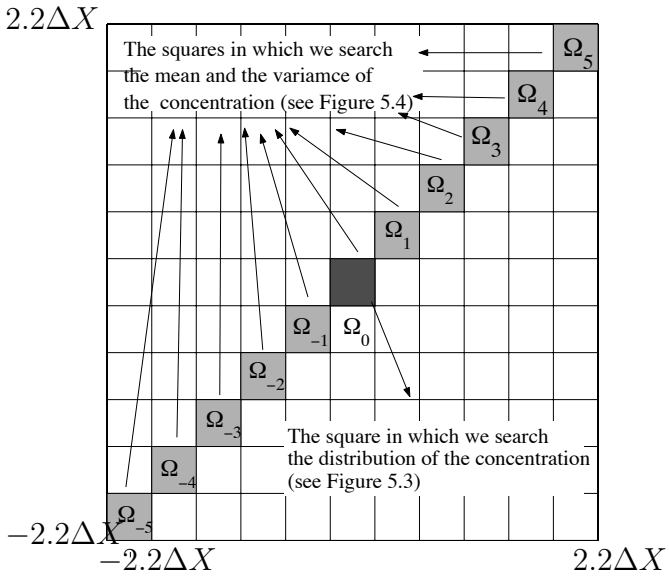


Figure 6.1: The domain of the test problem

6.3.1 100-particle model

Figure 6.2 shows an example of a simulation of the dispersion process for a 100-particle model after 1, 3 and 7 days (Figures (a), (c), (e)) with the correlation function (6.2), $\alpha = 5/(\Delta X * \Delta X)$ (see Equation (6.5)) and an one-particle model with 100 realizations (Figures 6.2 (b), (d), (f)). If we repeat the simulation many times and then average, the center of the cloud of particles will be in the origin for both the multiple and the one-particle models. However, from Figures 6.2(a) and 6.2(c) it is seen that the center of the individual clouds can diverge from the origin. If we wait long enough the center of each cloud will tend to return to the origin (Figure 6.2(e)).

Suppose we want to know the distribution of the concentration in the square Ω_0= $[-0.2\Delta X,\ 0.2\Delta X] \times [-0.2\Delta X, 0.2\Delta X]$ (see Figure 6.1) after 2.5 days. We release 100 particles in the origin $\boldsymbol{x} = \mathbf{0}$ and fix the number of particles that occur in this square after 2.5 days. Figures 6.3(a), (b) and (c) show the concentration distribution for different values of parameter α ($\alpha = 1/(\Delta X * \Delta X)$, $\alpha = 5/(\Delta X * \Delta X)$ and $\alpha = 10/(\Delta X * \Delta X)$). In Figure 6.3 (d) all these distributions are shown together and can be compared with each other. It is clear that with increasing the parameter α, the number of observations with "zero" concentration decreases and, therefore, the standard deviation of the concentration goes down (see figure 6.3(d)). For the distant regions this is not true. However, in this case, the number of experiments is not enough to provide accurate results. Therefore, the disagreement of the results for the distant regions can be explained by the computational error.

Usually one is only interested in the ensemble mean and the standard deviation of the concentration. The diffusion process in this test case is symmetric in all directions (it

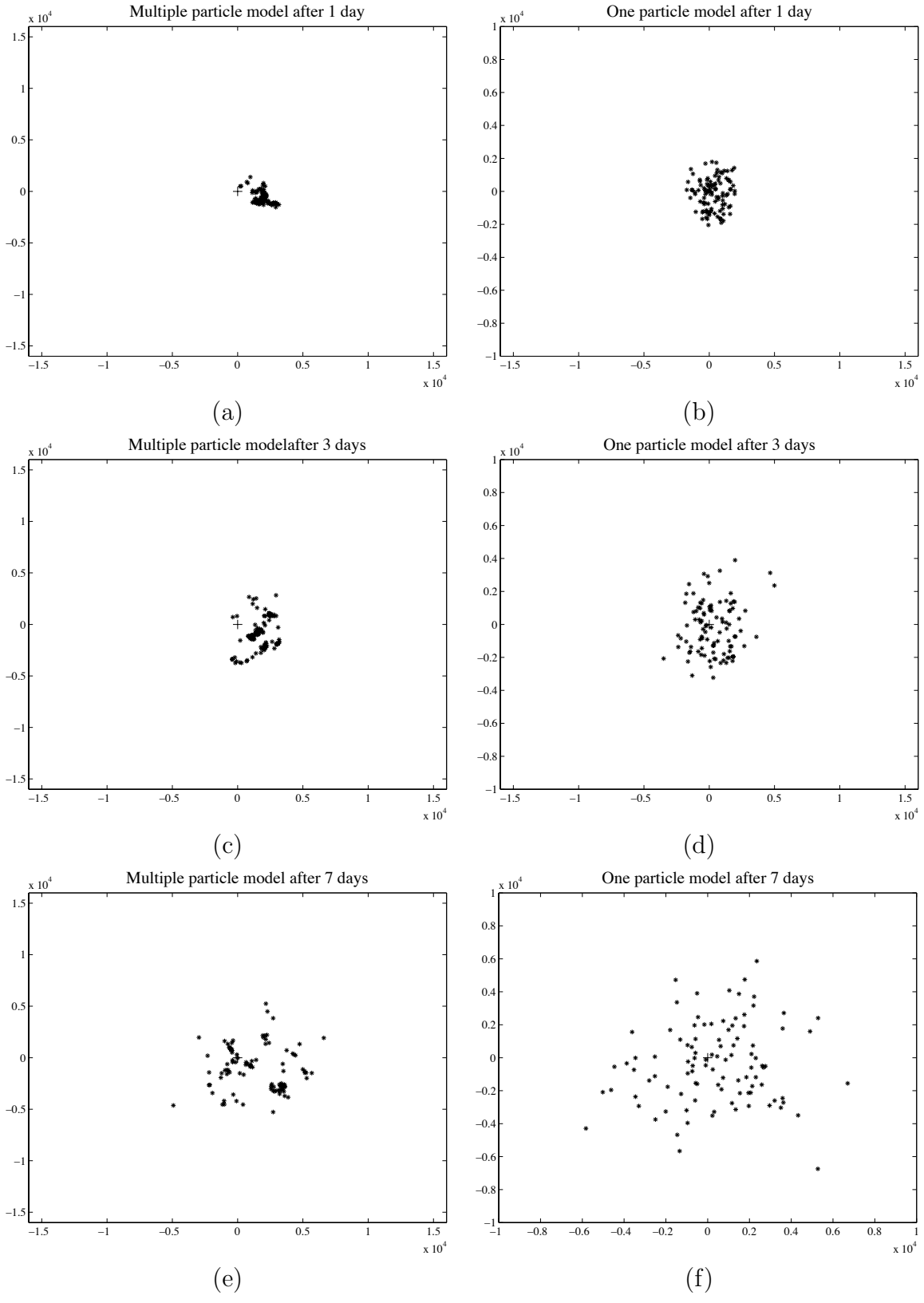


Figure 6.2: The multiple particle and one-particle models

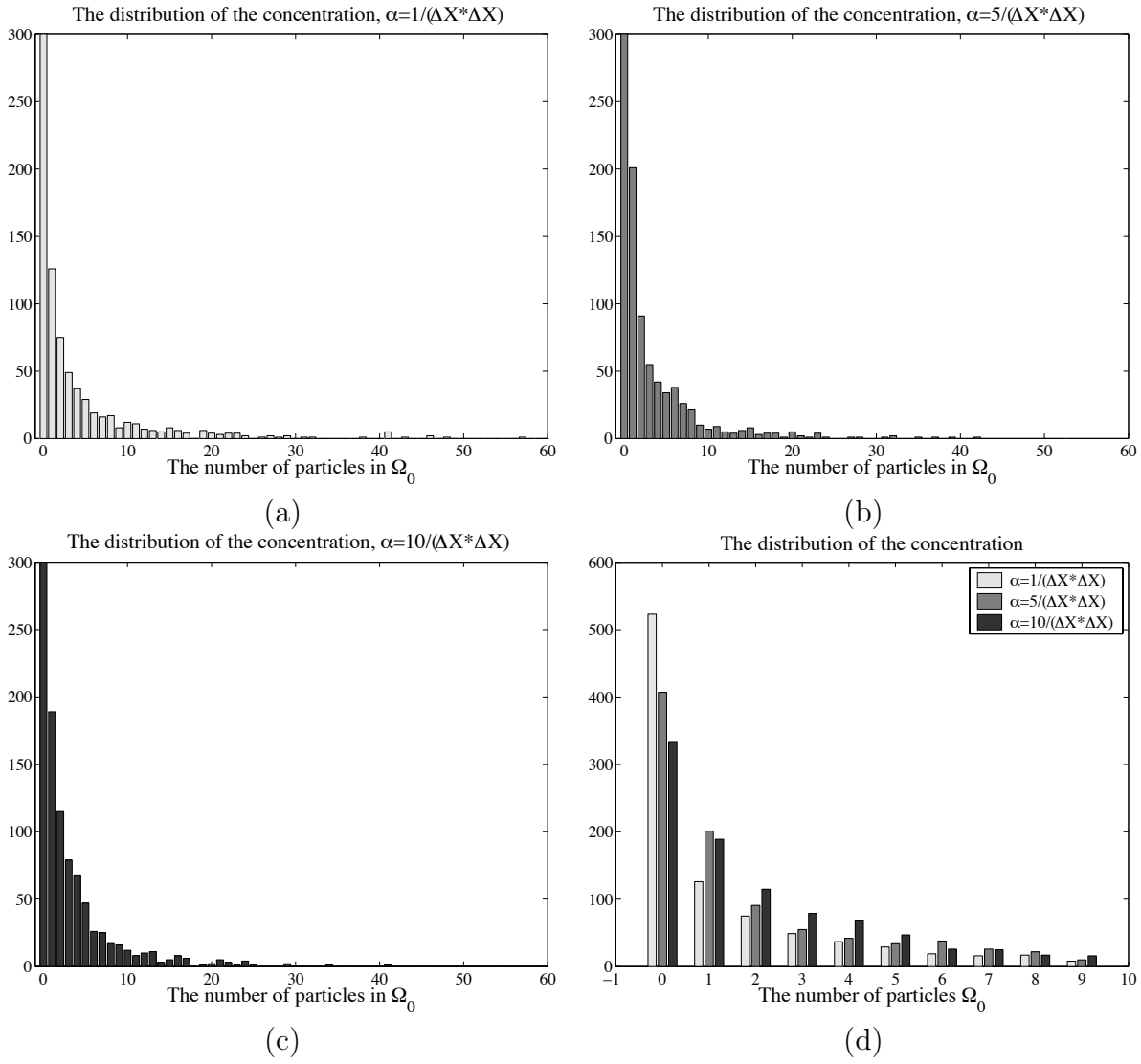


Figure 6.3: The distribution of the concentration

means that the concentration at some point depends only on the distance between this point and the origin). Therefore, we study the mean and the standard deviation of the concentration not in the whole domain, but only in the squares Ω_i, $i = -5, \ldots, 5$.

Figure 6.4(a) shows the ensemble mean of the concentration distribution. For instance, we can see that in average, after a number of observations, there will be 3 particles in the region Ω_0. However, if we want to predict how many particles will occur in this region in the next simulation we should take into account the standard deviation of the concentration distribution (see Figure 6.4(b)). Let us consider the case $\alpha = 1/(\Delta X * \Delta X)$: the standard deviation of the concentration in Ω_0 is almost 7. Therefore, we can expect that a number of particles will be somewhere between 0 and 10.

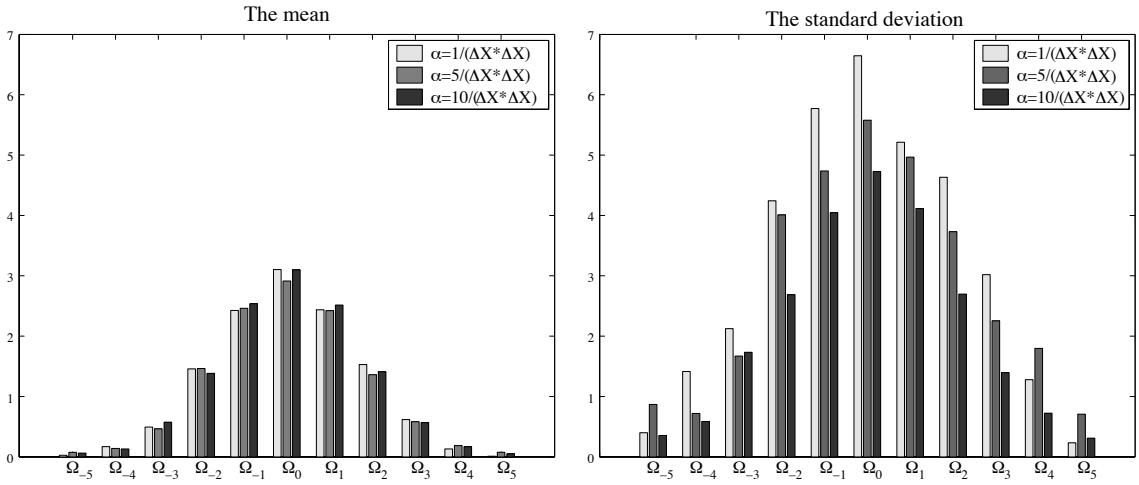


Figure 6.4: The mean and the standard deviation

When the value of the parameter α increases, the standard deviation decreases. It is a quite natural phenomena, because for large α the particles separate easier. We also expect that the standard deviation of the concentration distribution converges to zero when the time increases and the difference between independent and correlated simulations disappears, as it was suggested in the literature (see, for example, [43], [87]). To show this effect the following experiment has been considered.

We fix the parameter $\alpha = 10/(\Delta X * \Delta X)$, but we change the size of the central area Ω_0 in such way, that in average 5 particles from 100 will occur in this area. We consider 4 time steps

- T=2.5 days
  $\Omega_0 = [-0.26\Delta X, 0.26\Delta X] \times [-0.26\Delta X, 0.26\Delta X]$

- T=5 days
  $\Omega_0 = [-0.37\Delta X, 0.37\Delta X] \times [-0.37\Delta X, 0.37\Delta X]$

- T=7.5 days
  $\Omega_0 = [-0.45\Delta X, 0.45\Delta X] \times [-0.45\Delta X, 0.45\Delta X]$

- T=10 days
  $\Omega_0 = [-0.52\Delta X, 0.52\Delta X] \times [-0.52\Delta X, 0.52\Delta X]$

and for these moments of time we compare the distribution of the concentration for correlated and independent simulations (see Figure 6.5). We have repeated each experiment 100 times. All distributions in Figure 6.5 have the same mean, but, in case of multiple particle model, this mean is the result of many 'zero' concentrations and a few very high

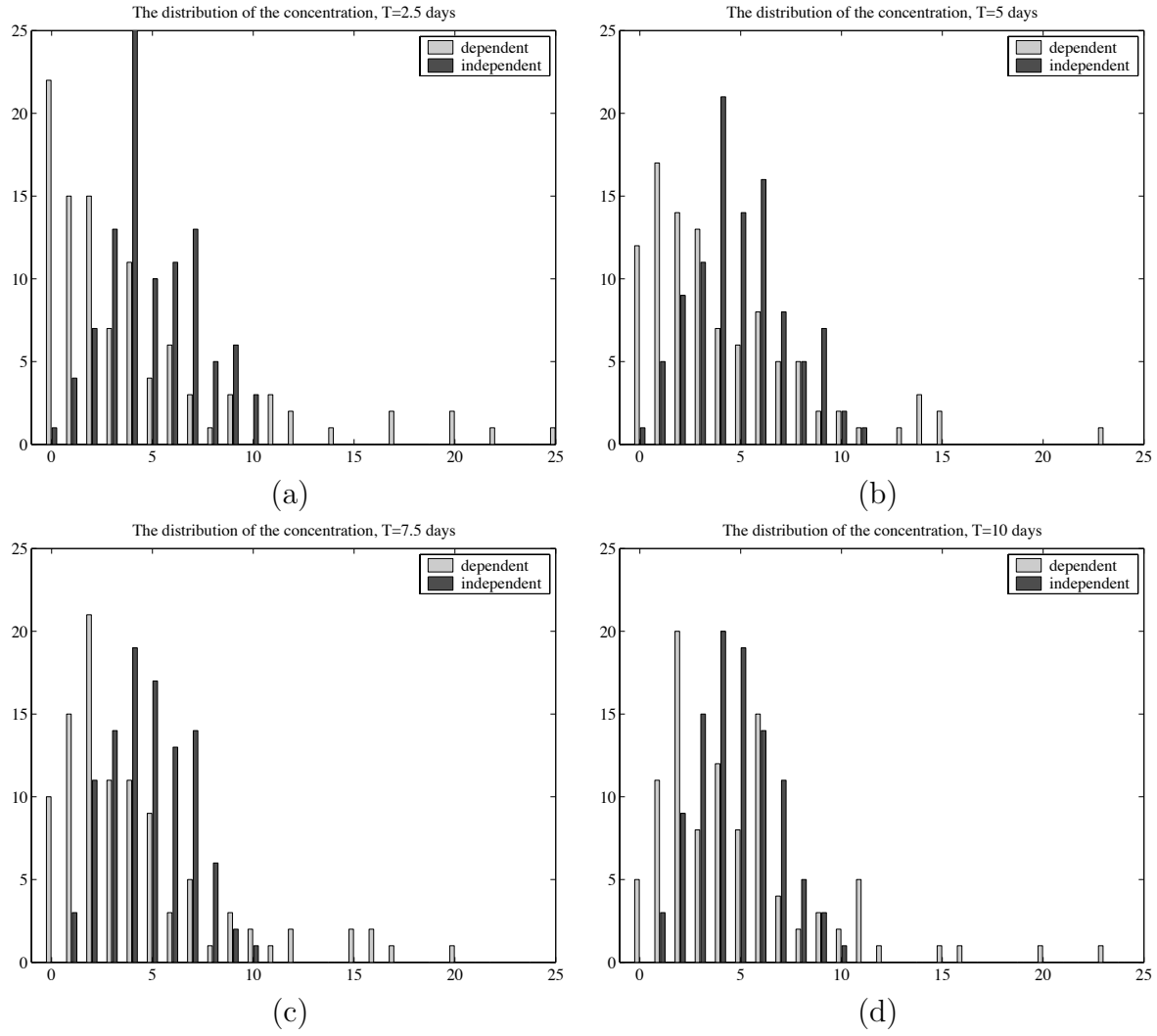


Figure 6.5: The distribution of the concentration with different time steps

concentrations. From Figure 6.5 it can be seen that the number of the 'zero' concentrations goes down with the time and the difference between multiple particle model and one-particle model disappears.

6.3.2 One- and two-particle models

In practice one does not need to start the multiple particle simulation to find the ensemble mean concentration and the concentration deviation. It is enough to consider only one- or two-particle models. The one-particle model has the following form

$$\begin{aligned} d\boldsymbol{X}(s) &= \sqrt{2D}\boldsymbol{I}_2 d\boldsymbol{B}(s) \\ \boldsymbol{X}(t) &= \boldsymbol{x} \end{aligned} \tag{6.11}$$

where the diffusion coefficient D is assumed to be constant. Equation (6.11) can be solved exactly and the density function $p(t, \boldsymbol{x}; T, \boldsymbol{y})$ of the stochastic process $\boldsymbol{X}(T)$ is a normal distribution with mean $\boldsymbol{x}$ and variance $2D(T-t)$

$$p(t, \boldsymbol{x}; T, \boldsymbol{y}) = \frac{1}{4\pi D(T-t)} exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{4D(T-t)}\right) \tag{6.12}$$

We will consider the averaged concentration in the domain Ω_0 defined as

$$\begin{aligned} C(T, \Omega_0) = \frac{1}{\mu(\Omega_0)} \int\limits_{\boldsymbol{y}\in\Omega_0} C(T, \boldsymbol{y}) d\boldsymbol{y} = \frac{1}{\nu(\Omega_0)} \int\limits_{\boldsymbol{y}\in\Omega_0} \frac{p(t, \boldsymbol{x}; T, \boldsymbol{y})}{H(\boldsymbol{y})} d\boldsymbol{y} = \\ \frac{1}{\nu(\Omega_0)} \int\limits_{\boldsymbol{y}\in\Omega_0} p(t, \boldsymbol{x}; T, \boldsymbol{y}) d\boldsymbol{y} = p(t, \boldsymbol{x}; T, \Omega_0) \end{aligned} \tag{6.13}$$

where $\mu(\Omega_0)$ is the area of Ω_0 and the water depth $H(\boldsymbol{y}) \equiv 1\,\mathrm{m}$ is taken to be constant in our test problems. The exact value of the integral in the equation (6.13) can be found from standard tables. We can use the forward estimator to find $p(t, \boldsymbol{x}; T, \Omega_0)$ in the following way (λ)

$$\widehat{p}(t, \boldsymbol{x}; T, \Omega_0) = \frac{1}{N\lambda^2} \sum_{n=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{x}} - \boldsymbol{\eta}^{(n)}}{\lambda}\right) \tag{6.14}$$

where $\eta^{(n)}$, $n = 1, \ldots, N$ are independent random numbers, uniformly distributed in the square Ω_0. In a similar way we can rewrite the forward-reverse estimator (5.16)

$$\widehat{p}(t, \boldsymbol{x}; T, \Omega_0) = \frac{1}{N^2\lambda^2} \sum_{n,m=1}^{N} K\left(\frac{\overline{\boldsymbol{X}}^{(n)}_{t,\boldsymbol{x}} - \overline{\boldsymbol{Y}}^{(m)}_{t^*,\boldsymbol{\eta}^{(m)}}(T)}{\lambda}\right) \overline{\mathcal{Y}}^{(m)}_{t^*,\boldsymbol{\eta}^{(m)}}(T) \tag{6.15}$$

We consider a two-particle forward system of the form

$$\begin{aligned} &d\boldsymbol{Z}(s) = \sqrt{2D}\boldsymbol{I}_4 \left(\sqrt{1-\beta^2} d\boldsymbol{B}(s) + \beta d\boldsymbol{W}(s)\right) \\ &\boldsymbol{Z}(t) = (\boldsymbol{x}, \boldsymbol{x})^T \end{aligned} \tag{6.16}$$

where $\boldsymbol{Z} = (\boldsymbol{X}^{[1]}, \boldsymbol{X}^{[2]})^T$, $\boldsymbol{B} = (\boldsymbol{B}^{[1]}, \boldsymbol{B}^{[2]})^T$, $\boldsymbol{W} = (\boldsymbol{W}^{[1]}, \boldsymbol{W}^{[2]})^T$. The Brownian motions $\boldsymbol{W}^{[1]}$ and $\boldsymbol{W}^{[2]}$ are assumed to be correlated with covariance matrix (6.4). From here on we will use the notation r instead of r_{12} to denote the distance between particles $\boldsymbol{X}^{[1]}$ and $\boldsymbol{X}^{[2]}$. The reverse time system corresponding to (6.16) has the following form

$$\begin{aligned} &d\boldsymbol{R}(s) = d\begin{pmatrix} \boldsymbol{Y}^{[1]}(s) \\ \boldsymbol{Y}^{[2]}(s) \end{pmatrix} = \begin{pmatrix} 2D\beta^2 \dfrac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[2]}} \\ 2D\beta^2 \dfrac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[1]}} \end{pmatrix} ds + \sqrt{2D}\boldsymbol{I}_4(\sqrt{1-\beta^2} d\boldsymbol{B}(s) + \beta d\boldsymbol{W}(s)) \\ &d\gamma(s) = 2D\beta^2 \left(\frac{\partial^2 f}{\partial \tilde{y}_1^{[1]} \partial \tilde{y}_1^{[2]}} + \frac{\partial^2 f}{\partial \tilde{y}_2^{[1]} \partial \tilde{y}_2^{[2]}}\right) \gamma ds \\ &\boldsymbol{R}(t^*) = (\boldsymbol{y}, \boldsymbol{y})^T,\ \gamma(t^*) = 1 \end{aligned} \tag{6.17}$$

method	N	$p(t, \boldsymbol{x}; T, \boldsymbol{y})$
FE	10^4	0.1879 ± 0.0131
FE	10^5	0.1895 ± 0.0070
FE	10^6	0.1880 ± 0.0025
FRE	10^3	0.1875 ± 0.0050
FRE	10^4	0.1889 ± 0.0016
FRE	10^5	0.1885 ± 0.0007

Table 6.1: The results of the one-particle model, the exact value $p(t, \boldsymbol{x}, T, \boldsymbol{y}) = 0.1886$ ($t = 0$, $T = 1$, $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{0}$)

It should be noted that the reverse-time system introduced in Section 5.2 is based on a system of uncorrelated Brownian motions. It is not difficult to see that the reverse system has an equivalent representation with respect to correlated Brownian motions which is given by Equation (6.17) for this example.

The deviation of the concentration in the area Ω_0 at time T is defined as (we use the fact that $H \equiv 1\text{m}$)

$$Dev(T, \Omega_0) = \sqrt{|p(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0) - p^2(t, \boldsymbol{x}; T, \Omega_0)|} \tag{6.18}$$

where

$$p(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0) = \frac{1}{\nu^2(\Omega_0)} \int_{\Omega_0} \int_{\Omega_0} p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}^{[1]}, \boldsymbol{y}^{[2]}) d\boldsymbol{y}^{[1]} d\boldsymbol{y}^{[2]} \tag{6.19}$$

Similarly to the one-particle model we can use the forward and the forward-reverse estimators to find the value of the joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0)$

$$\widehat{p}_{FE}(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0) = \frac{1}{N\lambda^4} \sum_{n=1}^{N} K\left(\frac{\overline{\boldsymbol{Z}}^{(n)}_{t,\boldsymbol{x},\boldsymbol{x}}(T) - \boldsymbol{\zeta}^{(n)}}{\lambda} \right) \tag{6.20}$$

$$\widehat{p}_{FRE}(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0) = \frac{1}{N^2\lambda^4} \sum_{n,m=1}^{N} K\left(\frac{\overline{\boldsymbol{Z}}^{(n)}_{t,\boldsymbol{x},\boldsymbol{x}}(t^*) - \overline{\boldsymbol{R}}^{(m)}_{t^*,\boldsymbol{\zeta}^{(m)}}(T)}{\lambda} \right) \gamma_{t^*,\boldsymbol{\zeta}^{(m)}}(T) \tag{6.21}$$

where $\boldsymbol{\zeta}^{(n)} \in \mathbb{R}^4$, $n = 1, \ldots, N$ are random numbers uniformly distributed in $\Omega_0 \times \Omega_0$. In Tables 6.1-6.5 the results of the one- and two-particle model are shown. The time of the simulation is 2.5 days for all experiments. Each experiment was repeated 30 times in order to find the error due to the numerical simulation. In Tables 6.1 and 6.3 one can see the results of the one-particle model for the point $\boldsymbol{y} = \boldsymbol{0}$ and for the area Ω_0 (see the previous section). In this case we know the solution of the stochastic differential equation (6.11) and can compare the numerical results with the exact solution. From these tables it can be seen that the forward-reverse estimator is at least two orders of magnitude more

method	N	$p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}, \boldsymbol{y})$	$p(t, \boldsymbol{x}, \boldsymbol{x}, T; \boldsymbol{y}, \boldsymbol{y})$
		$\alpha = 1/(\Delta X * \Delta X)$	$\alpha = 5/(\Delta X * \Delta X)$
FE	10^5	0.2405 ± 0.0259	0.2303 ± 0.0218
FE	10^6	0.2468 ± 0.0163	0.2622 ± 0.0118
FE	10^7	0.2520 ± 0.0066	0.2933 ± 0.0070
FRE	10^3	0.2546 ± 0.0209	0.3138 ± 0.4193
FRE	10^4	0.2568 ± 0.0070	0.3715 ± 0.1077
FRE	10^5	0.2604 ± 0.0023	0.3567 ± 0.0353

Table 6.2: The joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}, \boldsymbol{y})$ ($t = 0$, $T = 1$, $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{0}$)

method	N	$p(t, \boldsymbol{x}; T, \Omega_0)$
FE	10^4	0.1797 ± 0.0141
FE	10^5	0.1859 ± 0.0077
FE	10^6	0.1844 ± 0.0030
FRE	10^3	0.1863 ± 0.0051
FRE	10^4	0.1853 ± 0.0020
FRE	10^5	0.1857 ± 0.0006

Table 6.3: The results of the one-particle model, the exact value $p(t, \boldsymbol{x}; T, \Omega_0) = 0.1856$ ($t = 0$, $T = 1$, $\boldsymbol{x} = \boldsymbol{0}$)

method	N	$p(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0)$	$p(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0)$
		$\alpha = 1/(\Delta X * \Delta X)$	$\alpha = 5/(\Delta X * \Delta X)$
FE	10^5	0.1606 ± 0.0232	0.1239 ± 0.0157
FE	10^6	0.1662 ± 0.0128	0.1360 ± 0.0092
FE	10^7	0.1705 ± 0.0070	0.1367 ± 0.0067
FRE	10^3	0.1682 ± 0.0244	0.1250 ± 0.1953
FRE	10^4	0.1705 ± 0.0049	0.1250 ± 0.0585
FRE	10^5	0.1725 ± 0.0017	0.1440 ± 0.0185

Table 6.4: The joint probability function $p(t, \boldsymbol{x}, \boldsymbol{x}; T, \Omega_0, \Omega_0)$ ($t = 0$, $T = 1$, $\boldsymbol{x} = \boldsymbol{0}$)

	in the point $\boldsymbol{y}$		in the area Ω_0	
α	$C(T,\boldsymbol{y})$	$Dev(T,\boldsymbol{y})$	$C(T,\boldsymbol{y})$	$Dev(T,\boldsymbol{y})$
$1/(DX*DX)$	0.1886	0.4737	0.1856	0.3682
$5/(DX*DX)$	0.1886	0.5142	0.1856	0.3249

Table 6.5: The concentration and the standard deviation of the concentration

accurate than the pure forward estimator. One can also see that the value of the function $p(t,\boldsymbol{x};T,\boldsymbol{y})$ does not differ too much from $p(t,\boldsymbol{x};T,\Omega_0)$.

In Tables 6.2 and 6.4 the results of the two-particle model are presented for different value of the parameter α. There are two aspects of the approach that we will discuss in the next two sections. Firstly, it can be seen from these tables that the forward-reverse estimator is much more efficient (three orders of magnitude) than the forward estimator when the value of parameter α is relatively small, but for a large value of α the forward-reverse estimator is not so efficient. We will discuss this in more detail in the next section. Secondly, if one compares the joint probability function at the point $p(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$ and the average value of this function on Ω_0, then it can be seen that these values differ significantly. Moreover, the results from the table 6.5 show that the deviation in the point $Dev(T,\boldsymbol{y})$ grows with larger values of α, while the averaged value in Ω_0 of deviation $Dev(T,\Omega_0)$ decreases. This behavior will be discussed in more detail in the next section.

6.3.3 The definition of the standard deviation at a point

In this section we discuss some properties of the joint probability function $p(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y}^{[1]},\boldsymbol{y}^{[2]})$, in particular, how this function depends on the choice of the correlation function $f(\rho)$. Figure 6.6 (a) represents the correlation function (6.5) for $\alpha = 1/(\Delta X*\Delta X)$ and for $\alpha = 5/(\Delta X*\Delta X)$. It is clear that the larger α is the faster the correlation function goes down and as a result the particles become independent much faster. Therefore it is natural to expect that the deviation of the concentration decreases when α increases, but the results of the tables 6.2, 6.4 and 6.5 show the opposite. To explain this behavior we consider the joint probability function $p(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y}^{[1]},\boldsymbol{y}^{[2]})$ in some neighborhood of the point $\boldsymbol{y}$. Let us introduce the following function

$$\overline{C(\boldsymbol{y})C(\boldsymbol{y};r)} = p(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{u})\mid_{\|\boldsymbol{y}-\boldsymbol{u}\|=r} \tag{6.22}$$

Function (6.22) can be interpreted as the dependence between the concentration at a point $\boldsymbol{y}$ and at a location at a distance r from this point $\boldsymbol{y}$. In Figure 6.6 (b) the function $\overline{C(\boldsymbol{y})C(\boldsymbol{y};r)}$ is shown for $\alpha = 1/(\Delta X*\Delta X)$ and for $\alpha = 5/(\Delta X*\Delta X)$. The value of this function for $\alpha = 5/(\Delta X*\Delta X)$ at $r=0$ is larger than for $\alpha = 1/(\Delta X*\Delta X)$. This function decreases very fast and, as a consequence, it is not enough to study this function when $r=0$, but also in some neighborhood of this point. We propose to define

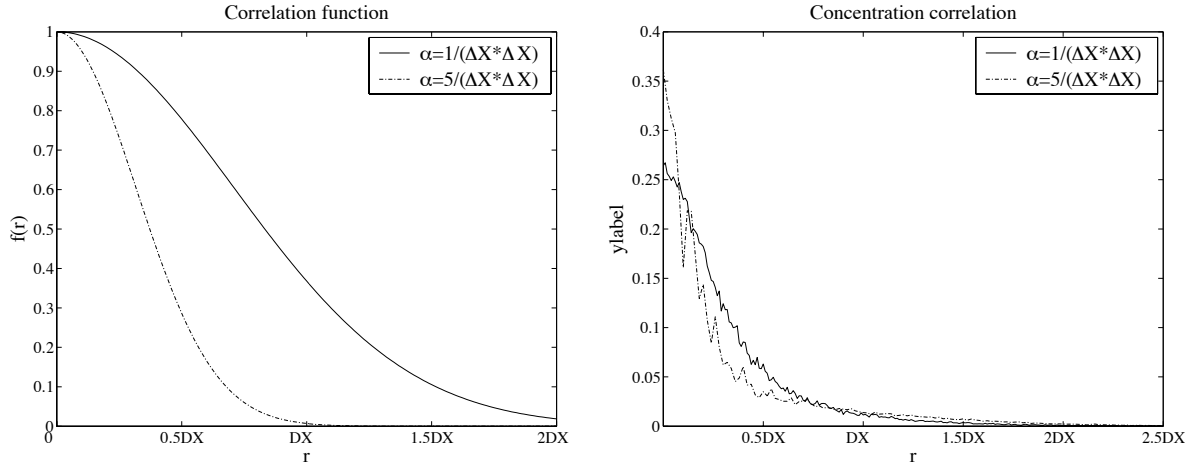


Figure 6.6: (a) The correlation function and (b) the correlation of the concentration

the deviation at the point $\boldsymbol{y}$ at time T $\widehat{Dev}(T, \boldsymbol{y})$ as follows

$$\widehat{Dev}(T, \boldsymbol{y}) = \frac{1}{H(\boldsymbol{y})} \left| \int\limits_{\mathbb{R}^2} \int\limits_{\mathbb{R}^2} p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{u}, \boldsymbol{v}) g(\boldsymbol{u}) g(\boldsymbol{v}) d\boldsymbol{u} d\boldsymbol{v} - \left(\int\limits_{\mathbb{R}^2} p(t, \boldsymbol{x}; T, \boldsymbol{u}) g(\boldsymbol{u}) d\boldsymbol{u} \right)^2 \right|^{\frac{1}{2}} \tag{6.23}$$

where

$$g(\boldsymbol{u}) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{\|\boldsymbol{y} - \boldsymbol{u}\|^2}{2\sigma^2} \right) \tag{6.24}$$

The function $g(\boldsymbol{u})$ is a density function of a 2-dimensional normal distribution with the vector of means $\boldsymbol{y}$ and covariance matrix $\sigma^2 \boldsymbol{I}_2$. We can use the forward and forward-reverse estimators (2.51) and (5.16) to find the integral

$$\widehat{p}(t, \boldsymbol{x}; T, \boldsymbol{y}) = \int\limits_{\mathbb{R}^2} p(t, \boldsymbol{x}; T, \boldsymbol{u}) g(\boldsymbol{u}) d\boldsymbol{u}$$

assuming that $\boldsymbol{\eta}^{(n)}$, $n = 1, \ldots, N$ are independent realizations of the random variable $\boldsymbol{\eta}$ with density function $g(\boldsymbol{u})$. In a similar way we can use the formulae (6.20) and (6.21) to find the integral

$$\widehat{p}(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{y}, \boldsymbol{y}) = \int\limits_{\mathbb{R}^2} \int\limits_{\mathbb{R}^2} p(t, \boldsymbol{x}, \boldsymbol{x}; T, \boldsymbol{u}, \boldsymbol{v}) g(\boldsymbol{u}) g(\boldsymbol{v}) d\boldsymbol{u} d\boldsymbol{v}$$

In this case we assume that $\boldsymbol{\zeta}^{(n)}$, $n = 1, \ldots, N$ are independent realizations of the 4-dimensional random variable $\boldsymbol{\zeta}$, normally distributed with the vector of means $(\boldsymbol{y}, \boldsymbol{y})'$ and the covariance matrix $\sigma^2 \boldsymbol{I}_4$. The value of the parameter σ^2 should be relatively small as

method	N	$\widehat{p}(t,\boldsymbol{x};T,\boldsymbol{y})$	$\widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$	$\widehat{Dev}(T,\boldsymbol{y})$	$\widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$	$\widehat{Dev}(T,\boldsymbol{y})$
			$\alpha = 1/(\Delta X * \Delta X)$		$\alpha = 5/(\Delta X * \Delta X)$	
FE	10^5	0.1801	0.1074 ± 0.0162	0.2738	0.0822 ± 0.0126	0.2231
FE	10^6	0.1801	0.1161 ± 0.0091	0.2892	0.0866 ± 0.0079	0.2327
FE	10^5	0.1801	0.1146 ± 0.0060	0.2866	0.0847 ± 0.0050	0.2286
FRE	10^3	0.1801	0.1142 ± 0.0154	0.2859	0.0788 ± 0.0749	0.2153
FRE	10^4	0.1801	0.1148 ± 0.0047	0.2870	0.0871 ± 0.0446	0.2338
FRE	10^5	0.1801	0.1166 ± 0.0014	0.2901	0.0882 ± 0.0173	0.2361

Table 6.6: The concentration and the standard deviation of the concentration

compared with the variance of the stochastic process $\boldsymbol{X}(s)$ at time T. For example, in our simulation $\sigma^2 = 0.04\Delta X^2$, while the variance of $\boldsymbol{X}(T)$ is $2D(T-t) \approx 0.85\Delta X^2$.

It can be easily shown using the equation (6.12) and the properties of density function of a normally distributed random variable that

$$\widehat{p}(t,\boldsymbol{x};T,\boldsymbol{y}) = \frac{1}{2\pi(2D(T-t)+\sigma^2)} \exp\left(\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{2D(T-t)+\sigma^2}\right)$$

The results for the function $\widehat{p}(t,\boldsymbol{x},T,\boldsymbol{y})$ and for the function $\widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$ are presented in Table 6.6. Comparing this results with the results from tables 6.3.2 and 6.2 it can be seen that the differences between the functions $p(t,\boldsymbol{x};T,\boldsymbol{y})$ and $\widehat{p}(t,\boldsymbol{x};T,\boldsymbol{y})$ are not very big, while the functions $p(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$ and $\widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$ differ very strongly.

6.3.4 The implementation of the FRE for the two-particle model

As it was mentioned before the accuracy of the forward-reverse estimator strongly depends on the correlation function $f(r)$, in particular on the parameter α. The difficulties arise with the weighting coefficient $\mathcal{Y}$. First of all, the differential equation for the scalar function $\mathcal{Y}(s)$ is very sensitive to the choice of the time step Δt of the numerical integration of the reverse time system (6.17). This problem is discussed in detail in [111]. As a consequence we must choose the time step h sufficiently small. For example, one can compare the results for time step $\Delta t = 300$s and for time step $\Delta t = 30$s ($N = 10^4$, $\alpha = 5/(\Delta X * \Delta X)$, $t^* = 0.5(T-t)+t$)

$$\begin{array}{ll} \Delta t = 300s & \widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y}) = 0.0871 \pm 0.0446 \\ \Delta t = 30s & \widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y}) = 0.08435 \pm 0.0367 \end{array}$$

Unfortunately, decreasing the time step we also increase the time of the simulation and this affects the efficiency of the forward-reverse estimator. A possible solution is to compare the results of the simulations with time steps h and $2h$ (using the Talay-Tubaro method

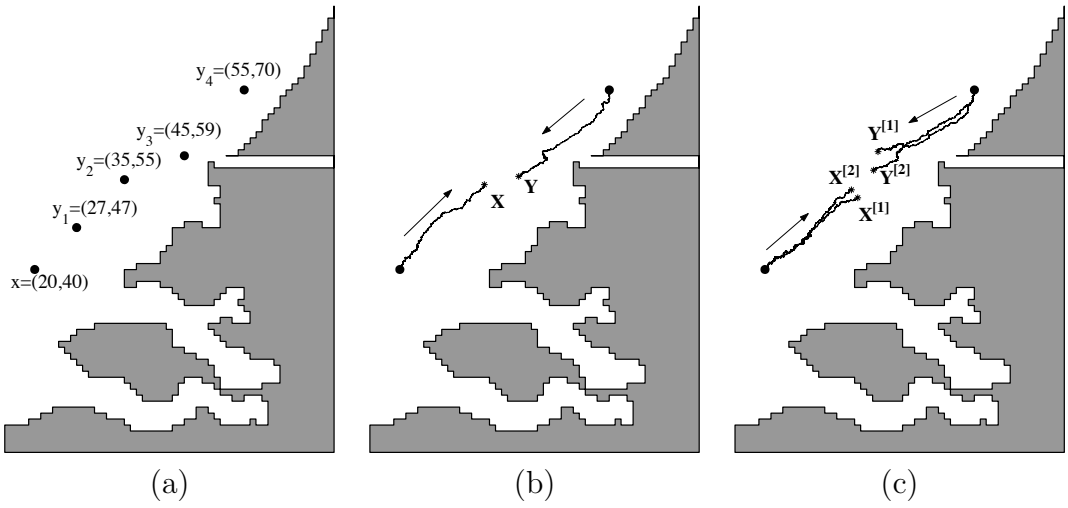


Figure 6.7: (a) The locations of interest. The example of the forward and reverse time simulation for the (b) one-particle model and for (c) two-particle model

we repeat the simulation with double time step) and, if the solutions of the reverse time system (6.17) differ very much, we do not take into consideration the results of these particles.

Another problem occurs when the weighting coefficient $\mathcal{Y}(T)$ varies very much for different particles. For example, when $\alpha = 5/(DX * DX)$ the values $\mathcal{Y}(T)$ can be around 300 and the contribution of this particle will dominate in the sum (5.16). What we can do here is to throw away the particles with "very large" coefficient. For instance, one can see the results of the forward-reverse estimator including all particles and without the particles with the coefficient $\mathcal{Y}$ larger than 250

all particles	$\widehat{p} = 0.0871 \pm 0.0446$
throw away particles	$\widehat{p} = 0.0782 \pm 0.0239$

These methods may help in many situations, but when the correlation function is too steep, the forward-reverse method will fail and we have to choose the internal point t^* closer to the end time T.

6.4 Application

In this section we study the mean ensemble concentration and the deviation of the concentration in some critical locations along the Dutch seaside using one- and two-particle models. Furthermore, we apply the forward-reverse estimator for both models and compare the results with the results of the classical forward estimator.

The forward system for tidally averaged numerical flow (see Section 5.4) has the fol-

method	N	$\widehat{p}(t,\boldsymbol{x},T,\boldsymbol{y})$	$\widehat{p}(t,\boldsymbol{x},T,\boldsymbol{y})$
		$\boldsymbol{y}=(27,47)\quad T=2.5$ days	$\boldsymbol{y}=(35,55)\quad T=5$ days
FE	10^4	0.1218 ± 0.0089	0.0398 ± 0.0040
FE	10^5	0.1211 ± 0.0054	0.0389 ± 0.0017
FE	10^6	0.1223 ± 0.0024	0.0392 ± 0.0010
FRE	10^3	0.1220 ± 0.0057	0.0392 ± 0.0018
FRE	10^4	0.1213 ± 0.0015	0.0393 ± 0.0007
FRE	10^5	0.1216 ± 0.0006	0.0393 ± 0.0002

Table 6.7: The results for the one-particle model

lowing form

$$\begin{aligned} &d\boldsymbol{Z}(s)=d\begin{pmatrix}\boldsymbol{X}^{[1]}\\ \boldsymbol{X}^{[2]}\end{pmatrix}=\begin{pmatrix}\boldsymbol{u}^{[1]}+\frac{D}{H^{[1]}}\frac{\partial H^{[1]}}{\partial \boldsymbol{y}^{[1]}}\\ \boldsymbol{u}^{[2]}+\frac{D}{H^{[2]}}\frac{\partial H^{[2]}}{\partial \boldsymbol{y}^{[2]}}\end{pmatrix}ds+\\ &\sqrt{2D}\boldsymbol{I}_4\begin{pmatrix}\sqrt{1-\beta^2}d\boldsymbol{B}^{[1]}(s)+\beta d\boldsymbol{W}^{[1]}(s)\\ \sqrt{1-\beta^2}d\boldsymbol{B}^{[2]}(s)+\beta d\boldsymbol{W}^{[2]}(s)\end{pmatrix}\\ &\boldsymbol{Z}(t)=(\boldsymbol{x},\boldsymbol{x})^T\end{aligned} \tag{6.25}$$

The reverse time system that corresponds to the two-particle forward system (6.25) can be written as follows:

$$\begin{aligned} &d\boldsymbol{R}=\begin{pmatrix}d\boldsymbol{Y}^{[1]}\\ d\boldsymbol{Y}^{[2]}\end{pmatrix}=\begin{pmatrix}-\boldsymbol{u}^{[1]}-\frac{D}{H^{[1]}}\frac{\partial H^{[1]}}{\partial \tilde{\boldsymbol{y}}^{[1]}}+2D\beta^2\frac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[2]}}\\ -\boldsymbol{u}^{[2]}-\frac{D}{H^{[2]}}\frac{\partial H^{[2]}}{\partial \tilde{\boldsymbol{y}}^{[2]}}+2D\beta^2\frac{\partial f}{\partial \tilde{\boldsymbol{y}}^{[1]}}\end{pmatrix}ds+\\ &\sqrt{2D}\boldsymbol{I}_4\begin{pmatrix}\sqrt{1-\beta^2}d\boldsymbol{B}^{[1]}(s)+\beta d\boldsymbol{W}^{[1]}(s)\\ \sqrt{1-\beta^2}d\boldsymbol{B}^{[2]}(s)+\beta d\boldsymbol{W}^{[2]}(s)\end{pmatrix}\\ &d\gamma(s)=\Big(\varphi_1(s,\boldsymbol{Y}^{[1]})+\varphi_2(s,\boldsymbol{Y}^{[2]})+\varphi_3(s,\boldsymbol{R})\Big)\gamma(s)ds\\ &\boldsymbol{R}(t^*)=(\boldsymbol{y},\boldsymbol{y})',\quad \gamma(t^*)=1,\end{aligned} \tag{6.26}$$

Here

$$\varphi_i(s,\boldsymbol{Y}^{[i]})=-\frac{\partial}{\partial(\tilde{\boldsymbol{y}}^{[i]})^T}\left(\boldsymbol{u}^{[i]}+\frac{D}{H^{[i]}}\frac{\partial H^{[i]}}{\partial\partial\tilde{\boldsymbol{y}}^{[i]}}\right),\quad i=1,2$$

represents the contribution of the flow and the function

$$\varphi_3(s,\boldsymbol{R})=2D\beta^2\left(\frac{\partial^2 f}{\partial\tilde{y}_1^{[1]}\partial\tilde{y}_1^{[2]}}+\frac{\partial^2 f}{\partial\tilde{y}_2^{[1]}\partial\tilde{y}_2^{[2]}}\right)$$

represents the contribution of the correlation between particles.

We suppose that the contaminant is released in the location $(20,40)$ and consider some critical locations in which we want to know the ensemble mean concentration and the standard deviation of the concentration (see Figure 6.7 (a)). In Figure 6.7 (b) one can see the example of the forward and reverse time simulation (one particle model) and

method	N	$\widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$	$\widehat{p}(t,\boldsymbol{x},\boldsymbol{x};T,\boldsymbol{y},\boldsymbol{y})$
		$\boldsymbol{y}=(27,47)\quad T=2.5$ days	$\boldsymbol{y}=(35,55)\quad T=5$ days
FE	10^5	0.0594 ± 0.0086	0.0125 ± 0.0015
FE	10^6	0.0648 ± 0.0049	0.0126 ± 0.0014
FE	10^7	0.0653 ± 0.0027	0.0126 ± 0.0067
FRE	10^3	0.0645 ± 0.0096	0.0119 ± 0.0076
FRE	10^4	0.0662 ± 0.0034	0.0123 ± 0.0023
FRE	10^5	0.0659 ± 0.0010	0.0126 ± 0.0010

Table 6.8: The results for two-particle model

Figure 6.7 (c) shows an example of the forward and reverse time simulation for a pair of particles during 10 days. The internal point point t^* is chosen in the middle of the time interval $[t,T]$.

In Tables 6.7 and 6.8 the results for the first moment (the probability density function) and the second moment of the particle distribution (the joint probability function of process $\boldsymbol{Z}(T)$) respectively are shown. We compare the forward and forward-reverse estimators for two locations.

From Table 6.7 we can see that the mean ensemble concentration calculated by the forward-reverse estimator using 10^3 particles is $C(T,\boldsymbol{y}) = 6.7\cdot10^{-6}\pm0.3\cdot10^{-6}$ kg/m^3. The deviation $\pm0.3\cdot10^{-6}$ kg/m^3 is caused by statistical error of the estimation process and converges to zero with increasing the number of particles. On the contrary, the standard deviation $\widehat{Dev} = 1.8\cdot10^{-5}$ kg/m^3 as calculated in Table 6.8, is the result of the spatial correlation of the turbulence and does not depend on the number of particles.

The results for the concentration and the standard deviation for different locations along the Dutch coast are shown in Table 6.9. Here we used the forward-reverse estimator with number of particles $N = 10^5$. The results show that even 10 days after the accident the concentration fluctuation is still very large and should be taken into account in order to assess the impact of the calamity.

$\boldsymbol{y}$	T, days	t^*	$\widehat{C}(T,\boldsymbol{y})$	$\widehat{Dev}(T,\boldsymbol{y})$
$(27,47)$	2.5	$0.5T$	$2.15*10^{-5}$	$4*10^{-5}$
$(35,55)$	5	$0.5T$	$6.7*10^{-6}$	$1.8*10^{-5}$
$(45,59)$	7.5	$0.6T$	$2.8*10^{-6}$	$7.1*10^{-6}$
$(55,70)$	10	$0.7T$	$2.3*10^{-6}$	$6.1*10^{-6}$

Table 6.9: The concentration $\widehat{C}(T,\boldsymbol{y})$, kg/m^3 and the standard deviation of the concentration $\widehat{Dev}(T,\boldsymbol{y})$, kg/m^3

6.5 Conclusion

In this chapter we studied two-particle models for computing the mean and standard deviation of the concentration in the Dutch coastal zone using the forward-reverse estimator. The results show that the actual concentration may become much higher than the ensemble mean concentration as computed by the traditional transport models. As a consequence, for providing an accurate prediction of the spreading of the pollutant, we need to use two-particle models and take into account the spatial correlation of the turbulence.

The results also show that the forward-reverse estimator is at least two orders of magnitude more efficient than the classical pure forward estimator.

Chapter 7

The backward Îto method for the Lagrangian simulation of transport processes with large space variations of the diffusivity

7.1 Introduction

For space-varying diffusivity the advection part of the random walk model requires an additional correction term, which is equal to the diffusivity gradient. Because of this correction term the particles do not accumulate in regions of low diffusivity [58, 141, 98]. This random walk model can be introduced by using the theory of stochastic differential equations (SDE) [52, 31, 118, 114]. The advection-diffusion equation is interpreted as the Fokker-Planck equation [91] and the corresponding SDE in Îto sense can be derived. As a result, the particle track is simulated by a stochastic process, whose transition density function coincides with the tracer concentration. The Îto formulation is not the only one way to introduce the particle tracking model. Another random walk model based on Stratonovich stochastic calculus is also quite popular.

Unfortunately, the standard random-walk methods for simulating transport can only be applied when the diffusivity is sufficiently smooth, otherwise the correction term in the advection part dominates the flow velocity. In many situations the rapid but continuous change in turbulence statistics that occurs may be represented by a discontinuity [127]. Recently LaBolle et al. [78] proposed the random walk model based on backward Îto calculus that requires no corrective velocity.

In this chapter, we discuss the random walk models based on Îto, Stratonovich and backward Îto calculus. The backward Îto random walk model is seen to be appropriate for dealing with discontinuity in the diffusivity field. It is applied for two test cases, for which key properties of the solutions can be derived analytically.

This chapter is based on article:

Spivakovskaya D., Deleersnijder E. and Heemink A.W. 2007 'The backward Îto method for the Lagrangian simulation of transport processes with large space variations of the diffusivity', Ocean Science Discussion, 4, pp.623-654

7.2 The Îto, Stratonovich and the backward Îto random walk models

Let us consider the following one-dimensional advection-diffusion problem:

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}\left(uC - k\frac{\partial C}{\partial x}\right) \tag{7.1}$$

Here $C(t,x)$ is the concentration of a passive tracer, u is flow velocity and $k(t,x)$ is a diffusivity term. Equation (7.1) can be interpreted as a Fokker-Planck equation (see [65, 91]) and the corresponding stochastic differential equation in Îto sense can be considered

$$\textbf{(I)}\; dX(t) = (u + k'(x))\, dt + \sqrt{2k} dW(t) \tag{7.2}$$

The solution of the advection-diffusion problem (7.1) is then the probability density function of the stochastic process $X(t)$. The stochastic differential equation (7.2) actually is not a 'differential' equation, but can be interpreted as an integral equation

$$\textbf{(I)} \quad X(t) = x_0 + \int_0^t (u + k'(x))\, dt + \int_0^t \sqrt{2k} dW(t) \tag{7.3}$$

The first term (advection) in the right hand side of (7.3) is a standard Lebesgue integral, while the second part (diffusion) of (7.3) may be introduced as the limit of the sum ([78])

$$\textbf{(I)} \quad \int_0^t f(X,t) dW(t) = \text{ms-}\lim_{n\to\infty} \sum_{k=1}^{n} f(X(t_{k-1}), t_{k-1})[W(t_k) - W(t_{k-1})] \tag{7.4}$$

Here $0 = t_0 < t_1 < \ldots t_{n-1} = t_n = t$ and $\text{ms} - \lim$ denotes the limit in the mean square sense. In general, to define a unique stochastic integral one needs to specify at which point the function $f(X,t)$ is evaluated. For instance, in the definition of the Îto integral the function f is always evaluated at the beginning of subinterval $[t_{k-1}, t_k]$ rendering $f(X(t_{k-1}), t_{k-1})$ statistically independent of $[W(t_k) - W(t_{k-1})]$ and thus ensuring that the Îto integral has zero mean. One well-known alternative, the Stratonovich integral, may be defined as a limit of the sum in which the function is evaluated at the middle of the time interval

$$\textbf{(S)} \quad \int_0^t f(X,t) dW(t) = \text{ms-}\lim_{n\to\infty} \sum_{k=1}^{n} f\left(X(\frac{t_{k-1}+t_k}{2}), \frac{t_{k-1}+t_k}{2}\right)[W(t_k) - W(t_{k-1})] \tag{7.5}$$

The corresponding random walk model can be written as follows:

$$\textbf{(S)} \quad dX(t) = \left(u + \frac{1}{2}k'(x)\right) dt + \sqrt{2k} dW(t) \tag{7.6}$$

The random walk models in the Îto or Stratonovich sense contain the diffusivity gradient in the advection part. For problems with large space variations of the diffusivity, this

gradient may be very high and, therefore, dominates in the advection term. As a result, the solution obtained by the random walk model in the Îto or Stratonovich sense will not be accurate. To circumvent this diffusivity, one may have recourse to a random walk model that does not require a diffusivity gradient in the advection part. This formulation is based on the backward Îto integral (see [65, 78])

$$\textbf{(bI)} \quad \int_0^t f(X,t)dW(t) = \text{ms-}\lim_{n\to\infty}\sum_{k=1}^{n} f(X(t_k),t_k)[W(t_k)-W(t_{k-1})] \tag{7.7}$$

Using the backward Îto SDE for modelling advection-diffusion processes with discontinuous diffusivity was proposed by [78]. The corresponding random walk model may be written as follows:

$$\textbf{(bI)} \quad dX(t) = udt + \sqrt{2k}dW(t) \tag{7.8}$$

The sensitivity of the limit of the integral sums to the choice of location at which the function is evaluated is a consequence of the unbounded variation of the Wiener process ([65]). However, each of the random walk methods introduced above is consistent with the advection-diffusion equation (7.1). For a continuous diffusion term, all these methods provide the same solution of the equation.

7.3 Numerical integration of the stochastic differential equations

It can be shown from the advection-diffusion equation (7.1) (see [58]) that the mean and variance of the tracer cloud spread during time interval $(t, t+\Delta t)$ are given by

$$\begin{aligned} N_1 &= u^i\Delta t + k'(X^i)\Delta t \\ N_2 - N_1^2 &= 2k(X^i)\Delta t + o(\Delta t) \approx 2k(X^i)\Delta t \end{aligned} \tag{7.9}$$

Here N_i, $i = 1, 2$ denote the ith moment of the concentration. Now we show that the first two moments of the displacement $\Delta X^i = X^{i+1} - X^i$, $i = 1, \ldots, L$ in the random walk models (in Îto, Stratonovich and backward Îto senses) are the same as in (7.9).

Specific numerical schemes are associated with each of the stochastic methods mentioned above. For instance, the SDE in the Îto sense can be numerically integrated by applying the explicit Euler method:

$$X^{i+1} = X^i + u^i\Delta t + k'(X^i)\Delta t + \sqrt{2k(X^i)\Delta t}R^i \tag{7.10}$$

Here, $X^i = X(t_i)$, $u^i = u(t^i, X^i)$, $t_i = i\Delta t$, $i = 0, \ldots, L-1$, $\Delta t = t/L$ and R^i are mutually independent normally distributed random numbers with parameters $(0, 1)$. We need only to find the probability law of the solution $X(t)$ of the SDE (in other words, solution in the weak sense), but not to approximate the solution itself. For these purposes, it is not necessary to chose the normally distributed random variables. We can use any distribution with the same mean and variance, for instance, random numbers uniformly varying between $-\sqrt{3}$ and $\sqrt{3}$.

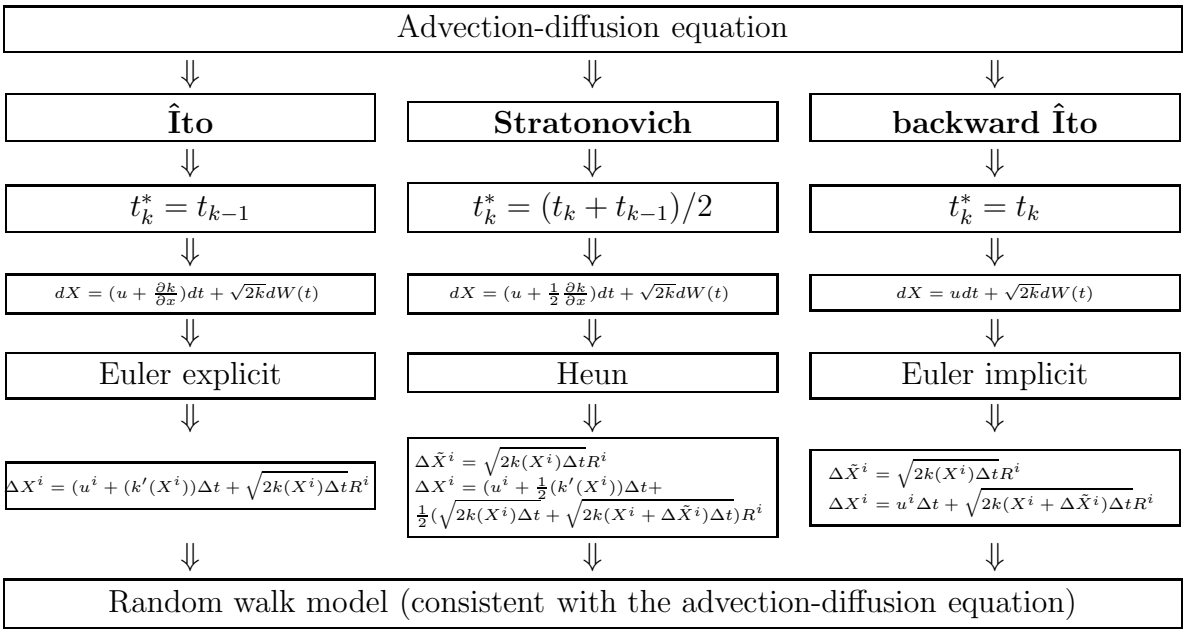


Figure 7.1: Comparison between Îto, Stratonovich and backward Îto formulations

The solution obtained by using the random walk model (7.10) has the same properties

$$\begin{aligned} &E(\Delta X^i) = E(u^i\Delta t + k'(X^i)\Delta t + \sqrt{2k(X^i)\Delta t}R^i) = u^i\Delta t + k'(X^i)\Delta t \\ &Var(\Delta X^i) = E\left(\Delta X^i - (E(\Delta X^i))\right)^2 = E(\sqrt{2k(X^i)\Delta t}R^i)^2 = 2k(X^i)\Delta t \end{aligned} \tag{7.11}$$

The Stratonovich formulation of the particle model can be approximated by using the Heun method

$$\begin{aligned} &\Delta\tilde{X}^i = \sqrt{2k(X^i)\Delta t}R^i) \\ &X^i = X^{i-1} + u^i\Delta t + \tfrac{1}{2}k'(X^i)\Delta t + \tfrac{1}{2}\left(\sqrt{2k(X^i)\Delta t} + \sqrt{2k(X^i + \Delta\tilde{X}^i)\Delta t}\right)R^i \end{aligned} \tag{7.12}$$

Let us consider the mean and variance of ΔX^i obtained by method (7.12)

$$\begin{aligned} &E(\Delta X^i) = E\left(u^i\Delta t + \tfrac{1}{2}k'(X^i)\Delta t + (\sqrt{2k(X^i)\Delta t} + \sqrt{2k(X^i + \Delta\tilde{X}^i)\Delta t})R^i\right) = \\ &u^i\Delta t + \tfrac{1}{2}k'(X^i)\Delta t + \tfrac{1}{2}E\left(\sqrt{2k(X^i + \Delta\tilde{X}^i)\Delta t}R^i\right) \end{aligned} \tag{7.13}$$

Let us expand the function $B(X^i + \Delta\tilde{X}^i) = \sqrt{2k(X^i + \Delta\tilde{X}^i)}$ into Taylor series

$$\begin{aligned} &B(X^i + \Delta\tilde{X}) = B(X^i) + B'(X^i)\Delta\tilde{X} + O(\Delta\tilde{X}^2) = \\ &\sqrt{2k(X^i)} + k'(X^i)\sqrt{\Delta t}R^i + o(\Delta t) \approx \sqrt{2k(X^i)} + k'(X^i)\sqrt{\Delta t}R^i \end{aligned}$$

In other words the following equation is valid

$$\sqrt{2k(X^i + \Delta\tilde{X}^i)} \approx \sqrt{2k(X^i)} + k'(X^i)\sqrt{\Delta t}R^i \tag{7.14}$$

Substituting (7.14) into (7.13) yields

$$E(\Delta X^i) = u^i\Delta t + \frac{1}{2}k'(X^i)\Delta t + \frac{1}{2}k'(X^i)\Delta t E(R^i)^2 = u^i\Delta t + k'(X^i)\Delta t \tag{7.15}$$

The variation of the displacement ΔX in the Heun scheme coincides with the variation of the concentration

$$\begin{aligned} &Var(\Delta X^i) = E\left(\tfrac{1}{2}((\sqrt{2k(X^i)\Delta t} + \sqrt{2k(X^i + \Delta\tilde{X})\Delta t})R^i - k'(X^i)\Delta t)\right)^2 = \\ &E\left(\sqrt{2k(X^i)\Delta t}R^i + \tfrac{1}{2}k'(X^i)(R^i - 1)\Delta t\right)^2 = 2k(X^i)\Delta t + o(\Delta t) \approx 2k(X^i)\Delta t \end{aligned} \tag{7.16}$$

We can conclude that the random walk model (7.12) has the same first two moments as a standard random walk model in the Îto sense and as in (7.9).

Finally, the backward Euler scheme is appropriate for the backward Îto formulation (see [78])

$$\begin{aligned} &\Delta\tilde{X}^i = \sqrt{2k(X^i)\Delta t}R^i \\ &X^i = X^{i-1} + u^i\Delta t + \sqrt{2k(X^i + \Delta\tilde{X}^i)\Delta t}R^i \end{aligned} \tag{7.17}$$

Using Eq. (7.14) we can again find the moments of the distribution of $X(t)$ obtained by backward Euler scheme

$$\begin{aligned} &E(\Delta X^i) = E\left(u^i\Delta t + \sqrt{2k(X^i + \Delta\tilde{X}^i)\Delta t}R^i\right) = \\ &u^i\Delta t + E\left((\sqrt{2k(X^i)} + k'(X^i)\sqrt{\Delta t}R^i)\sqrt{\Delta t}R^i\right) = u^i\Delta t + k'(X^i)\Delta t \end{aligned} \tag{7.18}$$

and

$$\begin{aligned} &Var(\Delta X^i) = E\left(\sqrt{2k(X^i + \Delta\tilde{X}^i)\Delta t}R^i - k'(X^i)\Delta t\right)^2 = \\ &E\left(\sqrt{2k(X^i)\Delta t}R^i + k'(X^i)\Delta t(R^i)^2 - k'(X^i)\Delta t\right)^2 = 2k(X^i)\Delta t + o(\Delta t) \approx 2k(X^i)\Delta t \end{aligned} \tag{7.19}$$

As a result the solution obtained by the backward Îto random walk model is consistent with the advection-diffusion equation (7.1). The main differences between the Îto, Stratonovich and the backward Îto formulation are shown in Fig. 7.1.

7.4 Illustrations

In this section the random walk models (in Îto and backward Îto senses) are applied for two test cases. In general, the analytical solution of the direct problem (7.1) cannot be found; however, the residence time of a tracer can be obtained [29, 26, 27]. The residence time of a water or tracer parcel in a control domain is usually defined as the time taken

by this parcel to leave the domain of interest [11, 133, 134, 15, 55, 122, 30]. As such, the residence is one of the most popular tool to describe and understand environmental issues.

Mathematically, the mean residence time $\theta(x)$ of the tracer of initial mass $m(t_0)$ released at time t_0 can be computed by monitoring the temporal evolution of the mass of the tracer in the control region [11, 122]

$$\theta(x) = -\frac{1}{m(t_0)} \int_{m(t_0)}^{0} t dm \tag{7.20}$$

Delhez, Heemink and Deleersnijder [29] introduced an alternative procedure designed for numerical models. They showed that the residence time can be found from the solution of the adjoint problem to the advection-diffusion equation.

For both examples, we assume that the diffusivity is discontinuous at some location. Such diffusivity profile does not exist in the nature; however, there are regions of large space variations of the diffusivity. The discontinuous diffusivity can be considered as a limit case for which it is generally easier to find the analytical solution. In addition, if the Lagrangian method under consideration can successfully handle a discontinuity in the diffusivity field, it is safe to assume that this method will be able to deal with region of high gradients of the eddy coefficient.

7.4.1 Illustration 1: Settling and diffusion problem

Let us consider the settling-diffusion model introduced in Section 4.5. We will use the same notations as Section 4.5, in particular, let z be a vertical coordinate that increases upwards. In [27] it is shown that the residence time may exhibit a discontinuity at the interval between the mixed layer ($0 < z < h$) and pycnocline ($z < 0$), for the eddy diffusivity is zero in the latter and positive in the former. Now we assume that the boundary of interest is $z = \delta$, rather than $z = 0$. δ is positive or negative according to whether the boundary is located in the mixed layer or the pycnocline, respectively. The corresponding residence time is hereinafter denoted

$$\tau(z_0, \delta) = \int_0^{\infty} \int_{\delta}^{h} C(t, z) dz dt$$

which may be recast as a function of z. There is no closed form solution for $C(t, z)$, but the residence time may be calculated analytically ([26]).

For the sake of simplicity, it is assumed that the eddy diffusivity is a positive constant λ, in the mixed layer and zero in the pycnocline. It is also desirable to introduce dimensionless variables:

$$z' = \frac{z}{h}, \quad \theta' = \frac{\theta}{h/w}, \quad Pe = \frac{wh}{\lambda} \tag{7.21}$$

From now on, only dimensionless variables will be used, so that it is appropriate to drop the primes. If $\delta > 0$, the lower boundary is located at a level where the eddy diffusivity is nonzero, while $\delta < 0$ corresponds to the case, when the lower boundary is located

below the pycnocline. Let us assume that the lower boundary of the domain is pushed towards the bottom of the mixed layer $\delta > 0$, $\delta \to 0^+$, $\delta < 0$, $\delta \to 0^-$. [27] show that the corresponding residence times are different. In particular, for the chosen value of the diffusivity

$$\tau(z, 0^-) = z + \frac{1 - e^{-Pe(1-z)}}{Pe} \tag{7.22}$$

and

$$\tau(z, 0^+) = z - \frac{e^{-Pe(1-z)} - e^{-Pe}}{Pe} \tag{7.23}$$

We apply the Îto and the backward Îto random walk models to simulate the transport of the tracer in the proposed model. The Îto random walk model formulation corresponds to the case when the lower boundary of the domain is placed above the pycnocline, while the backward Îto random walk model provides the solution of the case when the lower boundary of the domain is placed under pycnocline.

The exact and the numerical solutions for $N = 10^4$ particles are shown in Fig. 7.2. From Figure 7.2 we can conclude that the residence times obtained by applying the Îto and backward Îto random walk schemes are different. One can wonder which scheme provides the right solution. In reality, both methods are correct, however they give answers for two different problems.

In Section 7.2 it was shown that for a smooth diffusivity function both random walk schemes are identical. In the Îto case, an additional drift due to the spatial variation of the diffusivity is present. Because of this additional drift particles cannot stay in regions with low diffusivity. In the backward Îto formulation the additional drift term has disappeared and is included in the random term by applying the two-steps backward Euler scheme.

The disadvantage of the Îto formulation is that it cannot handle the case of discontinuous diffusivity. By applying an Îto model in this case the diffusivity drift is zero everywhere except exactly at the boundary where it is infinite. By applying a numerical scheme, particles will never reach exactly the pycnocline and as a result the diffusivity drift becomes essentially zero. Therefore a particle that comes close to the boundary will never go back into the domain (see Figure 7.3(a)) and the residence time computed is in fact the residence time of a domain without the pycnocline.

By applying the backward Îto model the diffusivity drift is included in the random term of the model. Now a particle does get back into the domain even if it is very close to the boundary ((see Figure 7.3(b)) and it is now possible to compute the residence time of the domain including the pycnocline.

7.4.2 Illustration 2: The direct and adjoint problems for the residence time

In the previous section we considered a model, in which the diffusivity exhibits a discontinuity at the boundary of the domain. However, in practice the diffusivity can change rapidly inside the domain of interest. An example of such a problem is needed. In this respect, inspiration may be found in [30].

Let t and x denote time and a space coordinate, respectively. In the domain $-\mathcal{L} \leq x \leq$

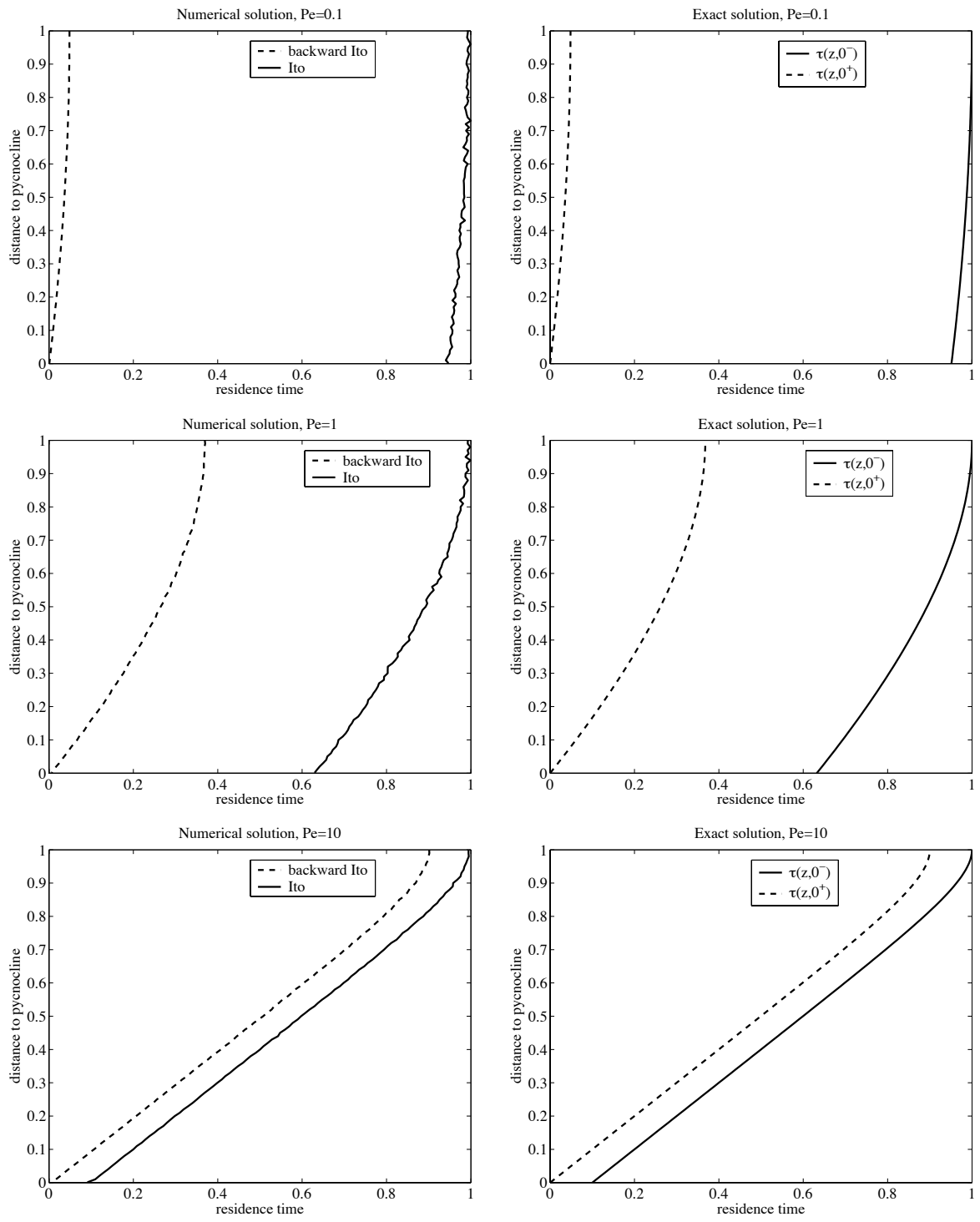
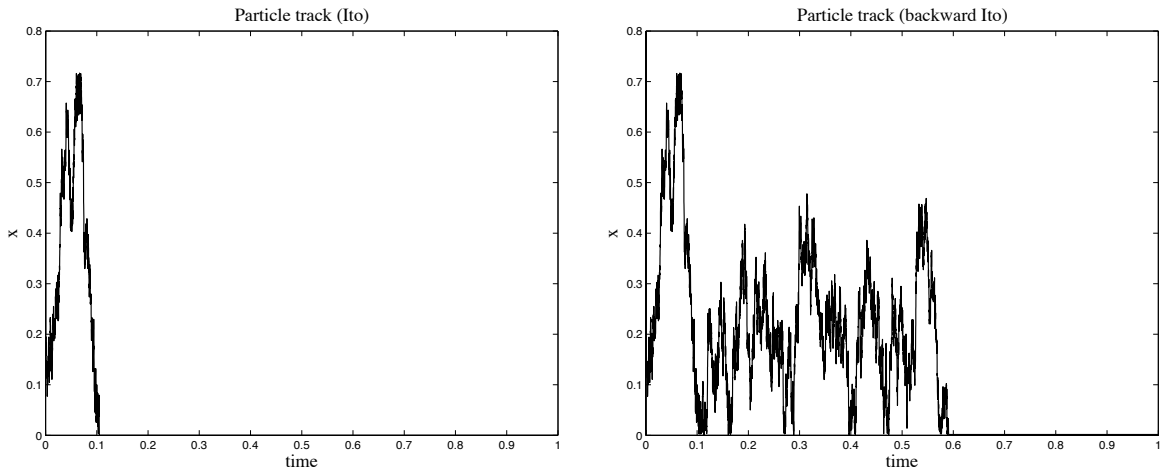


Figure 7.2: The profile of the residence times $\tau(z, 0^-)$ and $\tau(z, 0^+)$ in the surface mixed layer for various values of the Peclet number. Dimensionless variables are used and the eddy diffusivity is assumed to be constant in the mixed layer



(a) (b)

Figure 7.3: The track of particle released at location $z_0 = 0.9$ in case of (a) Îto and (b) backward Îto random walk models

$\mathcal{L}$, the concentration of the tracer $C(t, x)$ obeys the following partial differential problem

$$\begin{aligned} &\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}\left(uC - k\frac{\partial C}{\partial x}\right) \\ &C(0, x, x_0) = \delta(x - x_0), \quad -\mathcal{L} < x_0 < \mathcal{L} \\ &C(t, \pm\mathcal{L}, x_0) = 0 \end{aligned} \tag{7.24}$$

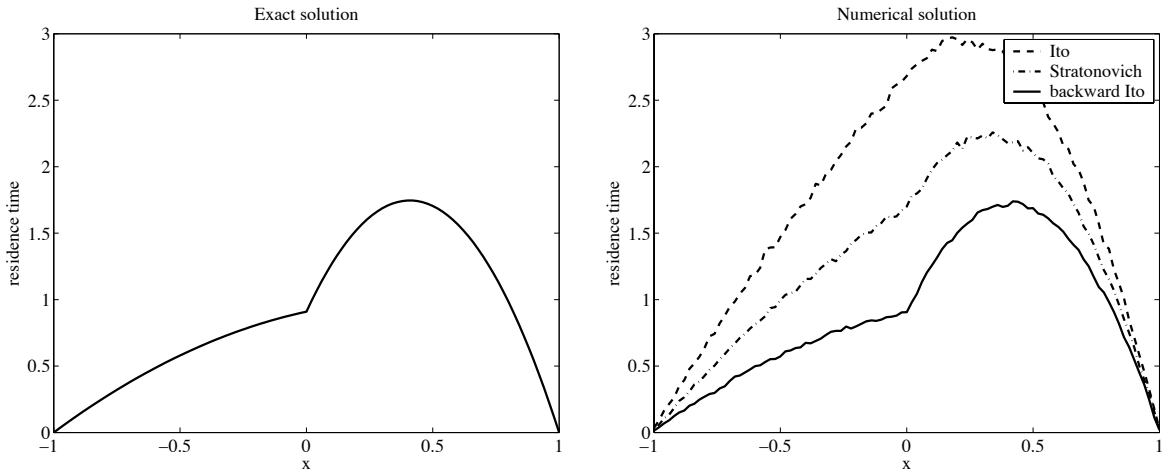
where the positive constant u is the fluid velocity, while $k(x) > 0$ denotes the eddy diffusivity. The residence time in the domain of interest of the tracer whose concentration obeys the partial differential problem (7.24) is ([29])

$$\theta(x_0) = \int_0^\infty \int_{-\mathcal{L}}^{\mathcal{L}} C(t, x_0) dx dt \tag{7.25}$$

In principle this value may be evaluated for any admissible value of x_0. The ensuing function may then be recast as a function of x, i.e. $\theta(x)$. However, obtaining the analytical solution of the direct problem (7.24) is usually considered as difficult. Fortunately, it is much easier to obtain the residence time by solving the adjoint problem ([29, 30]):

$$\begin{aligned} &\frac{d}{dx}\left(k\frac{d\theta}{dx} + u\theta\right) = -1 \\ &\theta(\pm\mathcal{L}) = 0 \end{aligned} \tag{7.26}$$

For the purposes of the present study, the eddy diffusivity must exhibit a discontinuity inside the domain of the interest. The simplest expression that satisfies this constraint



(a) (b)

Figure 7.4: The profile of the residence time in case of zero advection. Dimensionless variables were used. The value of the parameter: $\mu = 0.1$. (a) Analytical solution (b) Numerical solution for $N = 10^4$ particles with time step $\Delta t = 10^{-4}$

probably is the following piecewise constant function

$$k(x) = \begin{cases} k^+, & 0 < x \leq \mathcal{L} \\ k^-, & -\mathcal{L} \leq x < 0 \end{cases} \tag{7.27}$$

where k^+ and k^- are positive constants. Therefore, at $x = 0$, the residence time must satisfy two matching conditions:

$$[\theta(x)]_{x=0^-}^{x=0^+} = 0 \tag{7.28}$$

$$\left[k\frac{\partial\theta}{\partial x} + u\theta\right]_{x=0^-}^{x=0^+} = 0 \tag{7.29}$$

In the developments below, the residence time at $x = 0$ will be denoted θ_0. In other words, the latter satisfies the equalities

$$\theta(0^-) = \theta_0 = \theta(0^+) \tag{7.30}$$

The zero advection case

If the advection is zero ($u = 0$), then it is appropriate to introduce the dimensionless parameter $\mu = k^+/k^-$ and variables

$$\begin{aligned} t' &= \frac{t}{\mathcal{L}/(k^-)^2}, \quad (x', x_0') = \frac{(x, x_0)}{\mathcal{L}}, \quad k' = \frac{k}{k^-}, \\ C' &= \frac{C}{1/\mathcal{L}}, \quad (\theta', \theta_0') = \frac{(x, x_0)}{\mathcal{L}/(k^-)^2} \end{aligned} \tag{7.31}$$

For the sake of simplicity the primes can be dropped. Hence, the dimensionless diffusivity is

$$k(x) = \begin{cases} \mu, & 0 < x \leq 1 \\ 1, & -1 \leq x < 0 \end{cases} \tag{7.32}$$

After some calculations, the residence time is obtained:

$$\theta(x) = \begin{cases} -\dfrac{x^2}{2\mu} - \dfrac{2\mu\theta_0 - 1}{2\mu}x + \theta_0, & 0 < x \leq 1 \\ -\dfrac{x^2}{2} + \dfrac{2\theta_0 - 1}{2}x + \theta_0, & -1 \leq x < 0 \end{cases} \tag{7.33}$$

with

$$\theta_0 = \frac{1}{1+\mu} \tag{7.34}$$

The analytical and numerical solutions obtained by Îto, Stratonovich and backward Îto random walk methods are shown on Fig. 7.4. Clearly, the backward Îto solution is much better than Stratonovich, which, in turn, is better than that obtained by the classical Îto method.

The advection-diffusion case

If advection is present ($u > 0$), then it is appropriate to introduce the following dimensionless parameters and variables:

$$\begin{aligned} t' &= \frac{t}{\mathcal{L}/u}, \quad (x', x_0') = \frac{(x, x_0)}{\mathcal{L}}, \quad Pe^{\pm\prime} = \frac{u\mathcal{L}}{k^{\pm}}, \\ C' &= \frac{C}{1/\mathcal{L}}, \quad (\theta', \theta_0') = \frac{(x, x_0)}{\mathcal{L}/u} \end{aligned} \tag{7.35}$$

As in the previous example the primes can be dropped. It is also useful to define a piecewise constant Peclet number:

$$Pe(x) = \begin{cases} Pe^+, & 0 < x_0 \leq 1 \\ Pe^-, & -1 \leq x_0 < 0 \end{cases} \tag{7.36}$$

After some calculations, the residence time is obtained:

$$\theta(x) = \begin{cases} a^+ - x + b^+ e^{-Pe^+ x}, & 0 < x \leq 1 \\ a^- - x + b^- e^{-Pe^- x}, & -1 \leq x < 0 \end{cases} \tag{7.37}$$

with

$$a^{\pm} = \frac{e^{\mp Pe^{\pm}}\theta_0 \mp 1}{e^{\mp Pe^{\pm}} - 1}, \quad a^{\pm} = \frac{\pm 1 - \theta_0}{e^{\mp Pe^{\pm}} - 1} \tag{7.38}$$

and

$$\theta_0 = \frac{Pe^+ - Pe^-}{Pe^+ Pe^-}\frac{(e^{-Pe^+} - 1)(e^{-Pe^-} - 1)}{e^{-Pe^+} - e^{-Pe^-}} - \frac{e^{-Pe^+} + e^{-Pe^- - 1}}{e^{-Pe^+} - e^{-Pe^-}} \tag{7.39}$$

Figure 7.5 shows the analytical solution and the numerical solutions corresponded to Îto, Stratonovich and backward Îto formulations. One can see easily that only the solution obtained by backward Îto random walk model is very close to the analytical solution, while the Stratonovich and Îto solutions significantly differ from the exact residence time.

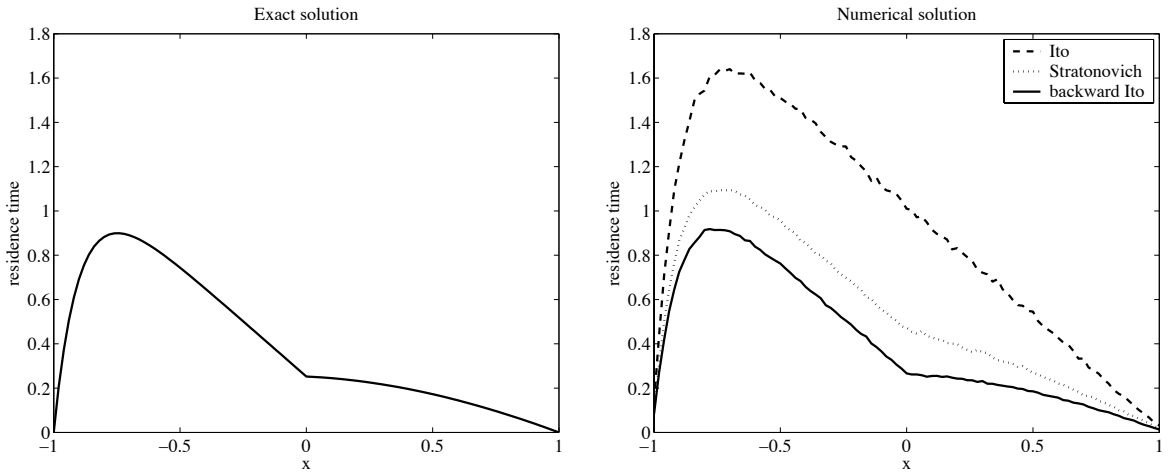


Figure 7.5: The profile of the residence time in case of non zero advection. The value of parameter: $Pe^+ = 0.5$, $Pe^- = 10$. Dimensionless variables were used. (a) Analytical solution (b) Numerical solution for $N = 10^4$ particles

7.5 Conclusion

In this chapter we considered the random walk model that can be applied to model the transport process in the regions with large space variations of the diffusivity, or as the limit case with discontinuous diffusivity. This model proposed by [78] is based on the backward Îto stochastic integral. It is consistent with the advection-diffusion equation and does not contain the diffusivity gradient in the advection part. Two test cases were analyzed: the sinking-diffusion model, in which the diffusivity exhibits a discontinuity at one boundary of the domain and the advection-diffusion problem with a discontinuity in the diffusivity inside the domain of interest. For both test cases the analytical solution of the indirect problem, e.g. finding the residence time, is known. The backward Îto random walk model was applied and the results show that this model provides the correct results for discontinuous diffusivities. The backward Îto random walk model was applied and the results show that this model provides the correct results for discontinuous diffusivities, while other, better known, random walk models perform rather poorly.

Chapter 8

Conclusions and recommendations

8.1 Main conclusions

The main aim of this thesis was to introduce and apply new random walk models and demonstrate that the Lagrangian approach may play an important role in environmental modelling. The main conclusions of this research are:

- **The Lagrangian approach is suitable for solving the advection-diffusion model for space-varying non-diagonal diffusivity.** Some diffusion processes in the ocean can be formulated using the non-diagonal space-varying diffusion term. An example is the diffusion process along the isopycnal surfaces, i.e. surfaces of equal density. The traditional Euler methods are susceptible to excessive numerical dispersion and artificial oscillations, while the Lagrangian approach provides an accurate and effective solution. Results of applying the random walk model for several idealized test cases show that the accuracy of the method does not depend on the diffusivity term.

- **The reverse-time diffusion is an efficient method to solve inverse problems.** In some practical situations we are more interested in solving the inverse problem, i.e. finding the possible source of contamination, rather than estimating the concentration. Using forward models for these problems is computationally very expensive, because we need to repeat the simulation for different initial conditions. An alternative is to use the reverse-time stochastic model that can be derived from the original one. The particle track is modelled in the reverse-time direction and, in this way, there is no need for many simulations. The reverse-time diffusion was applied to construct the risk map for several critical locations along the Dutch coastal zone.

- **The combination of the forward and reverse-time diffusions allows reduction of the computational time by orders of magnitude.** If it is required to describe the spreading of the pollutant after a calamity at sea, forward modelling is the most suitable method. However, sometimes, we are interested in the concentration of the pollutant in a certain location or area. When the concentration in the location of interest is relatively small, most particles hardly contribute to

the estimation. The solution is to use a combination of forward and reverse-time diffusion. We start the forward simulation from the location of the calamity and, at the same time, the reverse-time simulation from the location of interest. This technique reduces the CPU time significantly.

- **The fluctuations in the concentration of the pollutant due to spatial correlation of the turbulence may become high and should be taken into consideration**. The turbulence is correlated in space. As a result, the actual concentration at certain locations may be much higher or lower than the ensemble mean concentration. For instance, the ensemble mean concentration may be an average of a large number of zeros (realizations when the cloud of pollutant does not reach the location) and a few very large values. This type of averaging may be meaningless, because the few high concentrations may kill the organisms in a certain area and the large number of zeros cannot bring them to life again. The deviation of the concentration can be found from the two-particle model. Here, two particles are released at the same time at the same location and their movement correlated. The resulting model suffers from the so-called 'curse of dimensionality'; however, the accurate solution may be obtained in a reasonable time by using the forward-reverse estimator.

- **The random walk model based on the backward Îto calculus may be used for modelling advection-diffusion process in case of discontinuous diffusivity.** In some situations the diffusivity term may be discontinuous at the boundary. In such cases, the traditional random walk method is no longer applicable; however, we can use the stochastic model, which does not require the diffusivity gradient in the advection part. This model is based on the backward Îto calculus and is also consistent with the advection-diffusion equation.

8.2 Recommendations

The methods and results presented within this dissertation provide an essential framework for more effective use of Lagrangian methods (especially, methods based on reverse-time diffusion in operational transport simulators). In addition, the tools developed here can be used as a basis for future methods development. Future research may address the following topics:

- **Developing and applying Eulerian–Lagrangian methods.** The results obtained show that the Lagrangian methods may handle some problems more effectively than the Eulerian methods. However, the Lagrangian methods also have some limitations, for example they are computationally expensive. In the optimal transport simulator the advantages of both methods should be used. It is expected that the Eulerian and Lagrangian methods, incorporated within a single software code, can both be optimized for improved computational performance and accuracy.

- **It is desirable to compare the random walk method with the Eulerian methods for real-life applications in oceanography.** The Eulerian methods

are widely used for transport modelling in the ocean. In this research the stochastic methods were applied for some oceanographic problems, for example iso- and diapycnal diffusion and residence time of the tracer. These methods were applied for several idealized test cases and the numerical results correspond to the analytical solutions. However, the significance of the results would be increased if the developed random walk methods were applied to real applications and the computational performance of the new methods were compared with the performance of the traditional approaches.

- **The three-dimensional implementation of the algorithms in this dissertation is probably the most important 'next step' in the integration of the methods based on the reverse-time diffusion and contaminant transport simulators.** Three-dimensional transport modelling describes the pollutant spreading more accurately than the depth-averaged models. However, the usage of these methods for practical aims is limited by computational expensiveness. It is expected that the improvement in computational time given by the forward-reverse estimator will be better in three-dimensional models than in two-dimensional models.

- **Develop new random-flight methods based on the reverse-time diffusion.** As the random walk methods are valid only for subset of the situations, it is desirable to apply the reverse-time diffusion for the random-flight models. The random-flight models are usually used for transport modelling in cases of high Reynolds numbers [52]. Here the flow is modelled as the stochastic process, while in the particle models considered in this thesis the displacement of particles is simulated using random numbers. It is expected that the new methods based on reverse-time diffusion will lead to the significant improvement of the computational performance.

Bibliography

[1] Anderson B. D. O. (1982). Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12, pp.313-326.

[2] Arnold L. (1974). Stochastic Differential Equations. John Wiley & Sons, New York.

[3] Artemiev S. and Averina T. (1997). A Numerical Analysis of System of Ordinary and Stochastic Differential Equations. VSP, Utrecht, the Netherlands.

[4] Baas A. F., De (1988). Some Properties of the Langevin Model for Dispersion. Ph.D. thesis, Delft University of Technology.

[5] Baptista A.E. de M. (1987). Solution of Advection-Dominated Transport by Eulerian-Lagrangian Methods Using the Backward Method of Characteristics. Ph.D. dissertation, MIT, Cambridge.

[6] Batchelor G.K. (1953). The Theory of Homogeneous Turbulence. Cambridge University Press.

[7] Bear J. and Verruijt A. (1987). Modeling Groundwater flow and Pollution, D.Reidel, Dodrecht, the Netherlands.

[8] Beckers J. M., Burchard H., Campin J. M., Deleersnijder E. and Mathieu P.P. (1998). Another reason why simple discretizations of rotated diffusion operators cause problems in ocean models: Comments on "Isoneutral diffusion in a z-coordinate ocean model". *Journal of Physical Oceanography*, 28, pp.1552-1559.

[9] Beckers J. M,, Burchard H., Deleersnijder E. and Mathieu P.P. (2000). Numerical discretization of rotated diffusion operators in ocean models. *Monthly Weather Review*, 128, pp.2711-2733.

[10] van den Berg E., Heemink A.W., Lin H.X. and Schoenmakers J.G.M. (2005). Probability density estimation in stochastic environmental models using reverse representations. *Stochastic Environmental Research and Risk Assessment*, 20, 1-2, pp.126-134.

[11] Bolin B. and Rodhe H. (1973). A note on concepts of age distribution and transit time in natural reservoirs. *Tellus*, 25, pp.58-62.

[12] Borgas M.S. and Sawford B.L. (1994). A family of stochastic models for two-particle dispersion in isotropic homogeneous stationary turbulence. *Journal of Fluid Mechanics*, 279, pp.69-99.

[13] Bjork T. (1998). Arbitrage Theory in Continuous Time. Oxford University Press.

[14] Boughton B. A., Delaurentis J. M., and Dunn W.E. (1987). A stochastic model of particle dispersion in the atmosphere. *Boundary-Layer Meteorology*, 40, pp.147-163.

[15] Braunschweig F., Martins F., Chambel P. and Neves R. (2003). A methodology to estimate renewal time scales in estuaries: the Tagus Estuary case. *Ocean Dynamics*, 53, pp.137-145.

[16] Brzezniak Z. and Zastawniak W. (1998). Basic Stochastic Processes. Springer-Verlag, Berlin.

[17] Charles W.M. (2007). Transport Modelling in Coastal Waters Using Stochastic Differential Equations. Ph.D. thesis, Delft University of Technology.

[18] Cheng R. T., Casulli V. and Milford S. N. (1984). Eulerian-Lagrangian solution of the convection-dispersion equation in natural coordinates. *Water Resources Research*, 28(7), pp.944-952.

[19] Costa M. and Ferreira J. S. (2000). Discrete particle distribution model for advection-diffusion transport. *Journal of Hydraulic Engineering*, 126(7), pp.525-532.

[20] Cox M. D. (1987). Isopycnal diffusion in a z-coordinate model. *Ocean Modelling*, 74, pp.1-5.

[21] Cox R.A. and Nishikawa T. (1991). A new total Variation Dimishing Scheme for the Solution of Advective-Dominant Solute Transport. *Water Resources Research*, 27(10), pp.2645-2654.

[22] Cressie N. A. C. (1991). Statistics for Spatial Data. John Wiley & Sons, New York.

[23] Crone G.C. (1997). Parallel Lagrangian Models for Turbulent Transport and Chemistry. Ph.D. thesis, University of Utrecht.

[24] Dahl F.A. (2002). Variance reduction for simulated diffusions using control variates extracted from state space evaluations. *Applied Numerical Mathematics*, 43, pp.375-381.

[25] Deleersnijder E.,Campin J.-M. and Delhez E. J. M. (2001). The concept of age in marine modelling I. Theory and preliminary model results. *Journal of Marine Systems*, 28, pp.229-267.

[26] Deleersnijder E., Beckers J. M. and Delhez E. J. M. (2006a) The residence time of settling in the surface mixed layer. *Environmental Fluid Mechanics*, 6(1), pp.25-42.

[27] Deleersnijder E., Beckers J. M. and Delhez E. J. M. (2006b). On the behaviour of the residence time at the bottom of the mixed layer. *Environmental Fluid Mechanics*, 6, pp.541-547.

[28] Delhez E. J. M. (2006). Transient residence and exposure times. *Ocean Science*, 2, pp.1-9.

[29] Delhez E. J. M., Heemink A. W. and Deleersnijder E. (2004). Residence time in a semi-enclosed domain from the solution of an adjoint problem. *Estuarine, Coastal and Shelf Science*, 61, pp.691-702.

[30] Delhez E. J. M. and Deleersnijder E. (2006). The boundary layer of the residence time field. *Ocean Dynamics*, 56, pp.139-150.

[31] Dimou K. N. and Adams E. E. (1993). A random-walk, particles tracking models for well-mixed estuaries and coastal waters. *Estuarine, Coastal and Shelf Science*, 37, pp.99-110.

[32] De Jong K. (1998). Tidally averaged transport models. Ph.D. dissertation, Delft University of Technology.

[33] De Kok J. M. (1994). Tidal averaging and models for anisotropic dispersion in coastal waters. *Tellus*, 46A, pp.160-177.

[34] Douglas J., Jr. and Russell T.F. (1982). Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element of finite difference procedures. *SIAM Journal of Numerical Analysis*, 19, pp.871-885.

[35] Durbin P. A. (1980). A stochastic model of two-particle dispersion and concentration fluctuations in homogeneous turbulence. *Journal of Fluid Mechanics*, 100, pp.279-302.

[36] Durrett R. (2000). Stochastic Calculus: A Practical Introduction. CRC Press, Boca Raton.

[37] Dyke P. P. G. (2001). Coastal and Shelf Sea Modelling, Kluwer Academic Publisher, Dodrecht.

[38] Elliott R. J. and Anderson B. D. O. (1985). Reverse time diffusions. *Stochastic Processes and their Applications*, 19, pp.327-339.

[39] Ermak D.L. and Nasstrom J.S. (2000). A Lagrangian stochastic diffusion method for inhomogeneous turbulence. *Atmospheric Environment*, 34, pp.1059-1068.

[40] Ewing R.E. and Wang H. (2001). A summary of numerical methods for time-dependent advection-dominated partial differential equations. *Journal of Computational and Applied Mathematics*, 128(1-2), pp.423-445.

[41] Fay B., Glaab H., Jacobsen I. and Schrodin R. (1995). Evaluation of Eulerian and Lagrangian atmospheric transport models at the Deutcher Wetterdienst using ANATEX surface tracer data. *Atmospheric Environment*, 29, pp.2485-2497.

[42] Fick A. (1855). On liquid diffusion. *Philos.Mag.J.Sci.*, 10, 31-39.

[43] Fischer H.B., List E.J., Koh R.C.Y., Imberger J. and Brooks N.H. (1979). Mixing in Inland and Coastal Waters. Academic Press, New York.

[44] Fukunaga K. (1972). Introduction to Statistical Pattern Recognition. Academic Press, New York.

[45] Garder A.O.J., Peaceman D.W. and Possi L.J. (1964). Numerical calculation of multidimensional miscible displacement by the method of characteristics. *Society of Petroleum Engineers Journal*, 6(2), pp.175-182.

[46] Gihman I.I. and Skorohod A.V. (1972). Stochastic Differential Equations. Springer-Verlag, Berlin.

[47] Glasserman P. (2003). Monte Carlo Methods in Financial Engineering. Springer-Verlag, Berlin.

[48] de Graaf M., Jager Z., Vreugdehil C.B. and Elorche M. (2004). Numerical simulation of tidally cued vertical migrations of flatfish lavrae in the North Sea. *Estuarine, Coastal and Shelf Science*, 59, pp.295-305.

[49] Greengard L. and Strain J. (1991). The fast Gauss transform. *SIAM Journal of Scientific and Statistical Computing*, 12(1), pp.79-94.

[50] de Haan P. (1999). On the use of density kernels for concentration estimations within particles and puff dispersion models. *Atmospheric Environment*, 33, pp.2007-2021.

[51] Heath D. and Platen E. (2002) A variance reduction technique based on integral representations. *Quantative Finance*, 2(5), pp.362-369.

[52] Heemink A.W. (1990) Stochastic modeling of dispersion in shallow water. *Stochastic Hydrology and Hydraulics*, 4, pp.161-174.

[53] Heemink A.W. (1993) Tidally averaged models for dispersion in shallow water. *Water Resources Research*, 29(3), pp.607-617.

[54] Holley E. R. (1996). Diffusion and Dispersion. In: Singh V. P., Hager W. H. (eds), Environmental hydraulics, Kluwer Academic Press, Dodrecht, pp. 111-151.

[55] Holzer M. and Hall T. M. (2000). Transient-time and tracer-age distributions in geophysical flows. *Journal of the Atmospheric Sciences*, 57, pp.3539-3558.

[56] Huber E., Spivakovskaya D., Lin H.X. and Heemink A. W. (2006). The parallel implementation of forward-reverse estimator. *Proceedings of European Conference on Computational Fluid Dynamics ECCOMAS CFD 2006*, P. Wesseling, E. Oñate and J. Périaux (Eds).

[57] Hull J.C. (2000). Options, Futures, and other Derivatives. Prentice Hall.

[58] Hunter, J. R., Craig P. D. and Phillips H. E. (1993). On the use of random walk models with spatially variable diffusivity. *Journal of Computational Physics*, 106, pp.366–376.

[59] Huykorn P.C. and Pinder G.F. (1983). Computational Methods in Subsurface Flow, Academic Press, New York.

[60] Istok J. (1989). Groundwater Modeling by the Finite Element, American Geophysical Monograph 13, Washington DC.

[61] Îto K. (1951). On Stochastic Differential Equations. *Memoirs. American Mathematical society*,4, pp.1-51.

[62] Jazwinski A.W. (1970). Stochastic Differential Processes and Filtering Theory. Academic Press, New York.

[63] Jones M. C. (1993). Simple boundary correction for kernel density estimation. *Statistics and Computing*, 3, pp.135-146.

[64] Kaplan H. and Dinar N. (1987). A three dimensional stochastic model for concentration fluctuation statistics in isotropic homogeneous turbulence. *Journal of computational Physics*, 79, pp.317-335.

[65] Karatzas I. and Shreve S.E. (1998). Brownian Motion and Stochastic Calculus. Springer-Verlag, Berlin.

[66] Kinzelbach W. (1988). Groundwater Modelling, An Introduction with sample Programs in BASIC. Elsevier, Amsterdam.

[67] Kinzelbach W. (1988). The random walk method in pollutant transport simulation, In: Custodio et al. (eds), Groundwater Flow and Quality Modeling, D.Reidel Publishing Company, Dodrecht, pp.227-245.

[68] Kloeden P.E. and Platen E. (1999). Numerical Solution of Stochastic Differential Equations. Springer-Verlag, Berlin.

[69] Kloeden P.E., Platen E. and Schurz H. (2003). Numerical Solution of SDE Through Computer Experiments. Springer-Verlag, Berlin.

[70] Kolmogorov A. N. (1941). The local structure of turbulence in an incompressible fluid at very high Reynolds numbers. *Doklady AN SSSR*, 30(4), pp.299-304.

[71] Kolmogorov A.N. (1941). Energy dissipation in locally isotropic turbulence. *Doklady AN SSSR*, 32(1), pp.19-21.

[72] Kolodko A. and Sabelfeld K. (2003). Stochastic particle methods for Smoluchowski coagulation equation: variance reduction and error estimations. *Monte Carlo Methods Applications*, 9(4), pp.315-339.

[73] Konikow L. F. and Bredehoeft J. D. (1978). Computer model of two-dimensional solute transport and dispersion in ground water. U.S. Geological survey water-resources investigations book 7, Chapter C2.

[74] Kurbanmuradov O., Rannik U., Sabelfeld K. and Vesala T. (1999). Direct and adjoint Monte Carlo algorithms for the footprint problem. *Monte Carlo methods and Applications*, 5(2), pp.85-113

[75] Kurbanmuradov O., Rannik U., Sabelfeld K. and Vesala T. (2001). Evolution of mean concentration and fluxes in turbulent flows by Lagrangian stochastic models. *Mathematics and Computers in simulation*, 54, pp.459-476

[76] Leonard B. P., Lock A. P. andMacVean M. K. (1996). Conservative explicit unrestricted-time-step multidimensional constancy-preserving advection schemes. *Monthly weather Review*, 124, pp. 2588-2606.

[77] LaBolle E. M., Fogg G.E. and Tompson A. F. B. (1996). Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods. *Water Resources Research*, 32(3), pp.583-593.

[78] LaBolle E. M., Quastel J., Fogg G. E. and Granver J. (2000). Diffusion processes in composite porous media and their numerical integration by random walks: Generalized stochastic differential equations with discontinuous coefficients. *Water Resources Research*, 36(3), pp.651-662.

[79] Li S. G. and Wei C. F. (1998). Improved finite analytic methods for heterogeneous transport in multi-dimensions. *Journal of Hydraulic Engineering*, 124(4), pp.358-369.

[80] Lowry T. and Li. S. (2002). A characteristic-based finite analytic method for solving the two dimensional steady state advection-diffusion equation. *Water Resources Research*, 38(7), pp.1-15.

[81] Lu Z. and Zhang D. (2003). On importance sampling Monte Carlo approach to uncertainty analysis for flow and transport in porous media. *Advances in Water Resources*, 26, pp.1177–1188.

[82] Maruyama G. (1955). Continuous Markov Processes and Stochastic Equations. *Rend. Circolo Math. Palermo*, 4, pp.48-90.

[83] Mathieu P. P. and Deleersnijder E. (1998). What is wrong with isopycnal diffusion in world ocean models? *Applied Mathematical Modelling*, 22, pp.367-378.

[84] Milstein G.N. (1995). Numerical Integration of Stochastic Differential Equations. Kluwer Academic Publisher, Dodrecht.

[85] Milstein G. N., Schoenmakers J. G. M. and Spokoiny V. (2004). Transition density estimation for stochastic differential equations via forward-reverse representations. *Bernoulli*, 10(2), pp.281-312.

[86] Milstein G.N. and Tretyakov M.V. (2004). Stochastic Numerics for Mathematical Physics. Springer-Verlag, Berlin.

[87] Monin A. S. and Yaglom A. M. (1979). Statistical Fluid Mechanics. MIT Press, Cambridge.

[88] Müller H. G. (1991). Smooth optimum kernel estimators near endpoints. *Biometrika*, 78(3), pp.521-530.

[89] Naifar F. (2006). A Finite Volume Solver for the Simulation of Transport Processes. Ph.D. thesis, Delft University of Technology.

[90] Narasimhan T. N. and Witherspoon P. A. (1976). An Integrated finite difference method for analyzing fluid flow in porous media. *Water Resources Research*, 12, pp.57-64.

[91] Oksendal B. (2000). Stochastic Differential Equations. Springer-Verlag, Berlin.

[92] Pinder G. F. (1988). An overview of groundwater modelling. In: Custodio et al. (eds), Groundwater Flow and Quality Modeling, pp.119-134.

[93] Pinder G.F. and Gray W.G. (1977). Finite Elements in Surface and Subsurface Hydrology, Academic Press, New York.

[94] Redi M. H. (1982). Oceanic isopycnal mixing by coordinate rotation. *Journal of Physical Oceanography*, 12, pp.1154-1158.

[95] Remson I., Hornberger G.M. and Molz F.J. (1971). Numerical Methods in subsurface Hydrology. John Wiley & Sons, New York.

[96] Richardson L.F. (1922). Weather Prediction by Numerical Process. Cambridge University Press.

[97] Riddle A. M. (1998). The specification of mixing in random walk models for dispersion in the sea. *Continental Shelf Research*, 18, pp.441-456.

[98] Ross, O. N. and Sharples, J. (2004). Recipe for 1-D Lagrangian particle tracking models in space-varying diffusivity. *Limnology and Oceanography Methods*, 2, pp.289-302.

[99] Rubinstein R. Y. (1981). Simulation and the Monte Carlo Method. John Wiley & Sons, New York.

[100] Rubinstein R.Y. and Melamed B. (1998). Modern Simulation and Modeling. John Wiley & Sons, New York.

[101] Sabelfeld K. and Shalimova I. (2001). Forward and Backward Lagrangian stochastic models of turbulent transport. *Monte Carlo methods and Applications*, 7(3-4), pp.369-383.

[102] Schoenmakers J.G.M. and Heemink A.W (1997). Fast valuation of financial derivatives. *Journal of Computational Finance*, 1, pp.47-62.

[103] Schoenmakers J.G.M, Heemink A.W., Ponnambalam K. and Kloeden P.E. (2002). Variance reduction for Monte Carlo Simulation of stochastic environmental models. *Applied Mathematical Modelling*, 26, pp.785–795.

[104] Scott C.F. (1997). Particle tracking simulation of pollutant discharges. *Journal of Environmental Engineering*, 123(9), pp.919-927.

[105] Shao Y. (1992). Turbulent dispersion in coastal atmospheric boundary layers: An application of a Lagrangian model. *Boundary-Layer Meteorology*, 59, pp.363-385.

[106] Silverman B.W. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall, London.

[107] Spivakovskaya D., Deleersnijder E. and Heemink A. W. (2006). Random walk model in case of iso- and diapycnal diffusion. *Proceedings of European Conference on Computational Fluid Dynamics ECCOMAS CFD 2006*, P. Wesseling, E. Oñate and J. Périaux (Eds).

[108] Spivakovskaya D., Deleersnijder E., Heemink A. W. (2007a) Lagrangian modelling of multi-dimensional advection-diffusion with space-varying diffusivities: theory and idealized test cases. *Ocean Dynamics*, 57(3), pp.189-203.

[109] Spivakovskaya D., Deleersnijder E. and Heemink A.W. (2007b) The backward Îto method for the Lagrangian simulation of transport processes with large space variations of the diffusivity. *Ocean Science Discussion*, 4, pp.623-654.

[110] Spivakovskaya D., Heemink A. W., Milstein G. N. and Schoenmakers J. G. M. (2004). Stochastic modeling of transport in coastal waters using forward and reverse time diffusion, in *Proceedings of Computational Methods in Water Resources XVI, Chapel Hill, NC, 13-17 June 2004, Edited by: C.T. Miller, M.W. Farthing, W.G. Gray, and G.F. Pinder*, Elsevier, Amsterdam.

[111] Spivakovskaya D., Heemink A. W., Milstein G. N. and Schoenmakers J. G. M. (2005). Simulation of the transport of particles in coastal waters using forward and reverse time diffusion, *Advances in Water Resources*, 28, pp.927-938.

[112] Spivakovskaya D., Heemink A. W. and Schoenmakers J. G. M. (2005b). A two particle model for the estimation of the mean and standard deviation of concentrations in coastal waters. *Proceedings of International Symposium of Stochastic Hydraulics 2005*, Nijmegen, the Netherlands.

[113] Spivakovskaya D., Heemink A. W. and Schoenmakers J. G. M.(2006). The risk analysis in coastal zone using reverse-time diffusion, *Proceedings of the XVI International Conference on Computational Methods in Water Resources, edited by P.J. Binning P.K. Engesgaard, H.K. Dahle, G.F. Pinder and W.G. Gray. Copenhagen, Denmark.* http://proceedings.cmwr-xvi.org/getFile.py/access?contribId=100&resId=0&materialId=paper&confId=a051

[114] Spivakovskaya D., Heemink A. W. and Schoenmakers J. G. M.(2007). Two-particle models for the estimation of the mean and standard deviation of concentrations in coastal waters, *Stochastic Environmental Resources and Risk Assessment*, 21(3), pp.235-253.

[115] Stelling G.S. (1983). On the Construction of Computational Methods for Shallow Water Flow Problems. Ph.D. thesis, Delft University of Technology.

[116] van Stijn Th.L., Praagman N., van Eijkeren J. (1987). Positive advection schemes for environmental studies, In: C. Taylor et al. (eds.), Numerical Methods in Laminar and Turbulent Flow, Pineridge Press, Swansea, pp.1256-1267.

[117] Stijnen J.W. (2002). Numerical Methods for Stochastic Environmental Models. Ph.D. thesis, Delft University of Technology.

[118] Stijnen J. W., Heemink A. W. and Lin H. X. (2006). An efficient 3D particle transport model for use in stratified flow. *International Journal for Numerical Methods in Fluids*, 51, pp.331-350.

[119] Suh S.-W. (2006). A hybrid approach to particle tracking and Eulerian-Lagrangian models in the simulation of coastal dispersion. *Environmental Modelling and Software*, 21, pp.234-242.

[120] Sun N. Z. (1999). A finite cell method for simulating the mass transport process in porous media. *Water Resources Research*, 35(12), pp.3649-3662.

[121] Talay D. and Tubaro L. (1991). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4), pp.485-509.

[122] Takeoka H. (1984). Fundamental concepts of exchange and transport time scales in a coastal sea. *Continental Shelf Research*, 3, pp.311-326.

[123] Taylor G.I. (1921). Diffusion by continuous moments. *Proceedings of the London Mathematical Society*, 2(20), pp.196-211.

[124] Taylor G.I. and Batchelor G.K. (1960). The scientific papers of Sir Geoffrey Ingram Taylor. Vol.2. Meteorology, oceanography and turbulent flow. Cambridge University Press.

[125] Thomson D. J. (1987). Criteria for the selection of stochastic models of particles trajectories in turbulent flow. *Journal of Fluid Mechanics*, 180, pp.529-556.

[126] Thomson D. J. (1990). A stochastic model for the motion of particle pairs in isotropic high-Reynolds-number turbulence, and its application to the problem of concentration variance. *Journal of Fluid Mechanics*, 210, pp.113-153.

[127] Thomson D. J., Physick, W. L. and Maryon, R. H. (1997). Treatment of Interfaces in Random Walk Dispersion Models. *Journal of Applied Meteorology*, 36, pp.1284-1295.

[128] Thuburn J.(1997). TVD Schemes, Positive Schemes, and the Universal Limiter. *Monthly weather Review*, 125, pp.1990-1993

[129] Tompson A.F.B. and Gelhar L.W. (1990). Numerical Simulation of solute transport in three-dimensional randomly heterogeneous porous media. *Water Resources Research*, 26(10), pp.2541-2562.

[130] Uffink G. J. M. (1988). Modeling of solute transport with the random walk method, In: Custodio et al. (eds), Groundwater Flow and Quality Modeling, D.Reidel Publishing Company, Dodrecht, pp.247-265.

[131] Uffink G.J.M. (1990) Analysis of Dispersion by Random Walk Method, Ph.D. thesis, Delft University of Technology.

[132] Zienkiewicz O. C., Taylor R. L. and Nithiarasu P. (2005). The Finite Element Method for Fluid Dynamics. Elsevier, Amsterdam.

[133] Zimmerman J. F. T. (1976). Mixing and flushing of tidal embrayments in the Western Dutch Wadden Sea, Part I: distribution of salinity and calculation of mixing time scales. *Netherlands Journal of Sea Research*, 10, pp. 149–191.

[134] Zimmerman J. F. T. (1988). Estuarine residence time, in: Kjerve, B. (ed.) Hydrodynamics of estuaries, 1, pp.75-84, CRC Press, Boca Raton.

[135] Zimmermann S., Koumoutsakos P. and Kinzelbach W. (2001). Simulation of pollutant transport using a particle method. *Journal of Computational Physics*, 173(1), pp.322-347.

[136] Yang Y., Wilson L. T., Makela M. E. and Marchetti M. A. (1998). Accuracy of numerical methods for solving the advection-diffusion equation as applied to spore and insect dispersal. *Ecological Modelling*, 109, pp.1-24.

[137] Yeh G. T. (1990). A Lagrangian-Eulerian method with zoomable hidden fine-mesh approach to solving advection-dispersion equations. *Water Resources Research*, 26(6), pp.1133-1144.

[138] Yeh G. T. (2000). Computational Subsurface Hydrology: Reactions, Transport, and Fate, Kluwer Academic Publisher, Dodrecht.

[139] van Dop H., Niewstadt F.T.M. and Hunt J.C.R. (1985) Random walk models for particle displacements in inhomogeneous unsteady turbulent flows. *Physics of Fluids*, 28(3), pp.1639-1653.

[140] van Kempen N.G. (1981). Ito versus Stratonovich. *Journal of Statistical Physics*, 24(1), pp.175-187.

[141] Visser, A. W. (1997). Using random walk models to simulate the vertical distribution of particles in a turbulent water column. *Marine Ecology Progress Series*, 158, pp.275-281.

[142] Vreugdenhil C. B. (1993). Linear central finite-difference method, In: Vreugdenhil and Koren (eds), Numerical Methods for Advection-Diffusion Problems, Vieweg.

[143] Wagner W. (1988). Monte-Carlo evaluation of functionals of solutions of stochastic differential equations. Variance reduction and numerical example. *Stochastic Ananlysis and Applications*, 6, pp.447-468.

[144] Wand M. P. and Jones M. C. (1995). Kernel Smoothing. Chapman and Hall, London.

[145] Wang H. F. and Anderson M. P. (1982). Introduction to Groundwater Modelling: Finite Difference and Finite Element, Academic Press, New York.

[146] Wilson J. D. and Sawford B. L. (1996). Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. *Boundary-Layer Meteorology*, 78, pp.191-210.

[147] Zheng C. (1990). MT3D: A Modular Three-Dimensional Transport Model for Simulation of Advection, Dispersion and Chemical Reaction of Contaminants in Groundwater Systems, Report to the U.S. Environmental Protection Agency, Ada, OK.

[148] Zheng C. and Bennett G. D. (2002). Applied Contaminant Transport Modeling, John Wiley & Sons, New-York.

[149] Zheng C. and Wang P. P. (1999). MT3DMS: A Modular Three-Dimensional Multispecies Transport Model for Simulation of Advection, Dispersion and Chemical Reactions of Contaminants in Groundwater Systems; Documentation and User's Guide, Contract Report SERDP-99-1, U.S. Army Engineer Research and Development Center, Vicksburg, MS.

Summary

Reverse-time diffusion in environmental models

Dar'ya Spivakovs'ka

The modelling of the transport phenomena is a complex problem, because many parameters influence the fate of the pollutant spilled into the sea or ocean. In general, researchers construct the model by taking into account the main factors affecting the pollutant transport. However, a good model should not be too simple; otherwise the prediction of the spreading of the pollutant will be inaccurate. A good model also shouldn't be too complex, otherwise it becomes computationally expensive. The main goals of this research were to modify the existing random walk models by taking into account more physical phenomena and to reduce the computational time by applying new numerical methods.

The random walk models are based on the theory of stochastic differential equations. One can interpret the advection-diffusion equation as a Fokker-Planck equation and derive the corresponding system of stochastic differential equations. This system can be used for modelling the movement of an individual particle of the pollutant. By generating the tracks of many particles the spread of the pollutant can be described. In Chapter 2 the necessary mathematical background can be found, while Chapter 3 contains a detailed description of the two- and three-dimensional random walk models.

In certain situations we are interested in estimating the pollutant concentration only at some critical locations. The standard methods based on the forward modelling are, in this case, not very effective, because the majority of particles will move far from the critical location and do not contribute to the final result. A possible solution is to use the reverse-time diffusion concept. By introducing the reverse-time variable the reverse-time system of the stochastic differential equations, consistent with the forward model, can be derived. Using this model, one can simulate the backward particle trajectories starting from the critical location. While the original forward model provides results where the pollutant will go to, the reverse system gives information where about the pollutant did come from. A straightforward application of the reverse-time concept is computing the risk of very high concentrations of pollutants. In Chapter 4 the reverse-time concept is applied to construct the risk map of the Dutch coastal zone.For these problems it is more efficient to use realizations of the reverse system instead of the original forward model. However, in general, the most efficient implementation is obtained if the forward realizations and the reverse system realizations are combined. This method is called the forward reverse estimator and has been recently introduced by Milstein, Schoenmakers and Spokoiny [85]. In Chapter 4 the forward-reverse estimator is introduced and applied to estimate particle concentrations in coastal waters. The results of the simulations show that the CPU time compared with the classical Monte Carlo method is reduced at least

order of magnitude.

One of the advantages of random particle models is that it is a natural way to study not only the mean ensemble concentration, but also higher order moments of the concentration. For instance, the standard deviation of the concentration is connected with the statistics of the trajectories of pairs of particles. In Chapter 5, two-particle models that take into account the space correlation of the turbulence are introduced and some properties of the distribution of the particle concentration are studied. The results of applying the two-particle model for several locations along the Dutch coast show that the actual concentration may become much higher than the ensemble mean concentration as computed by the traditional transport model. The computational time of the two-particle model can also be significantly reduced by using the forward-reverse estimator.

The particle models may be used for modelling the advection-diffusion process along and across density surfaces. However, in these cases, we need to deal with a diffusivity tensor containing off-diagonal elements. In Chapter 6 the Lagrangian model in the case of a space-varying diffusivity tensor is developed. This random walk model is applied for two idealized test cases for which the analytical solutions are known. Results of the tests show that the Lagrangian approach provides accurate and effective solutions of advection-diffusion problems for a general diffusivity tensors.

The random walk methods are applicable only when the flow velocities and diffusivity are sufficiently smooth functions. In practice, there are some regions where the rapid but continuous change in diffusivity have to be represented by a discontinuity. The random walk model based on backward Îto calculus (see [78]) can be used for these problems. The latter is best suited for the problems under consideration. It is applied in Chapter 7 for two test cases with discontinuous diffusivity, highlighting the advantages of this method.

Samenvatting

Reverse-time diffusie in transportmodellen voor milieuvervuiling

Dar'ya Spivakovs'ka

Het modelleren van transport verschijnselen is een moeilijk probleem omdat verschillende processen de verspreiding van een in een rivier of zee geloosde verontreiniging beïnvloeden. In het algemeen maken onderzoekers modellen die rekening houden met de belangrijkste factoren die de verspreiding van de verontreinigingen bepalen. Een model mag niet te eenvoudig zijn; dan wordt de voorspelling van de spreiding van de verontreiniging onnauwkeurig. Een goed model mag ook niet te ingewikkeld zijn, want dan wordt de rekentijd te lang. De belangrijkste doelen van dit onderzoek zijn om de bestaande random walk modellen aan te passen door meer rekening te houden met de verschillende fysische processen die het transport bepalen en om de rekentijd te verkorten door nieuwe numerieke methoden toe te passen.

Random walk modellen zijn gebaseerd op de theorie van stochastische differentiaalvergelijkingen. De advectie-diffusie vergelijking kan geïnterpreteerd worden als een Fokker-Planck-vergelijking en vervolgens kan het overeenkomstige systeem van stochastische differentiaalvergelijkingen worden afgeleid. Dit systeem kan worden gebruikt om de verplaatsing van elk individueel deeltje van de vervuiling te modelleren. Door het pad van veel deeltjes uit te rekenen kan de verspreiding van de verontreiniging worden beschreven. In Hoofdstuk 2 wordt de theorie van de random walk beschreven en Hoofdstuk 3 bevat een gedetailleerde beschrijving van 2- en 3 dimensionale random walk modellen.

In sommige situaties zijn we alleen geïnteresseerd in het schatten van de concentratie van een verontreiniging op bepaalde (cruciale) locaties. De standaard methoden gebaseerd op voorwaartse modellering zijn in dit geval niet erg geschikt; het grootste deel van de deeltjes zal zich ver van de cruciale locatie verplaatsen en daarom niet aan het uiteindelijke resultaat bijdragen. Het gebruik van het reverse-time diffusie concept is een mogelijke oplossing. Door het introduceren van de reverse-time tijd variabele kan een met het voorwaartse model consistent systeem van stochastische differentiaalvergelijkingen worden afgeleid. Met dit model kunnen de deeltjes banen in omgekeerde richting gesimuleerd worden, beginnend bij de cruciale locatie. Terwijl het oorspronkelijke voorwaartse model uitrekent waar de vervuiling naar toe gaat geeft het omgekeerde systeem informatie over waar de stof vandaan komt. Een voor de hand liggende toepassing van het omgekeerdetijd concept is het uitrekenen van het risico van zeer hoge concentraties van vervuilende stoffen. In Hoofdstuk 4 wordt het reverse-time concept toegepast om een risico kaart van de Nederlandse kust te maken. Voor deze toepassingen is het efficiënter om realisaties

van het reverse-time systeem te gebruiken dan van het oorspronkelijke voorwaartse model. In het algemeen levert echter een combinatie van de voorwaartse realisaties en reverse-time systeem realisaties de meest efficiënte implementatie op. Deze methode wordt de forward reverse estimator genoemd en is recentelijk geïntroduceerd door Milstein et al. In Hoofdstuk 4 wordt de forward reverse estimator geïntroduceerd en toegepast om deeltjes concentraties in kustwateren te schatten. De resultaten van de simulaties laten zien dat de CPU tijd vergeleken met de traditionele Monte-Carlo methode met tenminste één orde-grootte wordt verminderd.

Één van de voordelen van deeltjes modellen is dat ze een natuurlijke methode zijn om zowel de ensemble-gemiddelde concentratie als hogere orde momenten van de concentratie te bestuderen. De standaardafwijking van de deeltjes concentratie hangt bijvoorbeeld samen met de relevante statistische eigenschappen van de paden van deeltjes paren . In Hoofdstuk 5 worden twee-deeltjes modellen geïntroduceerd die rekening houden met de ruimtelijke correlatie van turbulentie. Ook worden enige eigenschappen van de verdeling van de deeltjes concentratie bestudeerd. De resultaten laten zien dat de werkelijke concentratie op verschillende plaatsen voor de Nederlandse kust, uitgerekend met het twee-deeltjes model, veel hoger kan uitvallen dan de ensemble-gemiddelde concentratie voorspeld met het traditionele transport model. De rekentijd van het twee-deeltjes model kan ook sterk worden verminderd door toepassing van de forward reverse estimator.

De deeltjes modellen kunnen worden gebruikt om het advectie-diffusie proces langs en door dichtheidsoppervlakken te modelleren. In zulke gevallen moet echter rekening worden gehouden met een diffusiviteitstensor die niet diagonaal is. In Hoofdstuk 6 wordt het geval behandeld dat de diffusiviteitstensor niet diagonaal is en bovendien van de ruimte variabele afhangt. Het resulterende random walk model wordt toegepast op twee geïdealiseerde test situaties waarvoor analytische oplossingen bekend zijn. De test resultaten laten zien dat de Lagrangiaanse aanpak nauwkeurige en efficiënte oplossingen voor advectie-diffusie problemen met een willekeurige diffusiviteitstensor geeft. De random walk modellen zijn alleen toepasbaar als de stroomsnelheid en de diffusiviteit voldoende gladde functies zijn. In de praktijk zijn er echter gebieden waar snelle maar continue veranderingen in de diffusiviteit moeten worden beschreven met een discontinuïteit. Hiervoor kan het op backward Îto-calculus gebaseerde random walk model worden gebruikt. In Hoofdstuk 7 wordt deze techniek toegepast op twee test situaties met een discontinue diffusiviteit; in dit hoofdstuk worden ook de voordelen van deze methode gedeminstreerd.

Acknowledgments

Graham Greene said "to each man a city consists of no more than a few streets, a few houses, a few people". First of all, Delft for me is people whom I met here and who became real friends for me. As long as my friends live there I will think about Delft as my home.

I would like to express my sincere gratitude to my supervisor, Prof. A.W. Heemink for his support, help and guidance. He was always open to my ideas and encouraged me throughout my research. I am grateful for his kindness, patience and his efforts to understand my Russian English, which was not an easy task during the first year.

I am very grateful to Prof. E. Deleersnijder for his advice and very stimulating discussions. His valuable suggestions and constant enthusiasm kept me highly motivated. His advice to avoid "the black magic" in formulation of the mathematical concepts helped me to make the papers more readable. It was a great pleasure to work with him. I am very grateful to Dr. J.G.M. Schoenmakers for his help in implementing the forward-reverse estimator, which I adapted for my needs. I am grateful to a member of the defence committee, Prof. F.M. Dekking, for his valuable comments and remarks, which helped me to improve the thesis.

There are several people, who made my study in TU Delft possible. I am very grateful to Prof. I.V. Andrianov for helping me to come to Delft and also for his support during my studies. I am very grateful to Dr. W. van Horssen not only for organizing and arranging many administrative matters, but also for conversations and jokes. Keeping in mind his advice "You will have to pay taxes under any government" I don't take political changes in my country very much to heart. I would like to thank Evelyne Sharabi for organizing and helping with various matters. I thank to the CICAT project assistants Martine van der Laag-Hoogendijk, Manon Post and Theda Olsder; Manon, especially, for her kindness and friendship.

I spent five splendid and very interesting years at the Mathematical Physics group and I wish to thank people who work there. I am very grateful to Kees Lemmens for his help and support in studying programming languages and software and for his patience in answering all my questions and concerns. Thanks to him and Dr. H.X. Lin I was introduced to the subject of parallel processing, which helped to speed up my work. I am grateful to Eef Hartman, Henrick Corstens, Marleen Keijzer, Pieter Wilders and all my colleagues who helped me a lot with my daily life. I am grateful to Wilson Charles and Fahmi Naifar for the valuable and interesting scientific discussions. During five years at the TU Delft I had many roommates, who helped to create a pleasant and a very

international atmosphere in our office. I thank Auke for introducing me to the world of classical music, Laura for teaching me basic Italian, S. Xu for discussions about our common experience in an upbringing at communist society. I also thank to Safina and A. Zamani Foroushani.

My life in Delft would not have been so full and interesting without my friends with whom I shared all my ups and downs. First of all, I would like to thank my closest friends Maria, Andrey, Yulia and Marnix. We spent so much time together that I start thinking about us as a family. I miss you very much and I hope that the distance between us will not destroy our friendship. I am grateful to Alex Andrianov for his help and advice and to Alina for agreeable conversations and support. I am grateful to Xander, who helped me greatly with the Dutch language and to Dwi for her optimism, support and, of course, badminton lessons. Many thanks go to Olga and William for their hospitality and kindness. I wish to thank Ayse for her help, advice and very interesting conversations and discussions.

The group of friends from Delft would not be complete without mentioning Svetlana and Wies, Yulya and Andrey, Remus, Julius, Alex Chernetskyy, Alex Peschanskyy, Saad, Jelle, Nils, Timofey and Anna. I am grateful for the good time, that we spent together, for parties, laughs, going to the cinema, watching football and listening to the "gracht concerts".

I am very grateful to my high school teacher L.M. Spirkina, who showed me a wonderful world of mathematics and inspired to choose this field as a career path. I am grateful to Prof. V.P. Motorny, Prof. V.N. Turchin and Dr. S.A. Pichugov from Dnipropetrovsk National University. My special thanks go to my first scientific supervisor Dr. V.L. Velikin for his support, guidance and advice.

I am infinitely grateful to my family for their love, support and understanding, and for everything that they are doing for me. I am grateful to Oleg for his help regarding computers, to Vika and Timur, who allow me to return to the childhood. I am grateful to my sister Lena for her support, advice and sense of humor. When I was six, she told me that when I became a professor, she would listen to me. Probably, since that time I started thinking of being a scientist. And, most of all, I am grateful to my parents for their love and support. Дорогая мама, я хотела бы найти слова, которые могли бы выразить мою любовь и благодарность к тебе. В течении этих лет, я всегда знала, что есть люди, которые верят в меня и поддержать меня в любом моем начинании, люди которые любят меня такой, какая я есть и которые гордятся моими успехами и переживают за мои неудачи. Спасибо!

Curriculum vitae

Dar'ya Spivakovs'ka (Daria Spivakovskaya according to Russian spelling) was born on 28th October 1975 in Dnipropetrovsk, Ukraine, former USSR. From 1983 she attended the primary and secondary school N23 in Dnipropetrovsk, Ukraine. She finished the school in 1993.

In the same year she started her study at the Department of Functional Analysis, the Faculty of Mechanics and Mathematics, Dnipropetrovsk State University (now Dnipropetrovsk National University). In 1998 she defened her Master thesis entitled "On some properies of B-splines" under supervision of Dr. V.L. Velikin and obtained her Master degree (with honors). After graduation she joined as a staff member of the Department of Mechanics and Mathematics at the Dnipropetrovsk National University. She continued the research work there till February 2002.

From February 2002 Daria Spivakovskaya started a Ph.D. research program at the Department of Applied Mathematics (nowdays Delft Institute of Applied Mathematics), Delft University of Technology, The Netherlands. She worked on the project connected with stochastic modelling of the transport processes under supervision of Prof.dr.ir. A.W.Heemink. The results she obtained during next five years form the content of the current thesis entitled "Reverse-time diffusion in environmental models".

Since May 2007 she works as a petroleum engineer at A/S Norske Shell, Norway.